

8th International Conference Computer Ethics: Philosophical Enquiry

Ionian University:
Department of Archive and Library Sciences
Department of Informatics
INSEIT

Ionian Academy
Ionian University, Corfu, Greece June 26-28, 2009

8th International Conference Computer Ethics: Philosophical Enquiry
ISBN 978-960-272-654-9

ΝΟΜΙΚΗ ΒΙΒΛΙΟΘΗΚΗ



Μαυρομυχάλη 23, 106 80 Αθήνα
Τηλ.: 210 3678 800 • Fax: 210 3678 819
<http://www.nb.org> • e-mail: info@nb.org

Αθήνα: Μαυρομυχάλη 2, 106 79 • Τηλ.: 210 3607 521
Πειραιάς: Φίλωνος 107-109, 185 36 • Τηλ: 210 4184 212
Πάτρα: Κανάρη 28-30, 262 22 • Τηλ.: 2610 361 600
Θεσ/νίκη: Φράγκων 1, 546 26 • Τηλ.: 2310 532 134

Σύμφωνα με το Ν. 2121/93 για την Πνευματική Ιδιοκτησία απαγορεύεται η αναδημοσίευση και γενικά η αναπαραγωγή του παρόντος έργου, η αποθήκευσή του σε βάση δεδομένων, η αναμετάδοσή του σε ηλεκτρονική ή οποιαδήποτε άλλη μορφή και η φωτοανατύπωσή του με οποιονδήποτε τρόπο, χωρίς γραπτή άδεια του εκδότη.

ΔΗΛΩΣΗ ΕΚΔΟΤΙΚΟΥ ΟΙΚΟΥ

Το περιεχόμενο του παρόντος έργου έχει τύχει επιμελούς και αναλυτικής επιστημονικής επεξεργασίας. Ο εκδοτικός οίκος και οι συντάκτες δεν παρέχουν διά του παρόντος νομικές συμβουλές ή παρεμφερείς συμβουλευτικές υπηρεσίες, ουδεμία δε ευθύνη φέρουν για τυχόν ζημία τρίτου λόγω ενέργειας ή παράλειψης που βασίστηκε εν όλω ή εν μέρει στο περιεχόμενο του παρόντος έργου.

Art Director: Θεόδωρος Μαστρογιάννης
Υπεύθυνος Παραγωγής: Ανδρέας Μενούκος
Φωτοστοιχειοθεσία: Αριστέα Διακουμοπούλου
Παραγωγή: NB Production AM051109M23

NOMIKI BIBLIOTHIKI GROUP

23, Mavromichali Str., 106 80 Athens Greece
Tel.: +30 210 3678 800 • Fax: +30 210 3678 819
<http://www.nb.org> • e-mail: info@nb.org

All rights reserved. No part of this publication may be reproduced by any means, without prior permission of the publisher.



8th International Conference Computer Ethics: Philosophical Enquiry

**Ionian University:
Department of Archive and Library Sciences
Department of Informatics
INSEIT**

**Ionian Academy
Ionian University, Corfu, Greece June 26-28, 2009**

**Editor
Maria Bottis**



NOMIKI BIBLIOTHIKI 2009

CONTENTS

Forword	xiii
<i>Herman T. Tavani</i>	
The need for a new social mentality, tools and business practices in the modern information environment	1
<i>George Bokos</i>	
Philosophy and the Information Revolution	30
<i>Terry Bynum</i>	
Paradoxes of contemporary knowledge: between invention, creation, information and control	53
<i>Laymert Garcia Dos Santos, Konstantinos Karachalios and Jürgen Partenheimer</i>	
Classic yet Contemporary Gender Norm: is ICT an amplifier of gender bias?	67
<i>Ryoko Asai</i>	
Ethics at the Crossroads of Bioinformatics and Nanotechnology	79
<i>Athanasios Alexiou & Panayotis Vlamos</i>	
Modeling Aspects of Action Theory	96
<i>Saban Al-Fedaghi</i>	
A frame for ethical evaluation of (information) technologies	110
<i>Josep M. Basart & Montse Serra</i>	
The Greek scale of attitudes towards unethical behaviors on the Internet	119
<i>George Briskolas & Petros Roussos</i>	
Computer-mediated Friendship: Illustrating Three Tasks for a Computer Ethics of the Good	135
<i>Adam Briggie</i>	
Internet Research Ethics: Reports from US-Based Institutional Review Boards	148
<i>Elizabeth A. Buchanan & Charles Ess</i>	

Free software, economic “realities”, and information justice	157
<i>Samir Chopra & Scott Dexter</i>	
Living in the Eye of the Artificial Other Artificial Intelligence, Moral Imagination, and Disciplining.....	177
<i>Mark Coeckelbergh</i>	
The ethics of ambient computing for personal health monitoring.....	186
<i>Goran Collste</i>	
Trusting Invisible Strangers in Open Source Communities: About the Assumption, Inference and Substitution of Trust	208
<i>Paul B. de Laat</i>	
Ethical aspects of e-Government: Social Actors, Politics and the Digital Divide.....	234
<i>Georgia Foteinou & George Pavlidis</i>	
The Great Catopticon	252
<i>Jean-Gabriel Ganascia</i>	
Virtual Decisions: Just Consequentialism, Video game ethics, and Ethics on the fly	267
<i>Don Gotterbarn & James Moor</i>	
Developing artificial agents worthy of trust: “Would you buy a used car from this artificial agent?”	288
<i>Frances S. Grodzinsky, Keith W. Miller & Marty Wolf</i>	
Can the “Contextual Integrity” Model of Privacy Be Applied to Personal Blogs in the Blogosphere?	303
<i>Frances S. Grodzinsky & Herman T. Tavani</i>	
ICT and financial services: learning lessons from the environmentalists	312
<i>Mike Healy & Ben Fairweather</i>	
Moral Luck and Computer Ethics: Gaugin at the keyboard.....	326
<i>David Sanford Horner</i>	
Intellectual Property issues for digital libraries in the Internet networked public sphere	344
<i>Dionysia Kallinikou, Marinos Papadopoulos, Alexandra Kaponi & Vassiliki Strakantouna</i>	

The Mirror and the Spell: Order and Organization in the Age of the Digital Reserve	386
<i>Hannah Knox, Damian O' Doherty, Theo Vurdubakis & Chris Westrup</i>	
A brave new world for libraries: the sui generis right.....	413
<i>Rania Konsta & Manolis Gerolimos</i>	
Computerized support for ethical analysis.....	425
<i>Mikael Laaksoharju & Iordanis Kavathatzopoulos</i>	
Is the spammer evil?	438
<i>Eleftherios Loukis & John Papadakis</i>	
A Debordian analysis of Facebook	454
<i>Antonio Marturano & Sergio Bellucci</i>	
The Commodification of the Individual in the Internet Era: Informational Self-determination or "Self-alienation"?.....	466
<i>Lilian Mitrou</i>	
The Joy of Excellence: Intellectual Property and Hackers' Virtue.....	485
<i>Scott J. Molony</i>	
From Public Data to Private Information: The Case of the Supermarket	500
<i>Vincent C. Müller</i>	
Information ethics in business organisations	508
<i>Kyoshi Murata</i>	
Blogs and privacy in Seken as a Japanese life-world including indigenous moral norms	518
<i>Makoto Nakada</i>	
The counter-control revolution: silent control over individuals with dataveillance systems.....	543
<i>Yohko Orito</i>	
Being a 'Step Parent' to an 'Orphan' Work: Problems and Solutions Regarding the 'Adoption' of 'Orphans'.....	553
<i>Maria-Daphne Papadopoulou</i>	
Default Social Meaning: An Ethics of Video Game Play	591
<i>Stephanie Patridge</i>	

Ethical Issues in the Design and Use of Online Career Development Counseling Services.....	611
<i>Adamantia Pateli</i>	
Machine Self-Sacrifice.....	625
<i>Carson J. Reynolds & Alvaro Cassinelli</i>	
Mixed Feelings Computing.....	636
<i>Nikola Serbedžija</i>	
Ethical Issues of Pervasive Computing in the Insurance Industry	647
<i>Oliver Siemoneit</i>	
MyChoice & Traffic Lights of Trustworthiness: Where Epistemology Meets Ethics in Developing Tools for Empowerment and Reflexivity	655
<i>Judith Simon</i>	
The Rights and Duties of Lawful User in EU Copyright Law.....	671
<i>Tatiana-Eleni Sinodinou</i>	
Information, Knowledge and Wisdom: Groundwork for the Evaluation of Digital Information and Its Relation to the Good Life	694
<i>Edward H. Spence</i>	
Landscapes of Ethical Issues of Emerging ICT Applications in Europe	719
<i>Bernd C. Stahl & Simon Rogerson</i>	
Security, Privacy and Technophobia in the School of the Future	738
<i>Giannis Stamatellos</i>	
The role of Internet Service Providers. Ethics, Reality and the Law: The Example of Promusicae v. Telefonica.....	750
<i>Irini Stamatoudi</i>	
Virtually Good – What Can We Learn from the Argument from False Pleasures?	767
<i>Johnny Hartz Søraker</i>	
Remixing and Recoding: Revisiting the Copyright Wars	778
<i>Richard A. Spinello</i>	

The Digital Divide among Under-Age Individuals: An Economic and Legal Approach	794
<i>Stiakakis Emmanouil & Alexandropoulou-Egyptiadou Evgenia</i>	
Situational Crime Prevention and Insider Threat: Countermeasures and Ethical Considerations	808
<i>Marianthi Theoharidou & Dimitris Gritzalis</i>	
Everybody wins: challenges and promises of knowledge discovery through volunteer computing.....	821
<i>Lauri Tuovinen & Juha Rönning</i>	
Information Systems and Future Ethics: A Reflection of Students' Current Behaviour.....	843
<i>Aharon Yadin</i>	
Curia Novit Google: The Quest for Law in the Era of Internet Search Engines.....	853
<i>Georgios N. Yannopoulos</i>	
Constraints to the Application of ICT Implants: The Concept of Self-Ownership.....	862
<i>Karsten Weber</i>	
What Should We Share? Understanding the Aim of Intercultural Information Ethics.....	873
<i>Pak Hang Wong</i>	
"But the Data is Already Public": On the Ethics of Research in Facebook.....	885
<i>Michael Zimmer</i>	

Foreword

The papers included in the *Proceedings of the Eighth International Conference on Computer Ethics – Philosophical Enquiry: CEPE 2009* were originally presented during a three day-period – June 26-28, 2009 – at the Ionian University, Corfu, Greece. These papers were selected, via a systematic peer-review process, from the numerous papers received by the CEPE 2009 Program Committee.

Maria Bottis, a distinguished information law scholar and CEPE 2009 Conference Director, has done an outstanding job of bringing these papers together in a single volume. In editing a high-quality, juried proceedings of conference papers, she has also helped to preserve, and improve upon, an important CEPE tradition that began with the publication of the proceedings of papers originally presented at the first CEPE Conference (Erasmus University – Rotterdam, The Netherlands in June 1997). Many papers published in the seven previous CEPE Proceedings were also later included (in revised form) in special issues of *Ethics and Information Technology* and *Computers and Society* that were based on various CEPE conference themes. Other papers have subsequently appeared in journals such as *The Information Society*, *Minds and Machines*, and *Metaphilosophy*. Some of the papers included in the present volume will also be published later, in revised form, in special issues of journals and periodicals.

The CEPE 2009 Conference Proceedings is the largest volume of CEPE papers, to date. It includes 57 papers authored by computer-ethics scholars from around the world, thus perpetuating the truly international spirit of the CEPE conference series.

The success of the recent CEPE 2009 Conference in Corfu has also helped to reinforce the positive relationship that has developed over the years between the CEPE conference series (in general) and its sponsoring organization, INSEIT (International Society for Ethics and Information Technology). For this, INSEIT is extremely grateful to Maria Bottis and to her colleagues at the Ionian University.

Herman T. Tavani
President, INSEIT

The need for a new social mentality, tools and business practices in the modern information environment

George Bokos*

Professor, Department of Archive and Library Science,
Ionian University

Although for more than three decades now the environment of the production, publication, dissemination and use of information and information content and products (or better information carriers and media) has been under constant and continuously accelerating change, the Library and Information Services sector, like other, related or not, sectors of the modern social environment, seems rather unable to fully understand that such a change is not a simple evolution of media and practices, but a total and complete change of the operational environment and of the perspectives of this profession. I would like to quote here the eloquent metaphor used by Jason Epstein in his keynote speech in TOC 2009 Conference¹. According to this metaphor, “Like blind men in a room with an elephant, we cannot begin to imagine the eventual consequences as digitization and the Internet ignite a worldwide Cultural Revolution, orders of magnitude greater than Gutenberg’s inadvertent implementation of western civilization”. This is of course a problem, since the way we perceive a situation determines the way we react in order to cope with it. I am afraid that this is a more serious problem in the case of extreme and rapid technological and eventually social changes and this is the situation today.

* Professor *George Bokos*, former Head of the Department of Archive and Library Science, Ionian University, Greece, teaches Library Automation and New Information Technologies, while he is the Director of the Information Technologies Laboratory of this Department. He served as the Head, Cataloguing & Bibliographic Services Dept., National Library of Greece, 1988-1995 and as the Director of the National Library of Greece, 1996-1997. Professor Bokos has been a member of several technical committees and working groups, both at the national and international level and on subjects related mainly to the automated handling and management of bibliographical information. He has served as a national representative of Greece in the Telematics for Libraries Programme of the EU. Pr. Bokos has been a project manager and coordinator of several national and international research projects. He has been consulting libraries and other information services for many years on library automation and automated management of information. He is the author of several works and has participated in many conferences on the above subjects. His research interests include library automation and information technologies, electronic publishing, technologies and standards for data and content encoding.

1. Speech given by Jason Epstein at the 2009 O'Reilly Tools Of Change for Publishing Conference, viewed 3 June 2009, <<http://toc.oreilly.com/2009/02/full-text-of-jason-epsteins-to.htm>>1.

It should be pointed out here from the beginning that, although I come and, in a way, I represent here the Library and Information Services sector, I am using this sector, in the context of this presentation, rather as an example of a really peculiar situation, which is, more or less, of interest to many other fields of social activity or, better, to the modern society as a whole. This last claim is supported, I think, among other things, by the variety of topics covered the last three days in the context of this conference. This peculiarity is mainly related to the fact that the society, confronted with significant changes in the technological infrastructure of its activities, seems, up to this moment, to believe that it understands the quality, quantity and extent of the changes taking place and that it can cope with these changes with the suitable, mainly technological in nature, measures (like, for example, new software, new hardware, new or adapted legislation and practices e.t.c.). I am afraid that, to a certain extent, the modern society, confronted with the digital revolution, resembles the aforementioned metaphor of the blind people in a room, attempting to find out what is the newcomer by touching parts of his body, smelling it or listening to noises coming from it.

The problem is, as far as we can judge from our current knowledge and experience of the field, that, with measures of only this kind, the whole situation seems to become more and more perplexed and instead of solving problems, the new measures rather increase the complexity of the existing ones or create new of them. It seems, to put it in another way, that the modern society, ignoring the real nature of the elephant in the room and supposing that is something known and familiar, attempts to adopt and use the technology, disregarding its full potential, its consequences in both the nature and the production, management, dissemination and use of information and content and its repercussions in the mentality, expectations, habits and behaviour of the end user. The society tries, thus, to cope with a completely new and radically different from the past situation with its traditional tools and mentality, that is, with absolutely inappropriate for the specific task tools.

This peculiarity is the cause of a rather serious social problem, which can be briefly described as the inability of the society: a) to cope successfully with the problems of the rapid change and b) to fully exploit the potential of the new environment. This presentation will attempt to give some evidence indicating that the problem lies mainly to this inability of the modern society to understand, correctly appraise and, mainly, accept some essential characteristics of the changes under way. The first is that these changes were and are of an extremely rapid and radical nature and, thus, the changes in technology and tools resulted in a subsequent and equally radical change in mentality, that is, in our case, in the way we perceive information and content. The second is that these changes concern and influence a very wide range of social activities, or better the social activities as a

whole. The third, and perhaps the most important, is the fact that these changes had a significant impact on the user of information. In fact an absolutely new kind of information user is under formation. Talking about the user of information, we should not disregard the fact that a new kind of information producer is also under formation and, probably most important, that these roles, as well as other discrete roles in the traditional information environment can be easily and widely interchangeable today. And now to our example case for today, the Library and Information Services sector.

It is undeniable of course that there are a lot of references in the relevant bibliography mentioning terms like “revolution” or prophecies like “the end of libraries²”, the “death of the printed book³”, e.tc. Irrespective of the fact that such references were a characteristic of the first period of these radical changes and reflected the strong impact these changes had on the professionals of the time, we can be sure today that they were not entirely true or correct. In fact some of the terms, like revolution, were more than suitable for describing the changes of the last decades, although I am not sure whether there is, even today, a consensus regarding the real content of this “revolution”. The prophecies of course are in general more difficult to be equally successful and as a result we cannot be sure today, whether, for example, the printed book or the conventional libraries will die soon.

It is true that the Library and Information Services profession had the opportunity of an early introduction to this kind of changes, since libraries were among the first social organizations to adopt the new information technologies, starting to use computers in their work as early as the decade of the ‘60s⁴. This early involvement helped the field to proceed gradually and to absorb easily the subsequent rapid and radical changes that were in fact triggered by the convergence and the combined impact of the computing and telecommunications technologies. The digital content, the Internet and the World Wide Web are the more obvious, from a social point of view, products of this new situation. We should confess that the Library and Information Services sector, experienced and suitably educated, adopted easily and started to exploit very effectively, the tools offered by the new and advanced information technologies.

-
2. Although an attempt to help libraries, the paper by James Thompson, “The End of Libraries”, *Electronic Library*, v1 n4 p245-55 Oct 1983, is characteristic of the time.
 3. Sven Birkerts, *The Gutenberg Elegies: The Fate of Reading in an Electronic Age*, Faber and Faber, 1994, fears the eventual death of the book, a fact that will result in the subsequent death of many cultural values connected with the printed text. Although more optimistic, with regard to the impact of digital publishing, Richard A. Lanham. *The Electronic Word: Democracy, Technology and the Arts*, University of Chicago Press, 1993, expects also the end of the printed text.
 4. The personal memories of Maurice B. Line, “Forty years of library automation: a personal reflection”, *Program: Electronic Library and Information Systems*, 40, 2 (2006) 118-122, is a really good, brief and interesting account of the history of automation in libraries.

In fact the whole field seems, for more than a decade now, to struggle desperately and continuously, to incorporate in its relevant activities, as soon as possible, any new ability offered by the available technological environment. This is apparent in the gradual evolution of the Library Automation systems, which among other things, have been fully transformed to Web-based and Web oriented applications, and in the appearance of new fully Web based tools and applications⁵. We should also confess that the traditional libraries, empowered by these new technologies, became much more effective regarding their usual tasks in managing content and serving the user.

From this point of view, very briefly described here, the Library and Information Services sector is really successful in adopting and incorporating the new technological environment. The problem however lies not in the adoption of the technology, but in the adjustment of the Library and Information Services sector to the new environment. However, and as it has very successfully described, "Not since the age of Gutenberg has an information upheaval so thoroughly disrupted the processes of scholarly knowledge creation, management and preservation as the digital revolution currently under way"⁶. The whole thing can be said in a simpler way: the libraries today, using up to day technological infrastructure, continue to use their old traditional practices and service models. It is well known that, especially in our times, technology changes quickly, while people and institutions change slowly. That is because the way of thinking and perceiving the environment, the beliefs and the expectations, the mentality and the ethical standards of society, as these are manifested, and expressed by both the combined behaviour of the individuals and the official legal provisions and structure of society, they usually change at too low a pace. The society, thus, and the individual are usually lag behind, especially in times and cases of rapid and radical changes of the environment. Up to this moment the institutions in question, I mean the libraries and the information services in general, equipped with all the new tools and capabilities of the digital environment, remain, from a functional point of view, almost the same. I mean remain the same from the point of view, among other things, of the way they perceive their current operational environment and their respective role, the professional practices used, and the professional deontology, principles and ethics that determine their relations with their users, their content, their content providers and, finally, with other professionals, professional fields and the society as a whole.

5. Marshall Breeding, «Investing in the Future: Automation Marketplace 2009», *Library Journal*, 2009, Document available on the Web, URL: <http://www.libraryjournal.com/article/CA6645868.html>>, viewed 12 June 2009.

6. A very brief, succinct and successful description of the situation in the abstract for the article of Isaac Hunter Dunlap, "Going Digital: The transformation of Scholarly Communication and Academic Libraries", *Policy Futures in Education*, 6, 1, (2008) 132-141.

Of course, as already mentioned, adaptations of this kind are not an easy task and, above all, require time. However, it is, first of all, necessary for the Library and Information Services sector to fully understand the crucial characteristics of its new operational environment and subsequently to try to adapt to it, if libraries and similar information services are to continue to play an important role in the handling and management of information in the future. At least if they are to play a role really demanded by the society and its real needs. The interesting point here is that this is not really or only a problem of the libraries, but a problem of the society as a whole, which needs desperately such a professional tool, if we are to survive from the current, rather out of control, flood of information. I should mention again here that the problem described with regard to libraries is in fact a problem of many others social activities which have to do in one or the other way with information and content.

I will try, first, to point out very briefly what are, in my opinion, the most important features of the new operational environment of libraries. I mean here the features that constitute the radical change with regard to the conventional environment and which demand an analogous change on the part of the libraries. This parallel environment, which tends to create in fact a whole new parallel world, has succinctly described many years ago by Nicholas Negrepon⁷, as the antithesis between “bits” and “atoms”, meaning in fact the parallel existence of two worlds, a material and an immaterial one. This means, for example, that you can use today a traditional library or a source of information and content available in the Web. You can visit your local branch of a bank, or you can use the available Web banking. You can visit your local store of any kind, or you can go for Web shopping. There are many similar examples, but I think that the existence and use of a parallel social environment is obvious. These, of course, are not features related only to libraries, but features regarding the society as a whole, since, in any case, this is the operational environment not only of the Library and Information Services, but of, more or less, any other social sector. My purpose here is to show that the main key to successfully cope with the new situation is the need for an in depth understanding of the quantity and, mainly, quality, of the changes taking place, and, consequently, the need for a completely new perspective regarding the relations of the Library and Information Services sector with the rest of the society and the service models to be used in this new context.

A full analysis of the main features that characterize the modern operational environment of the Library and Information Services sector is not an easy task and of course is beyond the scope of this presentation. What I shall try, thus, to give here is some of these features, hopefully the most interesting for our purpose, described in terms of the status and trends of the Technology, the nature of the Con-

7. Nicholas P. Negrepon⁷, *Being Digital*, Vintage Books, 1996.

tent, the characteristics and attitudes of the User and the emerging new Mentality with regard to information production and use. These features may be briefly summarized as follows:

The key-points of the Technology first:

1. Any relevant discussion should start from the appearance of the digital form of information, combined with continuously altered and expanding production, publication, transmission and dissemination capabilities, as a result of advances in computing and telecommunications technologies. This feature alone means a tremendous change for people or services, like librarians and libraries, which are engaged in managing, organizing and supplying information and content to the end user. It is worth mentioning here the observation of the Librarian of Congress Dr. James H. Billington, in his Statement before the House Subcommittee on Legislative Branch, U.S. House of Representatives, March 20, 2007. In that statement Dr Billington observed that “It took two centuries for the Library of Congress to acquire today’s analogue collection—32 million printed volumes, 12.5 million photographs, 59.5 million manuscripts and other materials – a total of more than 134 million physical items. By contrast, with the explosion of digital information, it now takes only about 15 minutes for the world to produce an equivalent amount of information”⁸. This contrast cannot be overemphasized. It is a tremendous quantity change with severe quality repercussions in the attitudes of our society toward information production, dissemination, management and use. It is also a change that would, by itself and without taking into account any other change in the field, demand a complete reconsideration of the modern library’s business model and practice.
2. It should, also, be added here, as a second point, that this amount of information is created, published, transferred, disseminated and used in a completely new way, with a large set of new channels and media, with completely different standards regarding time and methods of access e.tc. I mean of course the new environment for the production, dissemination and use of information, provided by the Internet and the Web. All these changes had resulted in a gradual but nonetheless rapid and radical change of mentality, habits, behaviour and practices as concerns both the attitudes of people and the methods used for publishing and acquiring information and content. It is obvious, thus, that libraries, given this radical qualitative and quantita-

8. Testimony to Congress. Statement of Dr. James H. Billington The Librarian of Congress before the House Subcommittee on Legislative Branch U.S. House of Representatives, March 20, 2007. Document available on the Web, URL: <http://www.loc.gov/about/librarianoffice/speeches/032007.html>>, viewed 6 April 2009.

tive change of their main object of interest, the content and the collections, should adapt appropriately their attitudes and practices with regard to collection development and management. As I will try to show with some examples in the following, it is not enough for the libraries to adopt the technology and to try to improve their traditional services. What they should do is to adapt their objectives and transform their structure and services so that they conform to what is, wants and demands the user today.

3. The third point is the fact that computing and telecommunications capabilities are now becoming steadily an everyday tool for most people in our society. Although the discussion about technology may seem a banality today, we should not disregard the fact that the extent of the technology's penetration with regard to population and social activities and services will be the decisive factor concerning the final prevalence of the emerging new social practices, service models and mentality. The brief description of the current level of development and use of ICT is, in general, not an easy task, due to the fact that the field itself is both a complicated and a rapidly changing one. It is something more difficult, of course, in the context of this presentation. I will try however to give some evidence here, insisting mainly in the relevant trends, which are far more important from our point of view. Since we cannot describe the whole field, I will try to give an idea of the situation in terms of computers, internet, broadband and mobile devices penetration and use. This can be done here by simply supplying a selected set of charts and graphs, regarding the trends in the aforementioned fields.

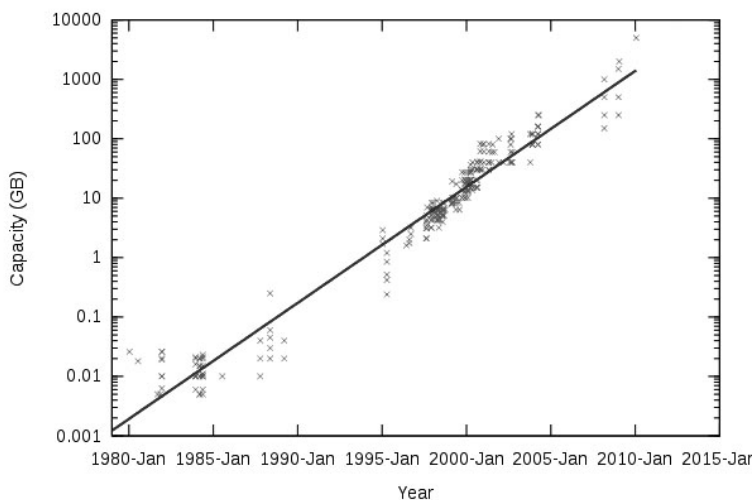


Fig. 1. Evolution of Hard Drive Capacity over Time 1980 – 2015 (Chart available on the Web: Wikimedia Commons, URL: < http://commons.wikimedia.org/wiki/File:Hard_drive_capacity_over_time.svg >)

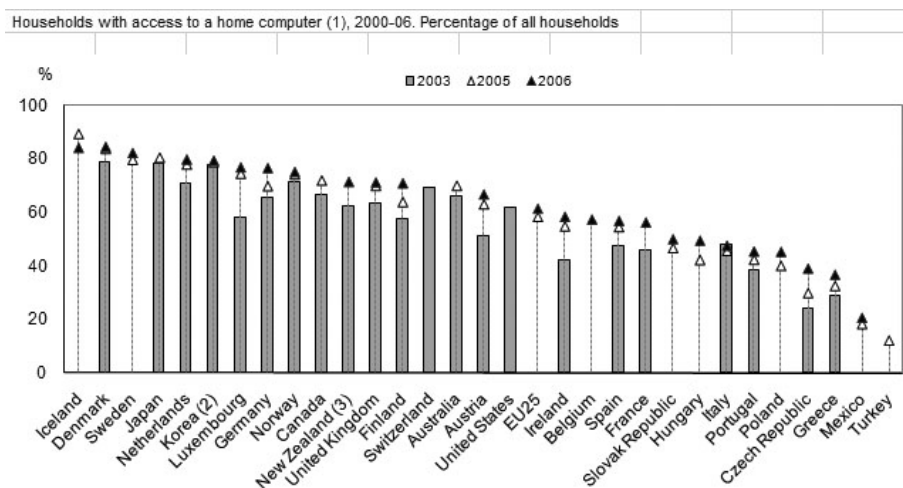


Fig. 2. Home computers penetration (Source: OECD, ICT database and Eurostat, Community Survey on ICT usage in households and by individuals, April 2007).

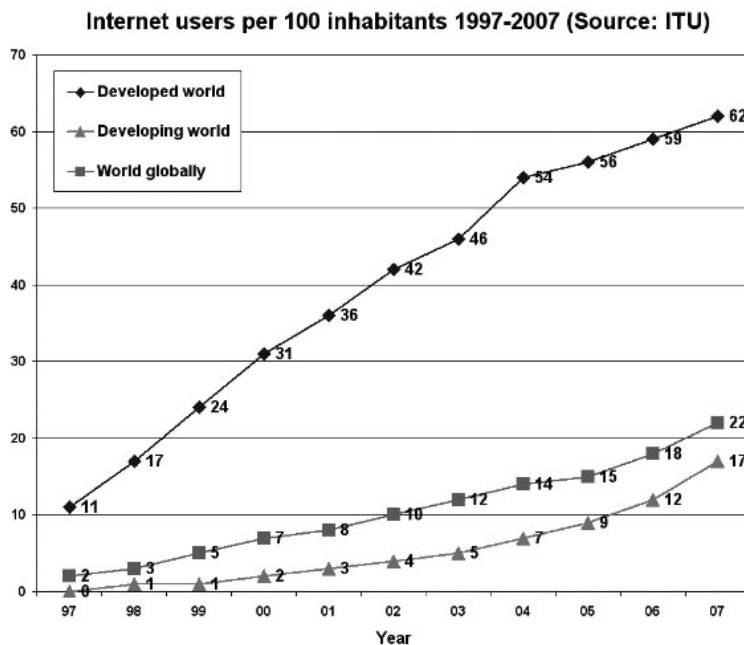


Fig. 3. Internet Users Trends 1997-2007.

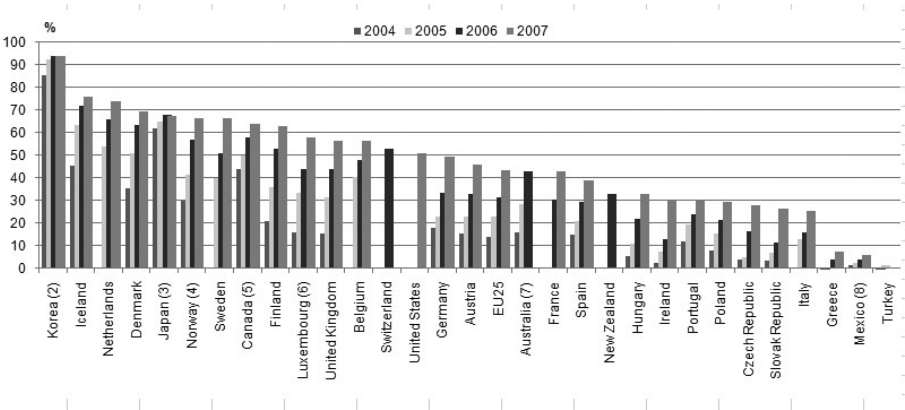


Fig. 4. Households with broadband access 2004-07. Percentage of all households. (Source: OECD Broadband Statistics, OECD Broadband Portal, URL: < http://www.oecd.org/document/54/0,3343,en_2649_34225_38690102_1_1_1_1,00.html>)

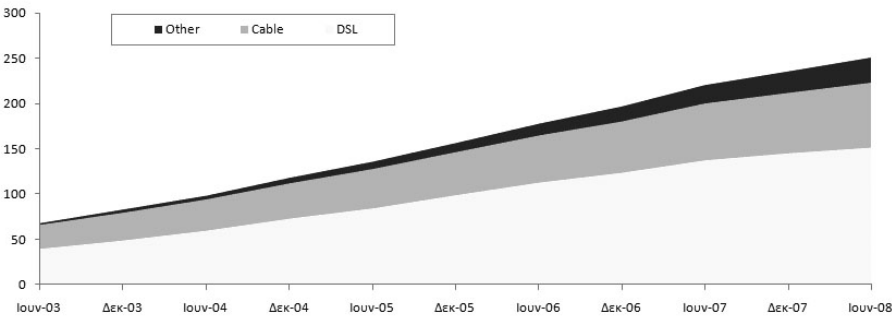


Fig. 5. Broadband Growth Total (2003-2008) Million Subscribers. (Source: OECD Broadband Statistics, OECD Broadband Portal, URL: < http://www.oecd.org/document/54/0,3343,en_2649_34225_38690102_1_1_1_1,00.html>)

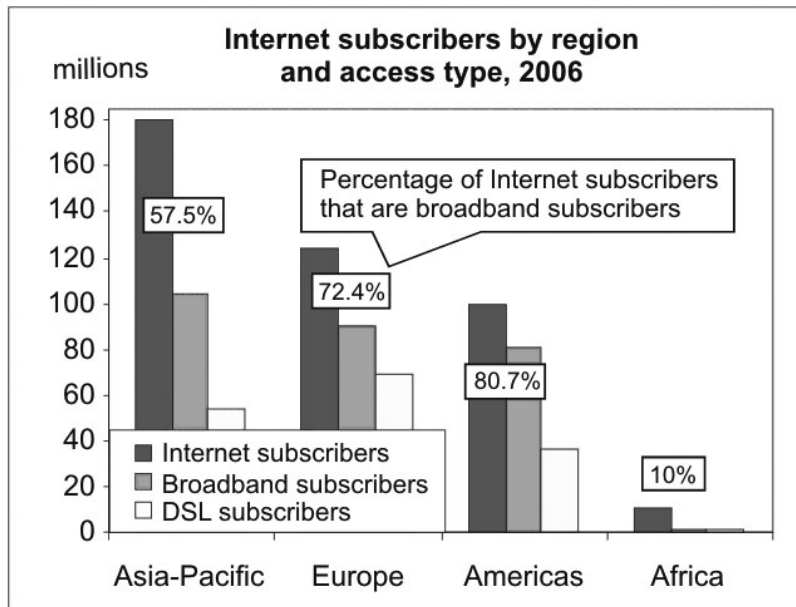


Fig. 6. Internet Subscribers by Region and Access Type, 2006 (Source: ITU, ICT Statistics, URL: < <http://www.itu.int/ITU-D/ict/statistics/maps.html>>).

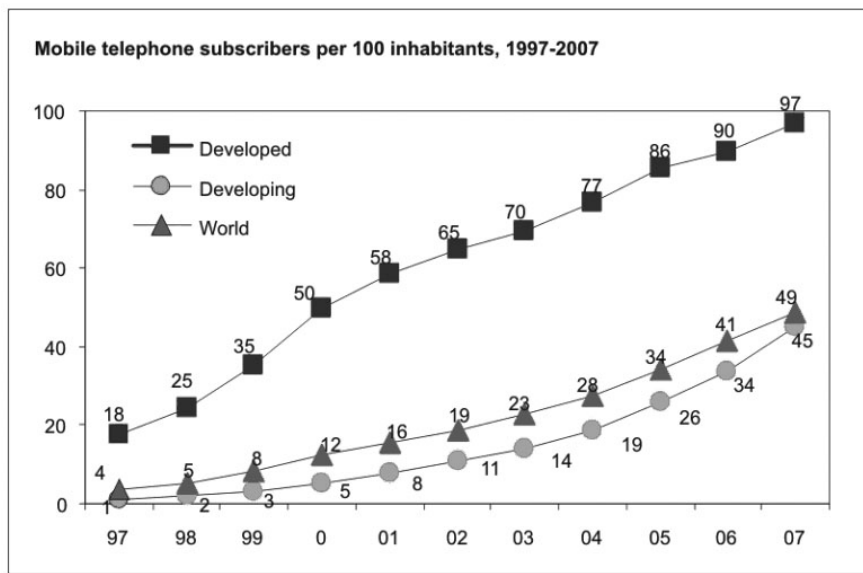


Fig. 7. Mobile cellular penetration rates worldwide (Source: ITU, ICT Statistics, URL: < <http://www.itu.int/ITU-D/ict/statistics/maps.html>>).

WORLD INTERNET USAGE AND POPULATION STATISTICS						
World Regions	Population (2008 Est.)	Internet Users Dec. 31, 2000	Internet Users Latest Data	Penetration (% Popula- tion)	Users Growth	Users % Of Table
<u>Africa</u>	975,330,899	4,514,400	54,171,500	5.6 %	1,100.0 %	3.4 %
<u>Asia</u>	3,780,819,792	114,304,000	657,170,816	17.4 %	474.9 %	41.2 %
<u>Europe</u>	803,903,540	105,096,093	393,373,398	48.9 %	274.3 %	24.6 %
<u>Middle East</u>	196,767,614	3,284,800	45,861,346	23.3 %	1,296.2 %	2.9 %
<u>North America</u>	337,572,949	108,096,800	251,290,489	74.4 %	132.5 %	15.7 %
<u>Latin America/ Caribbean</u>	581,249,892	18,068,919	173,619,140	29.9 %	860.9 %	10.9 %
<u>Oceania/Australia</u>	34,384,384	7,620,480	20,783,419	60.4 %	172.7 %	1.3 %
<u>World Total</u>	6,710,029,070	360,985,492	1,596,270,108	23.8 %	342.2 %	100.0 %

Fig. 8. World Internet Usage and Population Statistics (Source: Internet World Stats, URL: <http://www.internetworldstats.com/stats.htm>).

This few data and, especially, the suggested by the data trends are enough, I believe, to substantiate, beyond any doubt, the claim that ICT is already now an everyday tool, although there are differences between several societies or sectors of the same society. It is apparent and undeniable, I think, that it is simply a matter of time for the whole modern society, all over the world, to fully adopt and use the ICT as an everyday tool. I believe that the situation and the trends are obvious. I believe also that the resulted and expected changes in our mentality, attitudes and expectations should also be obvious. I remember that back in 1985 the National Library of Greece spent a large amount of money for a 10MB hard disk. Today you can buy a 2TB hard disk for less than 250 Euros (price checked in April 2009), and you can fill it up with content quickly, using your broadband connection to the Web. A 2TB hard disk can store almost 3.000.000 traditional 300-page books. This is more, in terms of number of items, than the whole collection of the National Library of Greece today. The described situation and trends suggest, I think, a radically different mentality for the information producer and user of today and tomorrow. The Library and Information Services sector, thus, and the society in general, should reconsider their attitudes and practices with regard to information and content use, in the light of this new type of relation of people with information. I am afraid, however, that people, even experienced people, fail sometimes to see the obvious. I remember today, in this concern, the famous remark of Ken Olsen, founder of Digital Equipment Corporation, back in 1977, that, quoting his words, "there is no reason for any individual to have a computer in his home". Although Olsen later explained that he "did not object to the concept of a PC, but to a computer in the home controlling everything", even

this clarified view of the situation can be considered today, in the light of the aforementioned evolutions and trends, as a misjudgement⁹.

The user of information services

The operational environment briefly described above means that the potential end user or rather the emerging potential user of library services today is, from this point of view, something absolutely different from the library user of the previous century, as concerns his capabilities, habits, practices, beliefs, desires and expectations.

The user in the real world belongs, of course, to several categories. They can be researchers, the academia, students, pupils, professionals of any field and kind e.tc. For each of these categories the respective user has its own characteristics, typical to the category he belongs to. These characteristics are of no interest to us today, although are of great importance to the Library and Information Services. What are of interest to us here are characteristics common, more or less, to all categories of users and related or arisen from the new digital information environment. These characteristics are common to the users or people involved in any information-related activities of today and, of course, not only to the users of Library and Information Services. It is obvious, however, that it is extremely difficult to examine here even these specific characteristics of what we can call, as a convention here, the “digital” user, although it may be better to call him the “hybrid” user. It is difficult, first of all, because such a presentation is not the suitable environment for a task like this. On the other hand, the “digital” user is still an evolving concept and an entity under development and formation. We can, thus, mention here only some of the more obvious features that are related to and characterize the user of the digital environment and, especially, the features that substantiate the claim for a radically different user mentality.

The potential user today operates in and uses extensively the tools and methods of the modern technological infrastructure. That is, computers, PDAs, the Internet, the Web, mobile phones and other communication devices, e.tc. Although it is true that not all potential users are accustomed to and use the available technological infrastructure, it is apparent I think, according to the aforementioned relevant trends, that this will be the case in the very near future. It should be stressed here again, that these are not the tools for accessing only the Library and Information Services, but the technological infrastructure is used heavily today for a wide range of social activities. This is especially important since the extent

9. A brief account of the story can be found on the Web, URL: <http://www.snopes.com/quotes/kenolsen.asp>, viewed, 4 April 2009.

and broadness of the use is the crucial factor for affecting decisively the mentality of the user.

For this evolving type of information services user the distinction between the several partners of the information and content production, dissemination, management and use cycle is not really apparent or interesting. The traditional distinction, also, between what we call Library and Information Services sector and other social sectors or groups of activity, is becoming less and less apparent or interesting. The user today seeks with the same tools and with the same methods for the prices, for example, of a telephone company and for an article in a scientific journal.

This user tends in an increasing way to recognize more and more only one point of service for his information related needs, and this is the screen of the device used for accessing the Web. He is becoming more and more accustomed to get the information and the content he needs quickly (if not instantly), cheaply (if not free at all) and easily from his home, or better, in any place and at any time, using his desktop computer, his notebook, his netbook, his mobile phone, his PDA, or any other device of this or of a similar kind as the point of access. Also, instead of the library staff, he seems to prefer the search engines of the Web, as his gateway to the world of information. For the user in the digital environment, thus, information, of any kind and for any use, content of any kind and digital objects of any type are all parts of a vast and vaguely defined tank. This evolving perception concerning information and content, and I mean, mainly, the vagueness of the content's location, provenance, ownership, status e.tc. is a very important factor with regard to the development of user attitudes toward content that seem to revolutionize the relevant field.

A mentality being developed under the influence of characteristics like speed, easiness and cheapness of access is a fact that should be taken into account seriously. The libraries and the society in general should not disregard the fact that people, in order to solve a problem, to find, for example, information on a specific matter, a certain book, a movie DVD, a musical work e.tc. will select the fastest, simplest, easiest and cheapest way available. They will also select the most familiar way available. Due to this undeniable truth people living and working in the context and with the support of the current technological infrastructure will select the Web instead of a library when searching for information or content, will select the Web instead of a video club when searching for a movie e.tc. Moreover and for this same reason people will opt for music in mp3 or other similar formats instead of SACD or DVD-AUDIO (and we all know now that these two high quality formats for digital music are now almost obsolete), they will opt for DIVX movies that they can get easily, quickly and cheaply from the Web,

instead of waiting to obtain the high definition blue-ray disk from the local video club e.tc. e.tc. This is an undeniable truth and, since the new technological environment can support it, the Library and Information Services field should adopt it and design its services accordingly. Moreover the society in general should accept and adopt it and redesign its traditional models for content dissemination and supply accordingly.

This is not only because service models designed on the basis of user's newly developed expectations, habits and practices will be more effective, but, and this is possibly much more important, because the aforementioned trend of the user presents, for the time being, serious side effects. First of all, in the context of the poorly organized and rapidly changing and expanding Web environment, this trend results in content piracy and violation of rights, a fact which in turn has a very negative and undesirable influence on the relevant market. Moreover it might be a bad influence as concerns the quality of the content and mainly the addiction of the end user to this kind of access to and quality of the content. And a long term addiction and use of practices of this kind will be an absolutely undesirable kind of education for the new generations, concerning the access to and the quality of the intellectual product.

Disregarding, thus, matters of quality, accuracy, reliability, integrity and authenticity of information and content, as well as violating property rights related to the digital objects, may become, sooner or later, an integrated part of the user mentality. Phenomena of this kind are signs and steps toward a kind of social entropy, and such a trend is not desirable at all. However, the solution is not to disregard the potential of the technology and the expectations of the user and, consequently, as is the case today, to try to forbid several methods of access to or kinds of material. The solution for the libraries and the society is to develop service models that adopt the full potential of technology, accept the justified attitudes of the end user and offer services, information and content in accordance with the user demands, desires, expectations and habits, without the aforementioned side effects. I am not in a position here to suggest detailed new business models of this kind, but I think that I am in a position to suggest the basic principle for future actions, and this is the conformance with the visible or envisaged trends.

A few remarks about the emerging changes in content matters today will clarify, I believe, the whole situation even better. The libraries of today, for example, cannot expect to continue using for a long time more only their traditional lending system for physical items, when the potential user of today is accustomed to get the content he needs quickly and easily from the Web. It is obvious that a new system is required and this has to do not only with the technology but also, or

better, more with the deontology, ethics, and practices of the whole process. The key-point here is that the libraries should try to address not or not only the technological improvement of their existing services, but to attempt, in collaboration with the other players in the field (I mean here content creators, publishers e.tc.), to develop new service models, exploiting the full potential of the new environment and in accordance with the expectation and demands of the end user. This, of course, is not an easy task, but a good starting point is a radical change of the society's attitudes toward these matters.

It is, of course, undeniable, that there are still vast collections of traditional printed material that the libraries can lend only with the use of their traditional lending system. But we should not disregard the extensive digitization projects under way all over the world. Thousands of libraries of all sizes have scanned books or other relevant material, catalogued them and made them available on the Web. We can mention here, just as an example The Library of Congress National Digital Library Program (NDLP)¹⁰, the Europeana Project¹¹, with more than 4 million digital objects of several kinds (images - paintings, drawings, maps, photos and pictures of museum objects, Texts - books, newspapers, letters, diaries and archival papers, Sounds - music and spoken word from cylinders, tapes, discs and radio broadcasts, Videos - films, newsreels and TV broadcasts, e.tc.), the Carnegie Mellon University Libraries Million Book Project¹², the well known Google Books Digitization Project¹³, e.tc., e.tc.

It is more than obvious that sooner or later a vast amount of the printed or other similar documents will be available in digital form and on-line, a fact that will allow, among other things, the exploitation of this material with the tools and capabilities provided by the digital infrastructure. But this is not the only benefit. These projects will make available a vast amount of printed material that is now out of print and accessible with great difficulty or, in certain cases, not accessible at all. But perhaps, and as far as our interests in this presentation are concerned, the most interesting consequence of these projects and of this widely spread trend will be its impact on the mentality of the user of content and information. Contrary to what is the situation today, the belief that there are or there should be no barriers or difficulties in locating and accessing information and content will become a fundamental part of the mentality of the user.

10. <http://memory.loc.gov/ammem/dli2/html/lcndlp.html>.

11. <http://www.europeana.eu/portal/>.

12. <http://www.ulib.org/>.

13. <http://books.google.com/googlebooks/library.html>.

This trend will have, without doubt, another serious consequence, since the libraries in the future will be more and more under great pressure to digitize their material, so that they conform to the demands of the general information environment. This means that such a digitization will be a prerequisite not only for coping with the user's demands, but also for their general ability to function in the context of the new information environment, by cooperating with other libraries or similar institutions, by exchanging material e.tc. e.tc. The libraries had experienced a similar situation some decades ago¹⁴, when, under the pressure of the emerging automated environment of the library's functions, they decided and carried out extensive projects for the retrospective conversion of their catalogues from manual card catalogues to automated ones. The main driving force behind that movement was the willingness of the libraries not to be isolated and to be able to function in the then emerging new automated environment. The gradual emerging, on the other hand, of this new generation of users, with attitudes and expectations completely incompatible with the traditional library lending systems, which we used here as an example of the whole situation, will render, sooner or later, these systems *de facto* obsolete.

From the point of view, however, of the existing business ethics and practices, the insistence of libraries in the traditional lending systems of the printed material, ignoring the potential of digitization and the development of new service models, so that they avoid an eventual support to content piracy, is not a solution at all. On the contrary, lending material of this type is an easy way for a typical library today to become, inadvertently, a precious source of illegal digital content. The solution, thus, is not ignoring the future and its potential, but the attempt for the development of new and suitable service models for content supply and use, in conformance with the potential of the technology, the expectations, beliefs and demands (that is, the mentality) of the user and the need to fully respect the rights of the content owners. The development of a new business model, however, requires a radical and extensive reconsideration of the practices, ethics, beliefs, attitudes, and deontology that constitute the models for content supply and use today. The libraries, for example, may reconsider the concept, the limits and the terms of application of the free access to information, while content providers may develop a new pricing model, since the production and other relevant costs, related to printed material, are considerably different and, of course, much higher, from the costs necessary for the digital content.

14. In fact and due to the vast amount of the relevant work that should be done, retroconversion projects are still under way in several countries all over the world. See, as an example, the Retroconversion of the Russian State Library Catalogue of Serials and Periodicals, Web document, URL: <http://www.rsl.ru/en/s7/s20/d23/>, viewed, 6 May 2009.

We should not disregard also the new features that have been attached to and transformed radically the concepts of information and content. This covers a wide range of characteristics, starting from the digital nature of information and content and regarding aspects such as its reliability, stability, accessibility, authenticity, legality e.tc. This suggests that the digital material is not another kind of material, parallel to the other kinds that constitute the collection of a traditional library. In fact it is another WORLD of material, which requires a completely new set of tools, methods and practices for its handling and exploitation. This new world of material, suggests a relevant study of OECD, "has become an increasingly important and pervasive factor shaping both economic and social development. Supported, as already mentioned, by high-speed communications, increasing upstream as well as downstream bandwidth, declining access prices, convergence of previously distinct networks, innovation in new devices and applications and lower entry barriers will drive new ways of creating, distributing, preserving, and accessing digital content. As economies move towards being more knowledge-intensive, information-rich activities in which content is created, collected, managed, processed, stored, delivered, and accessed are spreading into a broad range of industries, contributing to further innovation, growth and employment. Digital content is becoming central in research, health, education and social services, knowledge and cultural services and government. It is also stimulating increased participation and creative supply by users"¹⁵.

Digital content covers today fields of social activity and creation such as scientific publishing, music, on-line computer and video games, mobile content, public sector information, user-created content, film and video and on-line advertising. This widely spread in modern society presence of digital content in terms of creation and use, has a continuously increasing impact on the relevant mentality of the user. Specifically, information and content is becoming more and more, in the mind of the user today, a new type of commodity, with its own features and characteristics, and which has nothing to do with the distinctions between the several information products of the analogue environment (i.e. the book, the CD, the DVD e.tc.).

Moreover in the context of this diffusion of characteristics and roles among the several components and partners of the information cycle, the distinction between creator and user of information is gradually vanishing, as is also the distinction between supplier and consumer of information and information products. As already mentioned, the Web user is gradually becoming a Web producer and vice versa. This is an increasing trend, which is gradually transforming

15. OECD Policy Guidance for Digital Content, Web document, URL: <http://www.oecd.org/dataoecd/20/54/40895797.pdf>, viewed, 8 May 2009.

the Web into what is called participative Web¹⁶. Although this trend has in fact just begun, the user created content is considered a very important aspect of the emerging new information environment. The OECD defines User Created Content as the content that fulfils the following requirements:

- a) It is published in the Web,
- b) A certain amount of creative effort has been put into its creation either by creating a new work or by adapting existing works to construct a new one, and
- c) It is created outside of professional routines and practises. It often does not have an institutional or commercial market context and UCC may be produced by non-professionals without expectation of remuneration or profit. Motivating factors include: connecting with peers, achieving fame, notoriety or prestige, and expressing oneself.

Although conceptually useful, the last User Created Content feature, that of a creator who does not expect remuneration or profit and of a creation being outside of the normal or typical professional routines has been proved difficult to be maintained. Although UCC have begun as a grass roots movement, not focused on monetary rewards, monetisation of UCC has been a growing trend. I believe that this is normal and, also, a very interesting sign that business models, suitable for the new environment, will eventually replace, sooner or later, the current situation.

It should be noted, however, that an extensive research of the kinds and the variety of User Created Content, showed some very interesting characteristics, apart from its relation to monetary rewards. It showed, for example, that in many cases the basic motive of such a creation was a previous relation or connection of the Web creator with the field or the object for which he creates content. Moreover the quality of the content offered for free is sometimes unbelievably high and at a really professional level, while the extent and the amount of material uploaded is in many cases really enormous. It should also be noted that, as my relevant research has shown, in many cases the kind of work offered by the user Web creators is of a kind not normally offered by the traditional professional publishing channels, providing thus, material or forms of material not to become available in another way.

But the quality or the special character of the User Created Content is not the most important matter from our point of view. The important thing is the fact that the place, the role, and, consequently, the attitudes of the user toward con-

16. OECD, Participative Web and User-Created Content, 2007, Web document, URL: < <http://213.253.134.43/oecd/pdfs/browseit/9307031E.PDF>>, Viewed, 8 May 2009.

tent creation and use and, of course, toward all the issues related to this matter (e.g. Ways of availability, methods of access, property rights e.tc.) are undergoing a considerable transformation. These attitudes, which suggest a very important change of mentality, may be summarized here as a widely accepted belief that information and content should be available for a cheap (if not free), quick (if not instant) and easy access and use, without any limitation. This may seem and it is, of course, in some cases and with the standards of today, a problem from the legal point of view, since in many cases the rights of the copyright owners are violated.

However we should not disregard here the fact that an enormous number of anonymous people all over the world, or rather all over the Web world, spent a large amount of their time to process, digitize and upload content in the Web with remarkable assiduity and with no apparent personal profit. A really considerable amount of this content presents no legal problems, is of an excellent quality, requires a vast amount of time for its creation and it is offered for free to the users of the Web. The extent of this phenomenon is of a special importance, since it is an indication of the gradual development of very interesting new attitudes of people toward content creation, access and use. It is true that, within the current legal framework, many of these activities can be considered as illegal. This is so even when they offer for free material with considerable added value and/or material that is not available in any other way legal or not, difficult or not, time consuming or not e.tc. e.tc. From this point of view and for this part of the society, I mean here the part represented by the active users of the Web, we are witnessing today a continuously increasing deviation between what is the *de jure* and the *de facto* social views with regard to the right of access to information and content.

This is not a simple issue, since it implies a gradual and very important change in the mentality of the Web user, with already visible results in the every day practices of publishing and acquiring content. Moreover we should not disregard the fact that these changes in mentality concern a continuously and steadily expanding part of the society, since more and more new generations are added to the tank of the active Web users. Some charts concerning the various activities of users, with regard to the content creation for the Web, may be given here as an example of the situation and the trends.

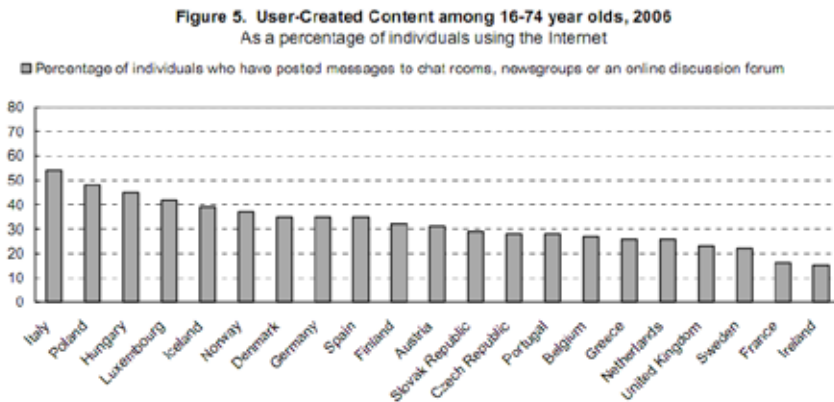


Fig. 9. Source: OECD. Working Party on Indicators for the Information Society, *Measuring User Created Content: Implication for the "ICT Access and Use by Households and Individuals" Surveys*, (DSTI/ICCP/IIS (2007) 3/FINAL).

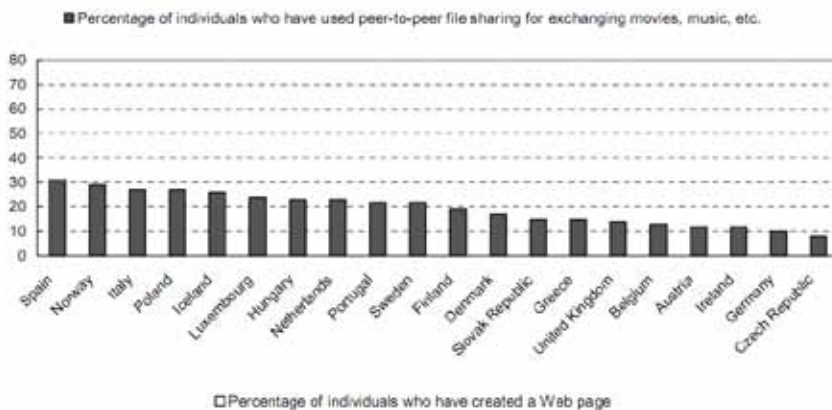
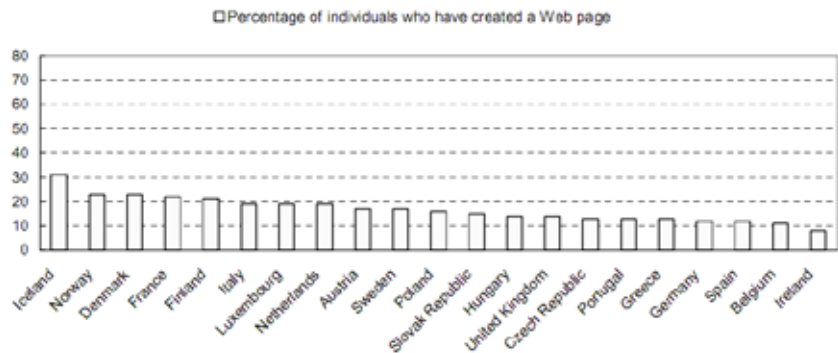
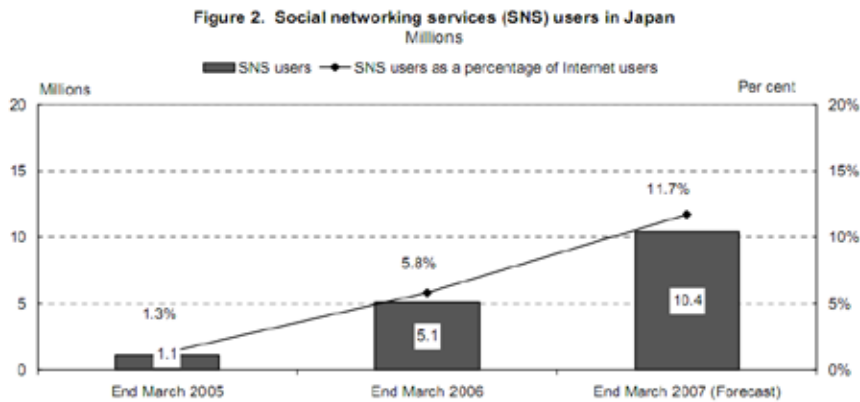


Fig. 10. Source: OECD. Working Party on Indicators for the Information Society, *Measuring User Created Content: Implication for the "ICT Access and Use by Households and Individuals" Surveys*, (DSTI/ICCP/IIS (2007) 3/FINAL).



Source: EUROSTAT, February 2007.

Fig. 11.



Source: MIC, Analysis on Current Status of and Forecast on Blogs/SNSs, May 2005 and April 2006.

Fig. 12.

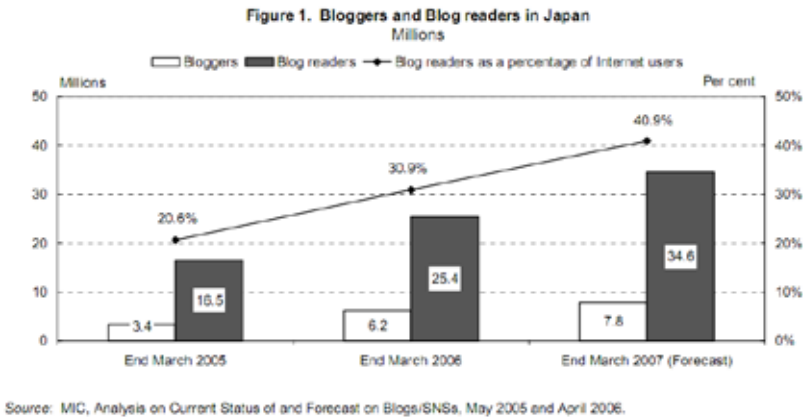


Fig. 13.

For the time being the reaction of the official representation of society is limited to legal measures against people violating intellectual property rights. This kind of reaction suggests that the society cannot understand the underlying factors of such “illegal” actions, which are not, in most cases, a “criminal” attitude, but a gradual and radical transformation of the society’s attitudes toward content production and use. Since the change in mentality cannot be stopped with legal actions, this approach is difficult to be successful in the end.

The new and parallel digital world, and of course, the emerging hybrid world, cannot be ruled with tools developed by and for the analogue world. We should not forget that the recent conviction of the people of Pirate Bay was followed by the election of a member in the European Parliament by the “Pirates Party”. The solution, as usual and as always, requires the acceptance of the facts, and the fact of a radically changed environment and of a respectively changed mentality is a very crucial one. These changes require not legal actions, based on the legal foundations of the old world, but rather new business models and practices, concerning the cycle of information production and use. They may also need a new legal framework, in conformance with recent and emerging social attitudes and practices. In any case we have examples proving the validity of such an approach. We could mention here, for example, that far-sighted people have already tried radically different approaches to information and content supply with very successful results. It is enough here to mention the article by Chris Anderson, “Free! Why

\$0.00 Is The Future Of Business”, published in Wired, in February 25, 2008¹⁷, or the article by Mike Masnick, “A Business Model Involving Free File Sharing”¹⁸.

Besides, users creating content for the Web are integrated, sooner or later, in suitably developed business models that offer services and products, not offered by the traditional information framework, they are not at all illegal, they correspond to the Web user expectations and needs, they secure a continuously increasing number of customers and, as a consequence, they make a profit. As a result they succeed in transforming gradually the relevant business models and environment.

On the other hand and as expected, established media and Internet businesses have increasingly acquired UCC platforms for commercial purposes. Some users are remunerated for their content and some become professionals after an initial phase of non commercial activity. The whole environment, thus, and the relevant business models are undergoing a gradual, although not so obvious, transformation.

The term UCC may, thus, cover content creation by those who are much more than just “users”. Still, the creation of content outside of a professional routine and organisation and potentially not for reward is a useful characteristic to separate it from content produced by commercial or quasi-commercial entities for commercial purposes. More importantly it denotes the existence of an entirely new component in the composition of the content creation world. But the most important point here is the radical change induced in the mentality of the user toward content. Assuming the two opposite roles, of the user and of the creator alternatively, the user is no more the simple customer of the information services, but an active partner of the whole process. This change of status has a serious impact on the user’s views and attitudes toward several aspects of the creation and use of content, such as the extent of the right to free access, the protection of the copyright of the digital objects, the right to use content in several ways (e.g. adapting it) without the permission of the original creator e.tc.

All these changes in attitudes suggest that the emerging new mentality of the user cannot be fully accommodated either by the existing business practices and models, or by the prevalent legal provisions. This emerging mentality also does not seem to conform to the existing ethical views with regard to these matters. It is obvious, thus, that the whole ethical, business and legal framework, supporting

17. Wired Magazine, Available on the Web, URL: < http://www.wired.com/techbiz/it/magazine/16-03/ff_free?currentPage=all>, Viewed, 12 May 2009.

18. Document available on the Web, URL: < <http://www.techdirt.com/articles/20030912/1032238.shtml>>, Viewed 6 May 2009.

the society today should be reconsidered and revised toward the lines suggested by what are or by what we are expecting to be the attitudes and expectations of the user and creator of content in the near future.

For the time being, however, the problem for the Library and Information Services is obvious, since as traditional providers of information and content, gateways to information and content and intermediaries between creators and users, cannot either disregard the existing and ever increasing content available on the Web or violate the laws concerning copyright. The Library and Information Services, thus, besides trying to improve their existing services using the available technology, should try to develop new models of access to and supply of information and content, in cooperation with the other interested partners in the field. In such an approach they should use as a guideline the objective of fully exploiting the potential of not only the technology, but also of the extensive human and social structures that are active and that are continuously emerging in several forms and types of interaction, within this new world, the World Wide Web, creating, transforming, publishing and exchanging information and content.

This last remark and fact, regarding the extensive human web that is being developed over and within the World Wide Web, in the form of human relations, of exchanging of material and ideas, of publishing, discussing or transforming content is something that should not be disregarded and something, I believe, with far reaching consequences in the mentality of the user. This is because, I believe, the impact of the new environment is not limited to a remote and theoretical influence, like, for example, the impact on society of discussions about global warming in newspapers, TV broadcasts etc. In this case the development of complex human and social structures, which are in various ways interrelated and which produce, exchange, transform and test content and content producing tools and practices is the object of everyday practice, discussion and exploitation. At the same time, of course, Library and Information Services should also try to transform their own material and the relevant methods, tools and policies of its supply, in a way that will satisfy the emerging user community, making thus this material really available to the widest possible user population. However the changes under way demand many more changes from the part of Library and Information Services.

This is due to the fact that the changes this new kind of information and content has induced in information and content handling, management, disseminating, using and exploiting are not limited to issues of handling, but they are also related to the nature of the digital content itself. We have already mentioned the radical change which transformed the analogue object of information to a digital one. This change of nature had tremendous consequences in the essential charac-

teristics of the information object itself. From the point of view of management, which is the main interest of the Library and Information Services sector, the digital object is something radically different from the physical object, which has been so far the object of interest for libraries. There are a lot of relevant changes from this point of view and there is an extensive literature on this subject defining the essential characteristics of the digital documents, comparing them with their analogue counterparts. It is enough here to mention the differences in document structure, stability, limitless nature, potential of accessibility, reliability, authenticity, integrity e.tc. e.tc.

These are more or less well known things. What, however, I should mention here is the very serious change induced by these features in the attitudes of people, especially young people, toward information and content. This means that the object of interest of the emerging generation of users has radically different characteristics from the objects that the Library and Information Services sector is prepared and manages successfully for decades now. The mentality of the emerging user generation cannot, for example, accept or understand difficulties or delays, concerning the accessibility to what is available, while, at the same time, has serious concerns about the integrity, the reliability and the authenticity of digital objects, as well as about the potential availability of this type of content in the medium or long term. The unpleasant experience, for example, of the appearance of the message "error 404 page not found" is becoming more and more frequent even in the case of references from serious scientific journals when the reference is made to a Web document.

This instability and liquidity of the Web digital objects, combined with the aforementioned concerns about the quality of the Web documents lead to the development of a mentality and a respective demand on the part of the user that requires all the advantages of the digital information with the less possible disadvantages. This is normal and expected. The point is that such a mentality should guide the attempts of the libraries to reconsider their role and to develop plans for the improvement of their services. The amount and the variety of the digital material produced every moment, the ways it is produced and published, the creators and producers of this material and the relevant place and demands of the user in this context outline a radically different role for libraries and information services.

In this context we should not underestimate the fact that the changes outlined so far created a whole new, and parallel to the conventional one, environment of social activities in general. Of course it is of interest to assess the current relative extent and use of these two environments and, most importantly, the rate of change of this relation between the two. However, although extremely important, there are still not enough data for such an assessment or rather for such

a comparison. On the other hand, collecting this kind of data is extremely difficult due to the tremendous rate of every day expansion of digital or Web based universe. This expansion should be considered in the form of the Web expansion (e.g. as the number of interconnected networks, hosts, Web sites e.tc.), of the Web-based activities, of the people using the Web, of the tools available for exploiting the Web and the digital form of information (e.g. Computers, PDAs, mobile phones, e.tc.), of the bandwidth and the speed available for accessing the Web content, of the gradual elimination of local or any other kind of barriers to access the “Web world”, of the available Web content itself e.tc., e.tc., e.tc.

In this sense we tried here to give some indications of the relevant sizes and trends, when discussing the technological environment. Although those selected data were limited, the relevant trends were, I think, obvious. In the context, however, of such a situation of coexistence between “bits” and “atoms”, or between a digital and a conventional world, we may think for a moment that the centre of these activities, the human being, consists only of “atoms”, belonging, thus, in fact, to the one of these two worlds. Fortunately or unfortunately this is not true. The human being consists of “atoms” but exists and acts, in the sense of thinking, assessing, deciding, e.tc. as a digital entity. He pertains, thus, to both worlds in the same manner this seems to happen in our social environment today.

As a further result of this new situation, the new form and nature of data and data dissemination and use have cancelled the usual conventional barriers between the several social and professional sectors, as concerns the tools, the practices, the habits, the techniques e.tc. concerning the acquisition, the management, the handling, the use, the dissemination and the exploitation of information. In fact there are no more really different worlds, as concerns, for example, tools, standards and practices regarding data and information, between several social or professional sectors. This means, among others, that data and information can be produced, processed, disseminated and used through several sectors and for several purposes, using the same tools, forms of data and channels of data dissemination in an absolutely transparent way for the interested user of information. Such a situation means, at least, a very extensive reconsideration of roles and responsibilities, throughout the modern social environment and between the conventional partners of this environment.

It is obvious that these developments, creating a whole new and parallel environment for human activities, but also demanding a new mentality and a new way of thinking and absorbing facts, have created a number of problems affecting and related to most of the sectors of the modern social activities of all kinds. There are problems of acceptance, problems of coping with the rate of change, problems of understanding, problems of lack of the necessary tools, practices and methods,

problems of an effective combining the old with the new, problems arising from unsuccessful attempts to handle the new with ways and attitudes coming from the old, e.tc., e.tc. It is, however, inevitable that the problems should be solved. And of course we are witnessing today an unprecedented effort by the society as a whole to handle these problems (e.g. By creating new tools, by discussing problems of conflict arising here and there, by pointing out issues arising from the emerging new environment, e.tc., e.tc.).

But if Libraries and Information Services, which are our example sector here, are to handle this situation in an effective and successful way, some essential facts and truths should be kept in mind and guide, more or less, their relevant approach toward solving the present problems and, especially, planning for the future. It should be mentioned, however, that the guidelines outlined below, with regard to the Library and Information Services sector, are, more or less, similar and equally important for any other sector or field of social activities today.

These facts and, I believe, truths, may be summarized as follows:

1. The problems seem more serious because they are due not to a conventional evolution of the existing world, but to an emerging completely new and parallel world and, consequently, a parallel new operational environment. Moreover, and using the simplest possible redaction of the situation, we could say that we have to cope not only with the problems of two different worlds at the same time, but also with the problem of combining them into a new viable world for the people of the future. We are accustomed to the traditional “atoms” world and, possibly, we will find our way in the emerging new “bits” world. But our perception of the future social environment, and, thus, our plans and actions should be guided by the principle that the foreseeable future will be a hybrid one in terms of both the content and the actions on the content.
2. The relevant problems might look more serious because Library and Information Services have to cope with a situation which, however strange and peculiar, is not stable, but, to the contrary, a still rapidly changing one. It should be pointed out here that this concept of “change” covers not only the obvious and rapid change of the technology and the tools, but also the gradual, although not so rapid, change of human mentality, attitudes and expectations. The latter presents, for several reasons, the more serious problems. First of all the human mentality changes slower than the technology and the relevant infrastructure. Secondly, mentality and attitudes cannot change with a decision of the state. Moreover it changes with different speed for the different human generations. This difference, in our case of the sudden,

radical and very rapid changes of the operational environment, resulted in a serious deviation between the professionals of the field, or more precisely, the older generations and the new ones, with regard to information and content production, dissemination and use. Policies, thus, and practices should be designed so that they allow for a continuous adaptation to new forms of content, new media for the dissemination of information, new expectations and demands of the potential users e.tc.

3. Any problem or situation cannot be considered or handled in isolation or with regard to a specific social sector or activity. The most important aspect here is that we should consider not the specific problem, but the specific problem in the context of a whole new world and, at the same time, with the heritage and determining power of the traditional world (I mean here in the sense of practices, mentality, tools, e.tc.). The future planning principle here is that any problem or issue is not concerned with a specific field only, but it is rather a problem of new attitudes and perceptions of the society, which are, more or less, similar for the society as a whole and the acceptance of this truth should be the guideline for any action toward solving problems or developing policies. The content creation, for example, in the new environment, faces the same problems, issues and challenges in any field related or involved in content creation, management and use. At the same time the tools used for every activity in this field are, more or less, the same.
4. The proper mode of action is not the attempt to use the new technology, in order to offer the existing services in a better way, but the attempt to design and offer new kinds of services and products in new ways, by fully exploiting the potential of the new technological infrastructure and by taking into account the real needs and expectations of the users. In a recent survey of research libraries in the USA their users were asked to rank several proposed measures and actions concerning the improvement of the services offered. Two of them and the corresponding ranking are of interest here. The first concerned the provision of the library so that its users have the ability to access information and content offered by the library from any place and at any time they want. The users ranked this proposed new feature of service improvement very high. The other concerned the proposal of the library to get new and suitably experienced personnel, so that the user could get better help and support within the library. The users ranked this proposed feature of improvement very low. In fact they showed that they were uninterested in using better local services by the library and what they really wanted was a better ability to use the contents and the services of the library with the way and the tools they are now accustomed to.

5. The proper way of action is not to insist on existing models of content dissemination and use, if these models directly oppose both the nature of the new content and the beliefs, expectations and everyday habits of users. The proper way is to work hard, not the libraries in isolation, but the whole field of content creation and use in cooperation, with extensive use of the available technology, so that new practices, tools, methods and business models are developed, in conformance with the characteristics of the new information environment and of the new mentality, attitudes and expectations of the user today. Some indications of the important changes under way, with regard to the mentality of the user, have already been described here. In the meantime the easy part of the change should be undertaken by the libraries and this is the transformation of their material and relevant practices so that they correspond to the potential of the technological environment and the demands of their users. Such an orientation will lead, without doubt, to a real improvement of the services offered.

Philosophy and the Information Revolution

Terrell Ward Bynum
Professor of Philosophy
Southern Connecticut State University

Introduction

In 1998, in our book *The Digital Phoenix: How Computers Are Changing Philosophy*, James Moor and I noted that computing and related technologies are providing philosophy with new and evolving subject matters, methods, and models for philosophical inquiry. [...] Most importantly, computing is changing the way philosophers understand foundational concepts in philosophy, such as mind, consciousness, experience, reasoning, knowledge, truth, ethics and creativity. This trend in philosophical inquiry that incorporates computing in terms of a subject matter, a method, or a model has been gaining momentum steadily. A Digital Phoenix is rising! (Bynum and Moor 1998, p. 1)

Now, more than a decade later, it is clear that, indeed, the Digital Phoenix has risen. Relevant philosophical books and articles have proliferated on a vast range

* *Terrell Ward Bynum* is Professor of Philosophy at Southern Connecticut State University, Director of the Research Center on Computing and Society there, and Visiting Professor at De Montfort University in Leicester, England. He is a lifetime member of Computer Professionals for Social Responsibility, Past Chair of the Committee on Professional Ethics of the Association for Computing Machinery, and Past Chair of the Committee on Philosophy and Computers of the American Philosophical Association. Professor Bynum's academic degrees include a Ph.D. (CUNY), M.Phil. (CUNY), M.A. (Princeton), B.A. (U. of Delaware) - all in Philosophy - and a B.S. (U. of Delaware) in Chemistry. He has been a Fulbright Fellow (U. of Bristol, England), Danforth Fellow (Princeton), Woodrow Wilson Fellow (Princeton), Mellon Fellow (CUNY), and Dartmouth Fellow (Dartmouth College). On the topic of computing and human values, Professor Bynum has published articles and books, created conferences and workshops, given speeches and addresses, produced and hosted video programs, and developed an internationally influential web site. His other works include books, monographs, articles and reviews in logic, psychology, history of philosophy, and education. In 1968, he created the scholarly journal *Metaphilosophy*, which he edited for twenty-five years. In 1991, Professor Bynum was co-creator and co-director (with Walter Maner) of the National Conference on Computing and Values funded by the National Science Foundation; and in 1995 he was co-creator and co-director (with Simon Rogerson) of the ETHICOMP series of international computer ethics conferences held in England, Spain, the Netherlands, Italy, Poland, Portugal, Greece, Japan and China. He is currently co-chair of ETHICOMP 2010 to be held in Barcelona, Spain in March 2010. This text here is Pr. Bynum's Weizenbaum Address.

of topics, such as agent ethics, computer modeling, trust on the Internet, the role of information in reasoning and logic, e-democracy, informational realism, cyborg ethics, computing and creativity, digital hermeneutics, Buddhist computer ethics, and many more. Research centers, research professorships, and new journals have been established, and relevant organizations and conference series have been founded. For example, in 2004 the International Society for Computing and Philosophy (IACAP) was created, and just five years later it had divisions and yearly conferences in North America (NA-CAP), Europe (E-CAP), the Asian Pacific Region (AP-CAP) and Latin America (LA-CAP).

It is not surprising that these broad philosophical developments are occurring. An “Information Revolution” has been changing the world more rapidly – and more radically – than the Copernican Revolution and the Industrial Revolution. Of special interest is the fact that the Information Revolution is altering our understanding of human nature, the nature of society, and even the nature of the universe. In the past, such changes have spurred significant philosophical rethinking and creativity.

In the present essay, I briefly discuss the birth of the Information Revolution and then examine some relevant philosophical ideas of two “philosophers of the Information Age”: Norbert Wiener and Luciano Floridi. In the formative years of the Information Revolution, Wiener was a pioneer and one of the most important figures, both scientifically and philosophically. Today, Luciano Floridi and his project “the philosophy of information” (his term) have generated significant developments in this new field of philosophy, which already is so vast that only a few key ideas can be explored here. The topics covered below are limited to a handful of Wiener’s and Floridi’s ideas on human nature, artificial agents, the nature of society and the nature of the universe.

The Birth of the Information Revolution

Seeds of the Information Revolution were sown in the 19th century and early 20th century with scientific discoveries and technological developments regarding electromagnetic radiation, electricity, telegraph, telephone and radio. As important as these were, however, it was the simultaneous development of cybernetic science, information theory and electronic computers in the 1940s that marked the point at which the Information Revolution began to grow exponentially. Scientist/philosopher Norbert Wiener was (1) the primary creator of the science of cybernetics, (2) a major participant in the invention of electronic computers, and (3) a major contributor to the development of information theory. In addition, (4) he had remarkable foresight and the philosophical background to anticipate many of the profound social, ethical, and philosophical impacts of his

scientific and technological work and that of his colleagues. His scientific and technological achievements resulted from careful and sustained efforts over many years, but his ethical and philosophical contributions were essentially byproducts – comments and after-thoughts contained mostly in his social and ethical commentaries after World War II, including two short books – *The Human Use of Human Beings: Cybernetics and Society* (1950, 1954) and *God & Golem, Inc.: A Comment on Certain Points where Cybernetics Impinges on Religion* (1964).

During World War II, as part of the allied war effort, Wiener and several others accepted the assignment of designing a new kind of antiaircraft cannon to shoot down high-speed, maneuverable airplanes. The new cannon was supposed to (1) detect the presence of an airplane, (2) determine its position and velocity, (3) predict where it would be a few seconds later, (4) decide where to aim and when to fire the cannon, and finally, (5) carry out its own decisions. These steps were to take place almost instantaneously and without human intervention. Wiener and his team decided to use radar to detect and track airplanes and newly invented electronic computers to make predictions and decisions.

While working on this project, Wiener realized that the new science that he and his team were developing – which he later named “cybernetics” – led him to view his proposed cannon as strikingly similar, in certain ways, to human beings or intelligent, non-human animals. Thus, for example, the cannon would need “eyes” (radar) to take in information about nearby airplanes, “nerves” (wires) for internal communications among its parts, and “a brain” (computer) to coordinate the parts, make predictions, make decisions, and carry out those decisions. Wiener believed that, after the War, similar cybernetic machines, controlled by electronic computers, would be created for various economic, military and social purposes. He predicted that there would be an “Automatic Age” or “Second Industrial Revolution” with momentous social and ethical consequences.

During the War, Wiener collaborated or conferred regularly with physiologist Arturo Rosenblueth, mathematician John von Neumann, and logician Walter Pitts, who had been a student of philosopher Rudolf Carnap. Near the end of the war and immediately afterwards, this circle of thinkers was joined, in famous “Macy conversations”, by psychologist Kurt Lewin, anthropologists Gregory Bateson and Margaret Mead, economist Oskar Morgenstern, philosopher of science F. S. C. Northrup, and several other scholars. While discussing Wiener’s new science of cybernetics, this group of thinkers came to believe that “a better understanding of man and society [...] is offered by this new field.” (Wiener 1948, p. 39)

Soon after the War, Wiener published *Cybernetics: or Control and Communication in the Animal and the Machine* (1948). In that book, he explained various aspects of his new science; but he also gave cybernetic analyses of human and

animal physiology, pathology, psychology, language, social relationships and social institutions. Since then, *Cybernetics* has become a classic publication of the Information Age; and the scientific, technological and philosophical ideas that it contains have had a continuing impact upon a wide diversity of fields from computer design and information theory to biology, sociology and ethics.

In *Cybernetics*, Wiener occasionally made strong ethical claims. For example, he said:

It has long been clear to me that the modern ultra-rapid computing machine was in principle an ideal central nervous system to an apparatus for automatic control; and that its input and output need not be in the form of numbers or diagrams but might very well be, respectively, the readings of artificial sense organs, such as photoelectric cells or thermometers, and the performance of motors or solenoids. [...] Long before Nagasaki and the public awareness of the atomic bomb, it had occurred to me that we were here in the presence of another social potentiality of unheard-of importance for good and for evil. (1948, p. 36)

Such comments led several of Wiener's friends to encourage him to write a follow-up book focussed upon social and ethical implications of computing and the new science of cybernetics. As a result, in 1950 he published *The Human Use of Human Beings* where, using various examples and predictions, he explored some likely impacts of emerging information technologies upon things of value that people hold most dear, such as life, health, security, happiness, abilities, opportunities, freedom, and knowledge. Even today, in the age of the world-wide-web and the search for a "global information ethics", the methods and procedures that Wiener employed in that book can be used to identify, analyze and resolve social and ethical problems associated with information technologies of all kinds – including, for example, computers and computer networks; radio, television and telephones; news media and journalism; books and libraries. (See Bynum 2000, 2004, 2005, 2008) Given the breadth of his topics, as well as the applicability of his ideas and methods to every kind of information technology, the term "information ethics" is an apt name for his concerns in *The Human Use of Human Beings*, as well as his later book, *God and Golem, Inc.* (1964). Computer ethics, as it is typically understood today, is a subfield of Wiener's information ethics; and computer ethics topics that Wiener analyzed, or at least touched upon, decades ago (in Wiener 1948, 1950, 1954, 1961, 1964) include computers in the workplace, computing and security, computing for persons with disabilities, decision-making machines, computing and religion, information networks, globalization, virtual communities, teleworking, responsibilities of computer professionals,

merging of human bodies with machines, robot ethics, artificial intelligence, and a number of other issues.

Wiener's Universe

Philosophically, Norbert Wiener was a materialist, and in 1947 he made a related discovery regarding the fundamental role of information in the universe. While working on problems in a field that would later be called "information theory", he announced to some colleagues and graduate students at MIT that "information is entropy" (Rheingold 2000, Ch. 5), or more precisely that entropy is a measure of information which is "lost" (i.e., rendered unavailable to form objects or processes) in virtually every physical change. At the same time, in the manuscript of his book, *Cybernetics*, which was circulating among some of his scientific colleagues, he stated that information is physical, but it is not matter or energy. Thus, while discussing thinking as information processing in the brain, he wrote that the brain

does not secrete thought "as the liver does bile", as the earlier materialists claimed, nor does it put it out in the form of energy, as the muscle puts out its activity. Information is information, not matter or energy. No materialism which does not admit this can survive at the present day. (Wiener 1948, p. 155)

The amount of information "lost" in virtually every physical change is determined by the second law of thermodynamics. Such physical information is sometimes called "Shannon information" – named for Claude Shannon, who had been a student and then a colleague of Wiener's. Shannon and Wiener simultaneously discovered that entropy is a measure of physical information; and then Shannon went on, shortly thereafter, to develop a mathematical foundation for information theory.

Physical information is syntactical. It is the kind that is carried by radio signals, telephone lines, and TV cables. It is the kind that digital computers process and DNA encodes in the cells of every living thing. In Wiener's view, matter-energy and physical information are different, but neither exists without the other. So-called "physical objects and processes" are actually patterns of information encoded within an ever-changing flux of matter-energy. Thus every physical object or process is part of a creative "coming-to-be" and a destructive "fading away", as current information patterns erode and new ones emerge. The discovery by Wiener and Shannon that entropy is a measure of information provided a new way to understand the nature of physical objects and processes: to use today's language, all are "information objects" or "information processes" – an account of the nature of the universe worthy of the Information Age!

Even living things are information objects that store and process physical information in their genes and use that information to create the building blocks of life, such as DNA, RNA, proteins and amino acids. Nervous systems of animals take in, store and process physical information making motion, perception, emotion and thinking possible. And, as science writer Charles Seife recently noted,

Each creature on earth is a creature of information; information sits at the center of our cells, and information rattles around in our brains. ... Every particle in the universe, every electron, every atom, every particle not yet discovered, is packed with information [...] that can be transferred, processed, and dissipated. Each star in the universe, each one of the countless galaxies in the heavens, is packed full of information, information that can escape and travel. That information is always flowing, moving from place to place, spreading throughout the cosmos. (Seife 2006, p. 3)

The ultimate “fading away” of a physical object or process results from to an irreversible loss of physical information – an increase in entropy – and this is governed by the second law of thermodynamics. According to that law, essentially all physical changes decrease available physical information and, as a result, every object or process that ever comes into existence eventually will be destroyed. This includes virtually anything that a person might value, such as life, wealth, happiness, great works of art, magnificent architectural structures, cities, cultures, civilizations, the sun and moon and stars. None of these can survive the ultimate decay and destruction that results from entropy, because every physical thing in the universe is subject to the second law of thermodynamics. It is for this reason that Wiener considered entropy to be the greatest natural evil. He made the traditional distinction between “natural evil”, caused by the forces of nature (for example, earth quakes, volcanoes, diseases, floods, tornados, and physical decay), and “moral evil” (for example, human-caused death, injury, and pain). The ultimate natural evil is entropy – the loss of available physical information.

Wiener on Human Nature

Like every other physical entity in Wiener’s universe, human beings can be viewed as fundamentally informational. Like other animals, they are essentially patterns of physical information that persist through an ongoing exchange of matter-energy. So Wiener says of human beings,

We are but whirlpools in a river of ever-flowing water. We are not stuff that abides, but patterns that perpetuate themselves. (Wiener 1954, p. 96)

The individuality of the body is that of a flame [...] of a form rather than of a bit of substance. (Wiener 1954, p. 102)

Through metabolic processes like breathing, drinking, eating, perspiring, and so on, the matter-energy that makes up a person's body is constantly changing. In spite of this continuous exchange of atoms and molecules with the outside world, the pattern of Shannon information encoded within a person's body remains similar over time, changing only very gradually. This preserves, for an extended period, a person's life, functionality and personal identity. So a person is an "information object" whose personal identity is constituted by a persisting pattern of physical information, rather than particular atoms or molecules that happen incidentally to make up one's body at any given moment. Eventually, of course, the information pattern that constitutes a person's identity, and accounts for his or her ability to think, live and function, changes significantly. The inevitable results are aging, increasing disability and, ultimately, death – that is, the erosion and ultimately the destruction of the relevant physical information pattern that constitutes the essence of one's being.

A person's informational nature enables him or her to interact with other informational entities (objects and processes) in the surrounding environment. Thus, in *The Human Use of Human Beings*, Wiener said the following:

Information is a name for the content of what is exchanged with the outer world as we adjust to it, and make our adjustment felt upon it. The process of receiving and of using information is the process of our adjusting to the contingencies of the outer environment, and of our living effectively within that environment. The needs and the complexity of modern life make greater demands on this process of information than ever before. [...] To live effectively is to live with adequate information. Thus, communication and control belong to the essence of man's inner life, even as they belong to his life in society. (Wiener, 1954, pp. 17-18)

The information processing activities that occur within a human being are similar to those in other animals; and the specific physical structure of any given animal (including a person), according to Wiener, determines the nature and complexity of the information processing in which that animal can engage. Regarding human beings, Wiener emphasized their tremendous potential for learning and creative action made possible by human physiology. He often drew a contrast with other animals, such as insects:

I wish to show that the human individual, capable of vast learning and study, which may occupy about half of his life, is physically equipped, as the ant is not, for this capacity. Variety and possibility are inherent in the human sen-

sorium – and indeed are the key to man’s most noble flights – because variety and possibility belong to the very structure of the human organism. (Wiener, 1954, pp. 51-52)

Cybernetics takes the view that the structure of the machine or of the organism is an index of the performance that may be expected from it. The fact that the mechanical rigidity of the insect is such as to limit its intelligence while the mechanical fluidity of the human being provides for his almost indefinite intellectual expansion is highly relevant to the point of view of this book. (Wiener 1954, p. 57, italics in the original)

[...] man’s advantage over the rest of nature is that he has the physiological and hence the intellectual equipment to adapt himself to radical changes in his environment. The human species is strong only insofar as it takes advantage of the innate, adaptive, learning faculties that its physiological structure makes possible. (Wiener 1954, p. 58)

Wiener saw an intimate relationship between the internal information processing of human beings and the purpose of a human life. He considered flourishing as a person to be the overall purpose of life – flourishing in the sense of realizing one’s full human potential in variety and possibility of choice and action. To achieve this purpose, a person must engage in a diversity of information processing activities, such as perceiving, organizing, remembering, inferring, deciding, planning, acting, and so forth. Human flourishing, therefore, is utterly dependent upon information processing.

Wiener on Artificial Agents

In his book *Cybernetics*, and in other publications thereafter, Wiener viewed human beings and other animals as dynamic information processing systems whose component parts communicate internally, by means of feedback loops, thereby unifying the human or other animal into a single functioning entity. Also beginning with *Cybernetics*, Wiener assumed that, in the future, there will be many machines that function in a similar manner as dynamic information processing systems. Some machines will be able to make decisions and carry them out by themselves, and some machines will even learn from their past activities and adjust their future behavior accordingly. Wiener expressed concern that machines which learn and make decisions could pose major ethical risks. He worried that some people, blundering like sorcerers’ apprentices, might create artificial agents which they later could no longer control, and which will act on the basis of values that humans do not share. It is risky, he noted, to replace human judgment with machine decisions, and he cautioned that a prudent man

will not leap in where angels fear to tread, unless he is prepared to accept the punishment of the fallen angels. Neither will he calmly transfer to the machine made in his own image the responsibility for his choice of good and evil, without continuing to accept a full responsibility for that choice. (Wiener 1950, pp. 211-212)

the machine [...] which can learn and can make decisions on the basis of its learning, will in no way be obliged to make such decisions as we should have made, or will be acceptable to us. For the man who is not aware of this, to throw the problem of his responsibility on the machine, whether it can learn or not, is to cast his responsibility to the winds, and to find it coming back seated on the whirlwind. (Wiener 1950, p. 212)

Wiener noted that, to prevent this kind of disaster, the world will need ethical rules for artificial agents.

By 1950, with the publication of *The Human Use of Human Beings*, Wiener had assumed that machines will join humans as active participants in society. For example, some machines will participate along with humans in the vital activity of creating, sending and receiving the messages that constitute the “cement” that binds society together:

It is the thesis of this book that society can only be understood through a study of the messages and the communication facilities which belong to it; and that in the future development of these messages and communication facilities, messages between man and machines, between machines and man, and between machine and machine, are destined to play an ever-increasing part. (Wiener 1950, p. 9)

Wiener predicted, as well, that certain machines – namely, digital computers with robotic appendages – would someday participate in the workplace, replacing thousands of human factory workers, both blue collar and white collar. In addition, he foresaw artificial limbs and other body parts – cybernetic “prostheses” – that would be merged with human bodies to help persons with disabilities – or even to endow able-bodied persons with unprecedented powers. Today, we would say that Wiener envisioned societies in which cyborgs would play a significant role and would establish ethical policies to govern cyborg behavior.

In summary, Wiener foresaw a “Machine Age” or “Automatic Age” in which machines would be integrated into the social fabric, as well as the physical environment. They would create, send and receive messages; gather information from the external world; make decisions; take actions; reproduce themselves; and be merged with human bodies to create beings with vast new powers. These were not mere speculations, because Wiener himself, by the early 1960s, had already

designed or witnessed early versions of devices, such as game-playing machines (checkers, chess, war, business), artificial hands with motors controlled by a person's brain, and self-reproducing machines such as non-linear transducers. (See especially Wiener 1964.)

Wiener's predictions about decision-making machines led some people to ask whether such machines would be "alive"; but Wiener considered such questions to be semantic quibbles, rather than genuine scientific questions:

Now that certain analogies of behavior are being observed between the machine and the living organism, the problem as to whether the machine is alive or not is, for our purposes, semantic and we are at liberty to answer it one way or the other as best suits our convenience. (Wiener 1954, p. 32)

Similarly, he thought that answers to questions about machine consciousness, thinking, and purposes are pragmatic choices. He did believe, though, that questions about the "intellectual capacities" of machines, when appropriately formulated, could be genuinely scientific:

Cybernetics takes the view that the structure of the machine or of the organism is an index of the performance that may be expected from it. [...] Theoretically, if we could build a machine whose mechanical structure duplicated human physiology, then we could have a machine whose intellectual capacities would duplicate those of human beings. (Wiener 1954, p. 57, italics in the original)

In his 1964 book, *God and Golem, Inc.*, Wiener expressed skepticism that the physical structure of a machine could ever duplicate the complex structure of a human brain, because electronic components were too large and impossible to cram together like the neurons packed into a human brain. Perhaps today, given recent developments in the field of microcircuitry, he would be less skeptical.

In summary, by viewing both animals and cybernetic machines as dynamic systems with internal communications and feedback loops, Wiener came to see traditional distinctions between mechanism and vitalism, living and non-living, human and machine as blurry and pragmatic, rather than unbreachable metaphysical "walls" between kinds of beings.

Wiener on the Nature of Society

According to Wiener, just as humans, animals and certain machines can be viewed as dynamic, cybernetic entities, so communities and societies can be analyzed in a similar way:

It is certainly true that the social system is an organization like the individual; that it is bound together by a system of communication; and that it has a dy-

namics, in which circular processes of a feedback nature play an important part. (1948, p. 33)

According to Wiener, societies, communities and certain groups of people – even bee hives, ant colonies and certain herds of mammals – can all be viewed as second-order cybernetic systems, because their constituent parts are themselves cybernetic systems. The processing and flow of information are crucial to their nature and their successful functioning. Communication, said Wiener, is “the central phenomenon of society” (1950, p. 229). As a result, Wiener’s analyses of society often included discussions of communication networks and their importance. During his later life, there already existed a crude worldwide telecommunications network of telephone, telegraph, cable and radio facilities. Thus, although he died several years before the creation of the Internet, Wiener already had considered, in the 1950s and early 1960s, some social and ethical issues that are commonly associated with today’s Internet. For example, one of Wiener’s thought experiments concerned the possibility of people working on the job by using long-distance telecommunication facilities (today’s “teleworking” or “telecommuting”). He illustrated the possibility by imagining an architect in Europe who oversaw the construction of a building in America without ever leaving Europe. The imagined architect used telephones, telegrams and an early form of FAXing called “Ultrafax” to send and receive blueprints, photographs and instructions. (Wiener 1950, pp. 104-105 and 1954, p. 98)

A related telecommunications topic that Wiener briefly considered in his writings was the possibility of “virtual communities”, as we would call them today. As early as 1948, he noted that “Properly speaking, the community extends only so far as there extends an effectual transmission of information.” (1948, p. 184) And in 1954, he pointed out that:

Where a man’s word goes, and where his power of perception goes, to that point his control and in a sense his physical existence is extended. To see and to give commands to the whole world is almost the same as being everywhere. [...] Even now the transportation of messages serves to forward an extension of man’s senses and his capabilities of action from one end of the world to another. (1954, pp. 97-98)

It was clear to Wiener that long-distance telecommunication facilities, especially when they become more robust, will create many possibilities for people to cooperate together “virtually” (as we would say today), either on the job, or as members of groups and communities, or even as citizens participating in government. (See Wiener’s discussion of possible world government in Wiener 1954, p. 92.) Given Wiener’s view that the exchange of messages is “the glue that holds society together”, the world today must be rapidly “morphing” into a global society, be-

cause people on every continent are exchanging billions of messages daily using cell phones, email, “texting”, instant messaging, “tweeting”, blogging, video posting, and on and on.

Floridi’s “Philosophy of Information” Project

In the formative years of the Information Revolution in the 1940s and early 1950s, Norbert Wiener was the most important “philosopher for the Information Age”. By the middle of the 1990s, the Information Revolution, which Wiener had distantly envisioned fifty years before, was well underway. A vast diversity of information and communication artifacts had been invented and were proliferating across the globe: mainframe, mini, desktop, laptop, and palmtop computers; software; databases; word processors; spread sheets; electronic games; the Internet; email; and more. Robots had joined or replaced many human workers in factories; a number of people had become “telecommuters” working online from home; “virtual communities”, with geographically dispersed members, were multiplying; and decision-making machines were replacing certain people in medical centers, banks, airplane cockpits, classrooms, and on, and on. At the same time, influential physicists -- like John Wheeler at Princeton University (see Wheeler 1990) -- had begun to argue that the universe is made of information.

In this context, philosopher Luciano Floridi launched an ambitious project to create a new philosophical paradigm, which he named “The Philosophy of Information”. He believed that other paradigms in philosophy -- such as, analytic philosophy, phenomenology, existentialism, and so on -- had become “scholastic”, and therefore stagnant as intellectual enterprises:

Scholasticism, understood as an intellectual topology rather than a scholarly category, represents the inborn inertia of a conceptual system, when not its rampant resistance to innovation. It is institutionalized philosophy at its worst. [...] It manifests itself as a pedantic and often intolerant adherence to some discourse (teachings, methods, values, viewpoints, canons of authors, positions, theories, or selections of problems, etc.), set by a particular group (a philosopher, a school of thought, a movement, a trend, etc.), at the expense of alternatives, which are ignored or opposed. (Floridi 2002, p.125)

Philosophy, said Floridi,

can flourish only by constantly re-engineering itself. A philosophy that is not timely but timeless is not an impossible philosophia perennis, which claims universal validity over past and future intellectual positions, but a stagnant philosophy.... (Floridi 2002, p. 128)

As an alternative to scholastic philosophical systems and communities, Floridi set for himself the ambitious task of creating a paradigm that would someday become part of the “bedrock” of philosophy (*philosophia prima*). At the heart of his project was to be the concept of information, a concept with multiple meanings, and also

a concept as fundamental and important as being, knowledge, life, intelligence, meaning, or good and evil – all pivotal concepts with which it is interdependent – and so equally worthy of autonomous investigation. It is also a more impoverished concept, in terms of which the others can be expressed and interrelated, when not defined. (Floridi 2002, p. 134)

Floridi had available to him many new methods and conceptual resources – developed after Wiener’s pioneering days – from computer science, system theory, logic, linguistics, semantics, artificial intelligence, philosophy of mind, philosophy of science, and theoretical physics. During the past decade, he has applied these new resources to great effect, and he has been joined by a number of philosophical colleagues and graduate students. As a result, his “philosophy of information” project has grown and matured into a broad research program addressing a wide diversity of philosophical questions. These range from the deceptively simple question, “What is information?”, to issues like the nature and ethics of artificial agents, the foundation and uniqueness of computer ethics, the semantics of scientific models, the nature and role of artificial companions in a human life, the informational nature of the universe, symbol grounding and consciousness, the role of information in reasoning and logic, and many, many more. In the limited space of the present essay, it is possible to cover only a few relevant topics, and so the sections below focus only upon a handful of Floridi’s many ideas on the nature of the universe, human nature, artificial agents, and the nature of society. In the closing section, relevant ideas from Floridi and Wiener are compared.

Floridi on Human Nature and the Nature of the Universe

Floridi’s philosophical method is that of constructionism. A constructionist adopts the view that ultimate reality (a Kantian would call it the “noumenal” world of “things-in-themselves”) is unknowable – a “black box” into which we can never see. Ultimate reality provides certain affordances and imposes certain constraints upon our experiences, observations, and experiments, but we are forever unable to know how and why it does so. The best that we can do is to construct models of reality, or parts thereof. Knowledge, truth and semantics apply to our models, and not to ultimate reality, since we can never know what that is. We can know how our models work, though, because we have constructed them. The world in which we live (Kant’s phenomenal world) is the sum total of

our models of reality. It follows that if we significantly change the objects and/or processes within our models, then we live in a different world. It is important to note, however, that this is not a version of relativism, because models can be compared with regard to their ability to accommodate the constraints and affordances of the unknowable “ultimate reality”. Floridi’s models are constructed using “the method of abstraction”, which he and his colleague J.W. Sanders adapted from Formal Methods in computer science. Their philosophical method involves the selection of a set of “observables” at a given “level of abstraction”. By attributing certain “behaviors” to the observables, one builds a model of the entity being analyzed, and this can be tested against our experiences, observations and experiments. The best models are those which most successfully achieve “informativeness, coherence, elegance, explanatory power, consistency, predictive power, etc.” (Floridi 2004)

In the article, “Informational Realism” (Floridi 2004 [for a revised, very technical version see also Floridi 2008a]), Floridi provides a rigorously developed argument that, at a certain level of abstraction, all objects in the universe are data structures composed of “mind-independent points of lack of uniformity.” This is true of every object in the universe, according to Floridi, and so

The outcome is informational realism, the view that the world is the totality of informational objects dynamically interacting with each other. (Floridi 2004, *italics in the original*)

At the informational level of abstraction, therefore, every existing entity is a “data structure” – an “informational object”. This includes human beings, and according to Floridi, seeing humans as informational objects is

part of a wide and influential informational turn, a fourth revolution in the long process of reassessing humanity’s fundamental nature and role in the universe. We are not immobile, at the center of the universe (Copernicus); we are not unnaturally distinct and different from the rest of the animal world (Darwin); we are far from being entirely transparent to ourselves (Freud). We are now slowly accepting the idea that we might be informational organisms among many agents (Turing), [...] not so dramatically different from clever, engineered artifacts, sharing with them a global environment that is ultimately made of information [...]. (Floridi 2008b, p. 654)

In addition to being composed of “informational objects dynamically interacting”, the universe, according to Floridi, is fundamentally good, and that goodness is independent of human moral judgments. This is a major metaphysical assumption of Floridi’s new “macroethics” (his term), which he calls INFORMATION ETHICS. [SMALL CAPS are used here to distinguish Floridi’s theory from the more

general field of information ethics in the broadest sense.] According to Floridi his INFORMATION ETHICS theory is similar to virtue ethics, deontology, consequentialism, and contractualism in that it is intended to be applicable to all ethical situations. On the other hand, INFORMATION ETHICS is different from these traditional theories because it is not intended to replace them, but rather to supplement them with further ethical considerations that can sometimes be overridden by more traditional ethical concerns (Floridi 2005).

What are the fundamental components of INFORMATION ETHICS? As explained above, according to Floridi, every existing entity in the universe, when viewed from a certain “level of abstraction”, is an “informational object” with a characteristic data structure that constitutes its very nature. For this reason, Floridi calls the universe, considered as a whole, “the infosphere”. Each entity in the infosphere can be damaged or destroyed by altering its characteristic data structure, thereby preventing it from “flourishing”. Such damage or destruction Floridi calls “entropy”, which results in the “empoverishment of the infosphere”. Entropy, therefore, constitutes evil that should be avoided or minimized. [It is important to note here that Floridi has borrowed the term “entropy” from physics, but what he means by that term is not thermodynamic entropy subject to the second law of thermodynamics.]

Since Floridian entropy is an evil that should be avoided or minimized, he developed four “fundamental principles” of INFORMATION ETHICS:

0. entropy ought not to be caused in the infosphere (null law)
1. entropy ought to be prevented in the infosphere
2. entropy ought to be removed from the infosphere
3. the flourishing of informational entities as well as the whole infosphere ought to be promoted by preserving, cultivating and enriching their properties

By viewing every existing entity as an “informational object” with at least a minimal moral worth, Floridi shifts the focus of ethical considerations away from actions, characters, and values of human agents, and toward the “evil” (damage, dissolution, destruction) that is suffered by objects in the infosphere. With this approach, every existing entity – humans, other animals, organizations, plants, non-living artifacts, electronic objects in cyberspace, pieces of intellectual property, stones, Platonic abstractions, possible beings, vanished civilizations – all can be interpreted as potential agents that affect other entities, and as potential patients that are affected by other entities. Thus, Floridi’s INFORMATION ETHICS can be described as a “patient-based” non-anthropocentric ethical theory instead of the traditional “agent-based” anthropocentric ethical theories like deontology, contractualism, consequentialism, and virtue theory.

A number of Floridi's critics have argued that his metaphysical presupposition that the universe is fundamentally good is unnecessary and unjustified. In reply, Floridi has responded that

The actual issue is whether Goodness and Being (capitals meant) might be two sides of the same concept, as Evil and Non-Being might be. [...] the reader sufficiently acquainted with the history of Western philosophy need not be told about classic thinkers, including Plato, Aristotle, Plotinus, Augustine, Aquinas and Spinoza, who have elaborated and defended in various ways this fundamental equation. For Plato, for example, Goodness and Being are intimately connected. Plato's universe is value-ridden at its very roots: value is there from the start, not imposed upon it by a rather late-coming, new mammalian species of animals, as if before evolution had the chance of hitting upon homo sapiens the universe were a value-neutral reality, devoid of any moral worth. (Floridi 2008b)

According to Floridi, seeing something in a particular way – that is, adopting a particular level of abstraction in order to model it – always has a purpose. If that purpose is fulfilled well and fruitfully, then one is justified in taking that perspective. By viewing the universe as fundamentally good, consisting of informational objects, their relationships and processes, Floridi is able to accomplish at least three major things:

1. Make sense of the awe and respect that one feels when confronted by the vast, beautiful universe (like Toaism, Buddhism, Platonism, Aristotelianism, Stocism, Spinozian philosophy, etc.).
2. Develop a way to apply moral reasoning and accountability to the rapidly increasing number of artificial agents (robots, webbots, cyborgs, virtual communities, etc.) that are being created in our midst by the billions. (See below.)
3. Develop a way to understand distributed moral accountability within complex social agents like corporations, organizations, virtual communities, governments, etc. (See below.)

Floridi and Sanders on Artificial Agents

In the late 1940s, when Wiener first predicted that there would be an “Automatic Age”, with machines that make decisions and even learn from their past “experiences”, many people considered his prediction to be science fiction or, perhaps, an extravagant exaggeration. Some even considered him to be an eccentric old man engaged in flights of fantasy (Conway and Siegelman 2005). Today, Wiener's prediction has been fulfilled so dramatically that it now seems to be an understatement. Some of today's “artificial agents” are indeed robotic hardware de-

vices like those that Wiener envisioned, but they also are software entities, such as webbots that “crawl” through computer networks, and “softbots” that reside in our laptops, cell phones, iPods, hearing aids, digital cameras, home appliances, and so on. Artificial agents now do many different things: correct spelling, delete “spam”, find and erase computer viruses, control nuclear power plants, fly airplanes, control rail networks on which high-speed trains travel, launch military missiles in the heat of battle, make medical decisions and inject patients with medicine, perform delicate surgery on living human beings, and on, and on. So Wiener was indeed correct when he predicted that a future society (ours!) would need ethical rules and procedures to govern artificial agents. This has become a major need of our time, one that is growing rapidly as artificial agents proliferate exponentially.

Floridi and his colleague J.W. Sanders (see Floridi and Sanders 1999, 2001, 2004), addressed this growing contemporary challenge. To develop an appropriate ethical theory for artificial agents, they sought to achieve the following three aims (Floridi and Sanders 2004):

1. Provide “an effective characterization” of an agent.
2. Provide an appropriate account of good and evil that artificial agents are capable of bringing about.
3. Provide an explanation of how and why to hold artificial agents morally accountable, even if they are “mindless” and thus without mental states.

What is an agent? – A human being is a paradigm example of an agent, so Floridi and Sanders’ characterization of an agent needs to fit humans. In addition, however, it must also fit softbots, robots, and other artificial agents such as virtual communities – even organizations, corporations and governments. The characterization that Floridi and Sanders developed includes three criteria that an entity must meet to be an agent (Floridi and Sanders 2004):

- i. Interactivity: The agent and its environment can act upon each other.
- ii. Autonomy: The agent is able to change its own state independently of its interactions with the environment. An agent, therefore, must have at least two states and be “decoupled” to some extent from its environment.
- iii. Adaptability: The agent’s interactions with the environment can change the transition rules by which it changes state; that is, the agent’s capacity to change its own states can evolve because of its own past interactions. (For humans or animals, we say that they “learn from their experiences”.)

Floridi and Sanders note that one must always specify the level of abstraction at which one views an entity before one can determine whether it is an agent, be-

cause something could be seen as an agent at one level of abstraction, but not at a different level. For example, given our everyday view of what a human being is, a person certainly is an agent at that level of abstraction; but viewed simply as a physical object situated in a particular region of space-time the person is not an agent at that level of abstraction.

What is a moral agent? – After providing a careful characterization of an agent, Floridi and Sanders introduced a criterion for being a moral agent:

An action is said to be morally qualifiable if and only if it can cause moral good or evil. An agent is said to be a moral agent if and only if it is capable of morally qualifiable action. (Floridi and Sanders 2004)

It is important to note here that the term “action”, as Floridi and Sanders use it, does not require that the agent have mental states like beliefs or intentions, or capacities like “free will” (whatever one may mean by that controversial term). An action in the relevant sense is simply an activity in which the agent causes an effect. For example, a computer “worm” that gets into a computer system in a nuclear power plant, makes and executes a decision, and thereby causes a catastrophe, has engaged in a morally evil action, even if it was completely “mindless” with no knowledge or intentions. Similarly, a computerized medical device that saves a patient’s life by injecting the patient with appropriate medicine in a crisis has engaged in a morally good action, even though it is completely “mindless”.

Distinguishing moral accountability from moral responsibility – Some critics of Floridi and Sanders have argued that it is inappropriate to call the activities of “mindless” beings moral or immoral, because such beings cannot be held responsible for what they do. Floridi and Sanders replied, however, that this objection fails to distinguish between holding an agent accountable – therefore subject to censure – on the one hand, and holding it responsible – therefore subject to blame and praise, punishment and reward – on the other hand:

Human moral agents who break accepted conventions are censured in various ways of which the main alternatives are: (a) mild social censure with the aim of changing and monitoring behavior; (b) isolation, with similar aims; (c) death. What would be the consequences of our approach for artificial moral agents?

Preserving consistency between human and artificial moral agents [in Cyberspace], we are led to contemplate the following analogous steps for the censure of immoral artificial agents: (a) monitoring and modification (i.e. ‘maintenance’); (b) removal to a disconnected component of Cyberspace; (c) deletion from Cyberspace (without backup). (Floridi and Sanders 2004, p. 376)

The distinction between accountability and responsibility not only enables the extension of ethical considerations to artificial agents like robots and webbots,

it also enables us to better understand the appropriateness of holding young children accountable for their behavior even if they are not yet morally responsible. In addition,

It facilitates the discussion of the morality of agents not only in Cyberspace but also in the biosphere – where animals can be considered moral agents [...] – and in what we have called contexts of ‘distributed morality’, where social and legal agents [like corporations, organizations and governments] can now qualify as moral agents. The great advantage is a better grasp of the moral discourse in non-human contexts [...] . (Floridi and Sanders 2004, p. 377)

Floridi on the Nature of Society

As explained above, Floridi sees the universe as the totality of informational objects dynamically interacting with each other – the “infosphere” – and this includes human beings as well as all other biological organisms; plus robots and all other artificial agents; every physical object; and even “Platonic” entities that do not exist in physical space-time. Also included in the infosphere are “second order” informational entities whose parts/members are themselves informational objects. These include families, organizations, corporations, communities (small and large), governments and whole societies. In summary, then, according to Floridi human beings are informational objects dynamically interacting with a world of other informational objects, and societies are large and complex dynamic second-order informational objects whose members are themselves dynamic informational objects.

In the past, people have not seen themselves as informational objects, nor have they considered most of the objects in their environment – homes, cars, clothing, cookware, highways, etc. – to be dynamic informational objects (even though, at the informational level of abstraction, that is what they are). Nevertheless, according to Floridi (2007), people soon will think of themselves as informational objects, because information and communication technologies are quickly being incorporated into everyday objects, making them interactive with us and with each other. They soon will be so profoundly re-engineered that Floridi has coined a term, “reontologisation”, for the re-engineering process. Our “reontologized” cookware will communicate with us and with each other as it cooks our food. Our refrigerators will learn our dietary preferences and notify us, or the grocery service, when we are running out of certain foods. Our physical belongings will stay in touch with us over the web when we travel, and many objects in our environment will learn from their “experiences”, make decisions and take actions accordingly. Our “reontologised” environment will be a complex society of humans, artificial agents and everyday objects, all wirelessly intercommunicating.

With such “ubiquitous computing” or “ambient intelligence” throughout the environment, the world will seem almost alive to us, and today’s distinction between being “off-line” in the “real world” and being “on-line” in cyberspace will vanish. We will think of ourselves as connected informational organisms – Floridi coined the term “inforgs” for this – living in a complex and dynamic society of other informational objects, both biological and artificial. When this process is completed, according to Floridi, “we shall increasingly feel deprived, excluded, handicapped or poor to the point of paralysis and psychological trauma whenever we are disconnected from the infosphere, like fish out of water.” (Floridi 2007) In such a circumstance, the “digital divide” between informationally advantaged societies and “have-not” societies will become a huge social and ethical chasm!

Concluding Remarks

The Information Revolution, though still in its early stages, has already led to a vast number of changes in the world – physical, political, economic, social, scientific, psychological, philosophical, and so on. In the present essay, because of limited space, it has been possible to explore only a few of the many philosophical contributions of Norbert Wiener and Luciano Floridi regarding the role and importance of information in the universe, human nature, artificial agents, and society generally. Both of these thinkers are among the most influential “philosophers for the Information Age”. Wiener, as a pioneer, helped to create much of the science and technology that made the Information Revolution possible, and he also used remarkable foresight to “see distantly” a number of important social, ethical and philosophical impacts. Half a century later, equipped with many new tools and examples from computer science, system theory, logic, linguistics, semantics, artificial intelligence, philosophy of mind, philosophy of science, and theoretical physics, Floridi has spearheaded an ambitious program to place the concept of information into the bedrock of philosophy. Wiener’s vision proved to be remarkable, and Floridi’s ambitious project appears to be headed for admirable success.

Upon first sight, the relevant views of Wiener and Floridi seem very similar. For example, both take the universe to be essentially informational – made of dynamically interacting informational objects and processes. Both consider human beings to be informational objects; and both say that “entropy” is a very important evil in the world. First appearances, however, are deceptive because Wiener was a materialist and Floridi is a Platonist, and they interpret information and entropy very differently.

As indicated above, Wiener considered the information of which the universe is made to be physical and therefore subject to the laws of physics. Such informa-

tion is syntactic, rather than semantic. It is the kind that is carried in radio waves, telephone lines and TV cables; and it is encoded in the DNA of every living entity and carried by every subatomic particle. Human beings and all other physical entities, according to Wiener, are dynamic patterns of such information that persist for a time, but gradually evolve, or even erode and dissipate. Entropy is a measure of that erosion and dissipation.

According to Floridi, the information of which the universe is composed is semantic, rather than merely syntactical; and it is non-physical, and thus it does not obey the laws of physics like the second law of thermodynamics. It is Platonic information – “mind-independent points of lack of uniformity” (Floridi 2008a) – that comprise the data structures not only of familiar objects like tables and chairs, humans and computers, but also of Platonic entities like possible beings, intellectual property, and unwritten stories from vanished civilizations.

Reconciling Wiener and Floridi? – Compared to Floridi’s philosophy of information, Wiener’s is very incomplete – a byproduct of his scientific research projects and social concerns. He did not work on his philosophy of information as a carefully developed research project, like Floridi has done. And it is quite possible that most of Wiener’s views, as explained above, are reconcilable with Floridi’s. For example, given Wiener’s belief that thermodynamic entropy – the “loss” of available physical information – is the greatest evil in the universe, he could accept the idea that, to the extent that the universe contains available physical information for creating good objects and processes, the universe is basically good. In addition, nothing in his philosophy of information would prevent his acceptance of Floridi’s analysis of artificial agents and artificial evil. In addition, Wiener could readily agree with Floridi’s analysis of today’s society that “ubiquitous computing” is placing more and more objects “online”, removing the distinction between “online” and “offline” existence, and turning us all into “inforgs”.

On the other hand, Wiener was a confirmed materialist, and Floridi is a committed Platonist, so philosophically their positions appear to be, in the end, unreconcilable. Perhaps a contemporary materialist philosopher will soon develop a rival philosophy of information as powerful and well developed as Floridi’s. In metaphysics, the more things change, the more they remain the same!

References

Bynum Terrell Ward, (2000), "The Foundation of Computer Ethics", *Computers and Society*, 30:2, 6-13.

Bynum Terrell Ward, (2004), "Ethical Challenges to Citizens of the 'Automatic Age': Norbert Wiener on the Information Society", *Journal of Information, Communication and Ethics in Society*, 2:2, 65-74.

Bynum Terrell Ward, (2005), "Norbert Wiener's Vision: the Impact of the 'Automatic Age' on our Moral Lives, in Cavalier, R.J. , Ed., *The Impact of the Internet on our Moral Lives*, Albany, NY: State University of New York Press, 11-25.

Bynum Terrell Ward, (2008), "Norbert Wiener and the Rise of Information Ethics", in van den Hoven, W.J. and Weckert, J. , Eds., *Moral Philosophy and Information Technology*, Cambridge, UK: Cambridge University Press, 8-25.

Bynum, Terrell Ward and James H. Moor, Eds. (1998, 2000), *The Digital Phoenix: How Computers are Changing Philosophy*, Oxford, UK: Blackwell. Revised edition 2000.

Conway Flo and Jim Siegelman, (2005), *Dark Hero of the Information Age: In Search of Norbert Wiener, the Father of Cybernetics*, New York: Basic Books.

Floridi Luciano, (2002), "What is the Philosophy of Information?", in James H. Moor and Terrell Ward Bynum, Eds., *CyberPhilosophy: The Intersection of Computing and Philosophy*, Oxford, UK: Blackwell, 115-138. (A Metaphilosophy Monograph)

Floridi Luciano, (2004), "Informational Realism" in John Weckert and Yeslam Al-Saggaf, Eds., *Computers and Philosophy 2003: Selected Papers from the Computers and Philosophy Conference CAP2003*, Australian Computer Society, Conferences in Research and Practice in Information Technology, pp. 7-12.

Floridi, Luciano (2005), "Information Ethics, Its Nature and Scope", *Computers and Society*, 36:3, 21-36.

Floridi Luciano, (2007), "A Look into the Future Impact of ICT on Our Lives", *The Information Society*, 23: 1, 59-64.

Floridi Luciano, (2008a), "A Defence of Informational Structural Realism", *Synthese*, 161: 2, 219-253.

Floridi Luciano, (2008b), "Artificial Intelligence's New Frontier: Artificial Companions and the Fourth Revolution", *Metaphilosophy*, 39: 4-5, 651-655.

Floridi Luciano and J.W. Sanders, (1999), "Entropy as Evil in Information Ethics", in L.Floridi, Ed., *Etica & Politica*, special issue on Computer Ethics, 1: 2.

Floridi Luciano and J.W. Sanders, (2001), "Artificial Evil and the Foundation of Computer Ethics", *Ethics and Information Technology*, 3: 1, 55-66.

Floridi Luciano and J.W. Sanders, (2004), "On the Morality of Artificial Agents", *Minds and Machines*, 14: 3, 349-379.

Rheingold Howard, (2000), *Tools for Thought*, Cambridge, MA: MIT Press, First MIT Press Edition. (Originally published in 1985 by Simon & Schuster.)

Seife Charles, (2006), *Decoding the Universe: How the New Science of Information is Explaining Everything in the Cosmos, from Our Brains to Black Holes*, New York: Viking Penguin.

Wheeler John A., (1990), "Information, Physics, Quantum: The Search for Links" in W. H. Zureck, editor, *Complexity, Entropy, and the Physics of Information*, Redwood City, CA: Addison Wesley.

Wiener Norbert, (1948, 1961), *Cybernetics: or Control and Communication in the Animal and the Machine*, New York: Technology Press, John Wiley and Sons. Second Edition, Cambridge, Massachusetts: MIT Press, 1961.

Wiener Norbert, (1950, 1954), *The Human Use of Human Beings*, Houghton Mifflin, 1950. Second Edition Revised, Garden City, NY: Doubleday Anchor, 1954.

Wiener Norbert, (1964), *God & Golem, Inc.*, MIT Press. Published in the UK in the same year by Chapman & Hall.

Acknowledgments

I am grateful to Southern Connecticut State University for funding and sabbatical leave time in support of this project. And I am especially thankful to Luciano Floridi, Krystyna Gorniak-Kocikowska, and Richard Volkman for reading a draft of this essay and making helpful suggestions for improvement. Any remaining shortcomings are, of course, my own.

Paradoxes of contemporary knowledge: between invention, creation, information and control

Laymert Garcia Dos Santos*,
Konstantinos Karachalios**
and Jürgen Partenheimer***

It seems nowadays to be apparent that the total technological and economical acceleration of the global capitalism, that followed the 'cybernetic turn' launched in the 1970s, has had as consequence a reconfiguration not only of the domain of work, but also of knowledge and life as such. Today, more than ever, what is requested from us is not a mere encyclopaedic or even critical knowledge, but one that enables us to act in function of two supplementary operations that the North-American inventor Buckminster Fuller characterised as information gathering and problem solving.

As a matter of fact, always faster and more flexible communication technologies, combined with an increasing digitalization of knowledge, with Internet, with the establishment of databases and the invention of search engines like Google have raised access to information to one of the most crucial questions, contributing to a radical transformation of what we understand knowledge to be. Because, even if it remains always important to be in possession of knowledge, it is more important

* Prof. Dr. Laymert Garcia Dos Santos

Born Itapopolis, Brazil in 1948. Essayist and university lecturer at the University of Campinas (São Paulo) in the field of the sociology of technology. He lived in Paris in the 1970s and gained a PhD in Information Technology at the University of Paris 7. He has published on the topics of culture, technology and the environment, and has written books, among them *Tempo de Ensaio* (Essay Time). He translated Etienne de la Boétie's *Abhandlung über die Freiwillige Knechtschaft* and presented studies on Lautréamont.

Dr. Konstantinos Karachalios

Born Athens in 1954. Gained a PhD in Physics and Engineering on nuclear reactor safety. He works in the European Patent Office in the field of external relations and is co-author of the work *Scenarios for the Future*. Publications deal with the topics of culture, politics, science and technology. Keynote speaker at the Computer, Freedom and Privacy Conference 2008 in Yale. He is founder and chairman of the Palladion, a cultural foundation established by immigrants in Germany.

Prof. Dr. Jürgen Partenheimer Interdisciplinary visual artist, poet. Biography see <http://www.jurgenpartenheimer.com/Html/Vita/Vita.php>

to be capable of 'treating' it with a maximum of efficacy and according to the complex context that is needed for its actualization. This means that the 'know what' is supplanted by the 'know how'.

In this sense, knowledge presupposes new links between culture and technology, because it is based on the relationships man-machine. However, in its new status, the 'know how' is like a double-edged sword that may be directed either towards management or towards transformation. In this latter case, knowledge becomes rather related to invention and creation.

No wonder thus, that the push for 'innovation' has emerged as a ubiquitous mantra, covering not only technical, but virtually all aspects of life, and linked to the notion of 'creative destruction'¹. Coupled with this comes the triumphal expansion of western-based Intellectual Property (IP) norms to the rest of the world (by means of the Trade Related aspects of IP rights System; TRIPS²), without doubt in all terms a singularly successful normative process in world history. Its declared rationale: to promote and reward innovative activities, by no means restricted to the technical domain anymore, as the practice of patenting business methods (including financial derivatives related ones), of natural material (such as genes) and of (so called transgenic) living plants and animals³ shows. A new globalised landscape of codified knowledge generation and control emerges, with very serious political and societal consequences.

The acceleration and excess of super modernity assist and provoke self-referential and mindless automatisms and contribute to imperative dilemmas to the regimes of intellectual property in a digital era that tries to balance the legitimate needs of both creators and the public.

- Striving for excess innovation turns to self-referential competitive and mindless innovation: "a culture is as good as its woods".

Additionally, innovation for the sake of innovation is increasingly provoked by inherent peer-pressure of digital presence, the desire of 'belonging', 'the transparency race'.

- 'Copy and Paste' generation of users who profit from technical deficiencies and lack of copyright security in the net. Poor technical standard provides unlimited opportunities to copyright violation.

1. This notion appears first in the writings of *M. Bakunin*, *F. Nietzsche* and *W. Sombart*. *J. Schumpeter* goes so far as to say that the «process of creative destruction is the essential fact about capitalism». The computers industry is often used as the best example to illustrate the innovative and in the same time destructive power of this process.

2. http://www.wto.org/english/tratop_e/TRIPS_e/TRIPS_e.htm.

3. For a literary fiction around these issues, see *Michael Crichton, Next*, Harper Collins, NY, 2006.

- Access to knowledge provided on a restricted, distorted or manipulated level assists a new kind of scattered knowledge leading to a system of 'patchwork-knowledge', a contemporary bastardization of Levi Strauss' technique of 'bricolage', however, based on random availability of pre-conditioned knowledge instead of conscious, intuitive or creative use of knowledge.
- IP in reference to art in the net offers a wide range of discussion, where the domain of invention, content and creativity are reduced to a forum of supply and demand, a market place of commodities, where intellectual property and artistic property are robbed of their true artistic neutrality and cut off of their intrinsic context of creativity and ideas.

Several questions emerge:

In a time that other western-sponsored values (democracy, autonomy of individual e.tc.) does not seem to penetrate that easily, what is the reason for this unique framing success of TRIPS and what could be the consequences?

To what extent are the current regimes of 'intellectual property' favouring and/or blocking access to this performing knowledge ?

How are art, as a domain of creativity, and techno-science, as a domain of invention, translated into 'know how' ?

Because the total acceleration does not give us anymore the time to critically reflect, to what extent does the operational knowledge suggest new relationships between intelligence and intuition on the one hand, and between perception and memory on the other?

What strategies and tools are needed to counteract the paralysing effects of acceleration on the projects of individual and political autonomy ?

Is there a path from current power relations, dominated by corporate strategies and geopolitics and the large scale misery they perpetuate, towards a new landscape of individual, social and political autonomy and a new culture of geopoetics leading to less alienation and to more global fairness ?

As the reader may feel, the following texts do not lead to a synthesis and are deliberately not conclusive. They create, instead, a sort of patchwork of issues, problems, questions raised in three perspectives which sometimes overlap, sometimes resonate, sometimes create a synergy from our impressions, sensations, thinking and findings related to the new ways of acquiring and "processing" knowledge in complex contemporary societies. Thus, in our views, access to knowledge nowadays is not a matter, or at least not only a matter of democratization of existing cultural matrixes and cultural heritage through new technological devices and media. For better or worse and, perhaps, for both at the same time, the so-called "knowledge society" and

“knowledge economy” seem to express a new historical formation requiring a new ontology and a new epistemology.

Two short stories as an insight to the new status of knowledge

Two real episodes which have taken place in June 2007, with almost one week of interval between them, could perhaps help us grasp why one has to explore the intuition that performative knowledge is acquiring a new status in contemporary social life.

The first event happened to a young man working in IT in a Brazilian bank. Spotted by head-hunters of a leading corporation of the telecommunications sector, he was invited to submit himself to a recruiting process for a new job. Accepting to play the game, he was introduced into an office, and then was given three minutes to answer the following question: “How many golf balls are in the air at the present moment?” The person who elaborated such a bizarre and apparently absurd quiz-challenge left the room immediately after, and the candidate was left alone with his laptop. Under such pressure, the young man started thinking how he could tackle the problem and said to himself: ‘Golf is a sport one plays only in day-light. This means that, in latitudinal and longitudinal terms, only a range of countries in the Northern and Southern hemispheres have to be taken into account.’ So, with the help of his laptop he drew the lines of his research field and developed a sort of strategy to find a probabilistic response to his question. Three minutes later, the corporation officer came in again. But the young man was not allowed to announce his findings; instead, he was asked to talk about all the steps made to reach a final figure – a procedure suggesting that what mattered was the process, not the “product”. After explaining how he had designed his problem, he was told he would be informed soon on the company’s decision. A week later, the corporation official contacted him, and told there were five divisions willing to hire his services and that it was up to him to say if and where he wanted to work. The young man chose the corporation’s Intelligence Division.

The second short story concerns a mature Sweden lady, also working in the financial markets. Since her husband was forced to undergo the transplant of spine marrow due to a rare leukaemia three years earlier, she had been following closely all the developments of his disease and treatment, including the infections to which his body was very vulnerable. And then, he suddenly got unusual bacteria, forcing him to go to one of the leading hospitals in Sao Paulo, for treatment. During a long conversation with the doctor who was in charge of her husband since he fell ill, the lady was provided with the most important data on his case and informed that a blood analysis would tell which sort of bacteria were causing such a big damage. Taking into consideration all the relevant information she had

gathered as keen observer of her husband's illness, and combining them with the new elements the doctor had dispensed, she went to Google, laid down the key words and started the search, asking the good questions. Before the hospital laboratories had released the outcomes of the blood test, she became aware that the infection was due to rare bacteria which were found in only six patients all over the world, and to which there was no medicine available. Yet, surprisingly, the test contradicted those findings, stating that the infection was related to another bacteria. Unfortunately, the laboratory and the physicians that had relied on the test were wrong: the patient died some days later. A new test had already proven that the search performed by the lady had delivered the correct results.

These two short stories seem important in the context of our discussion because they are enlightening and emblematic expressions of how the issue of performative knowledge might be viewed today. Information gathering and problem solving were here the crucial factors that were seriously taken in consideration in a new kind of relationship between man and machine. But also, one has to realize how sensitive were the issues of speed, time, intuition, perception, memory and context awareness in the whole knowledge process. As if the logic of recombination had acquired a new quality and had become operational in a new way...

Appropriation and control of codified knowledge in the techno-fix era

Before using the term 'Intellectual Property' for the following argumentation, we should remind that it is a neologism, used in expert circles since the mid 1960s⁴, coinciding with the creation of WIPO, in an attempt to link together, for political and administrative reasons, activity domains as different as patented technology, copyrighted literature, branding of companies and trade marks, e.tc. However, only with the Bay-Dole act⁵ in 1980, the term became 'intellectual property' and its acronym 'IP' notorious. The political and cultural consequences of this semantic invention, which helped considerably to turn not a priori sympathetic monopolies into a quasi human right (the 'it's my house' doctrine), proved rather important⁶.

4. Literally, it appeared already in 1845 in a US court decision (Davoll et al. v. Brown).

5. This US law permits universities or non-profit institutions to acquire ownership of an invention, the research for which was funded by public money.

6. *Peter Drahos, John Braithwaite: 'Information Feudalism: Who owns the knowledge economy?', W.W Norton & Co, NY, 2002; see also interview of Paul A. David to the EPO for the Scenarios for the Future project <http://www.epo.org/topics/patent-system/scenarios-for-the-future/interviews.html>.*

Indeed, what followed only 14 years later was breathtaking. The historical singularity of the world-wide expansion of these, essentially western, values and tools by means of the mentioned TRIPS agreement, one of the three WTO pillars (the other two being free trade in goods and services), can be explained neither by linguistic creativity nor by conspiracy theories. If this quite detailed, pervasive and robust normative system for inclusion of monopolies into a free trade agreement⁷, has been adopted by the vast majority of states in the world⁸, there must be not only negative reasons (coercion, manipulation, lack of experience of negotiators e.tc.), but also very strong 'positive' ones, at least within the imaginary of the 'norm-receivers' or at least among their elites, too.

It seems that - contrary to the solemn declarations of its political leaders - the west did not export aspiration for democracy and political emancipation, the other antinomic pillar of its imaginary to the rest of the world, but rather fascination by technology and its IT gadgets⁹. Even the most fanatic enemies of democratic ideas and aspirations, like fundamentalist religious movements in east and west, seem to have no ideological barriers towards hi-tech and seem quite fit in using it for their purposes. On the other hand, states with a totally different historical and cultural background to the west, like China, fully adhere to the prevailing growth model. Echoing the dilemma of many of his western counterparts, President Hu Jintao said (at the opening of the communist's party five-yearly congress on 12 Oct. 2008) that, compared to environment, "maintaining China's rapid economic growth remains the ruling Communist party's 'top priority'"¹⁰.

In the current debates about climate change and how to deal with the huge challenges, there is a lot of talk about technology transfer, in particular for mitigation strategies. Although the level of trust between the negotiating parties is not too high¹¹, everybody seems to trust into the omnipotential power of techno-science to solve all problems. Amazingly, although the imminent and generally accepted perception of threat could lead to a questioning of some fundamental assumptions of the western development model, it seems to rather reinforce the triumph of its golden boy, of the techno-fix ideology.

7. See http://www.wto.org/english/tratop_e/trips_e/trips_e.htm to get an impression of depth and level of detail.

8. From the 192 members of UN, 153, including China, India, Brazil and South Africa, have acceded to WTO and thus also to TRIPS, as a mandatory condition.

9. *Cornelius Castoriadis*, 'The West and the third world', Speech at Irakleion University, Crete, Greece, March 1991.

10. According to a report by FT on 15 October 2008.

11. Poisonous, according to a high ranking negotiator.

Of course, nobody is against human development; the point is that the prevailing western production and consumption model has conquered the minds of humanity, with the exception of some politically marginalised enclaves. And exporting the mobility model of Los Angeles to the rest of the world may not be the best for our lives and for the planetary ecological equilibrium¹². Although the signs on the wall become more and more clear, the debate about alternative development models is - to put it diplomatically - stagnating. This is an additional quite alarming sign: of the incapacity of our societies to think politically in broader terms, what leads to what Castoriadis calls a “fragmented world”¹³.

The information revolution with its massive ‘networking’ and increase in the global exchange of information does not seem capable of reversing this trend. Against its original promises, it seems to rather reinforce fragmentation and isolation. As Paul Virilio observed early in the process¹⁴, the information and communication technologies are causing the replacement of the geocentric perspective by an extreme egocentric, individualist one. Each one of us feels like the centre of the universe, we don’t need to move, the world must come to us, instantaneously. We are separated from real space and time and increasingly confuse real dimensions and feelings with their simulations.

Further, the transformation of data into information and then into knowledge – information that can be utilised to build capabilities – is far from being straightforward. This raises additional issues:

As information becomes increasingly abundant, what knowledge has value?

If the rules around access, management, production and ownership of knowledge are not chosen properly, more information could even equal less knowledge – and less innovation. This is the knowledge paradox¹⁵.

As codified knowledge is at the heart of production and control of cutting-edge technology that is a not only a major driving force for societal development, but

12. Even within organisations like the Transatlantic Consumer Dialogue TACD, which includes many critical streams and voices, the overall validity of the prevailing consumption mode does not seem to be systematically challenged (for topics and statements, see <http://www.tacd.org/index2.htm>).

13. *Cornelius Castoriadis* in the essay “L’époque du Conformisme généralisé”, *Les carrefours du Labyrinthe*, tome 3: *Le monde morcelé*. Seuil, Paris, 1990.

14. *Paul Virilio*, ‘Speed and Politics: An Essay on Dromology’, New York: Semiotext, 1977.

15. *Scenarios for the Future*, EPO, 2007.

also a key instrument for hegemony¹⁶, fierce battles around its ownership are intensifying.

Globally acting corporations, state governments and civil society groups will fight against or ally¹⁷ to each other, with often changing fronts, in an essentially Orwellian, nowhere leading conflict for control of codified knowledge, fitting to the mentioned information gathering and problem-solving purposes and its imaginary of omnipotence. This would be a new global Trojan war with IP as the Trojan horse and techno-fix as Helena, the object of desire¹⁸. The current, escalating conflicts around patented technology within most international ICT standardization bodies, with their societal and geopolitical implications, are clear signs of what is coming.

EPO's scenarios

These battles could get out of control with serious geopolitical consequences, as the governance structure of the global knowledge economy looks less compact than a block of Swiss cheese, leaving too much room for individual optimization strategies. As failures occur in ever shorter intervals and the externalities are rising steeply, the temptation to search for convenient scapegoats rises, too. Due to the growing intensity of the conflicts and volatility across all vital systems, attempts for quick and partial fixes, without looking at the essential issues at stake, in particular the social and geopolitical gaps in the perception of fairness and equity, will have no real impact.

The European Patent Office, as a broad regional patent office with a considerable human, cultural and scientific background and potential decided that retreat is not a good option now. One of the visible expressions of the will to understand, to reach out and to engage more pro-actively into the public debate is surely EPO's "Scenarios for the Future" project. The scenarios have the battles around knowledge, its production, appropriation and control at their heart and can be used to examine how these battles could shape the future world architecture or test specific policies and strategies. Four key scenarios, each one based on the predominance of a mega-driving force, relentlessly and uncompromisingly following its own logic¹⁹, have been published by the EPO as the outcome of this 3 years long, collective undertaking²⁰. Their findings in a nut shell:

16. Paul Virilio «History progresses at the speed of its weapons systems», supra 14.

17. For the bizarre form of these alliances, see Open Dynamic Alliances of the IGF.

18. According to an alternative narrative, Helena was not at all in Troy, so this was a bogus war.

19. Veritable «machines de guerre» according to G. Deleuze.

20. <http://www.epo.org/topics/patent-system/scenarios-for-the-future.html?banner=about> 1.

If the unfettered free market logic survives the collapse of the Washington Consensus and continues to rule, then globally sourcing and acting western or westernised corporations will establish themselves as the dominant driver. Current incumbents and IP title holders as well as 'domesticated' newcomers will aggressively assert their monopoly rights throughout the globe. Knowledge continues to be 'my house', trespassers receiving an 'armed response'.

When geopolitics becomes the dominant driver what counts would no longer be profits or shareholder value, but national interests and hegemony. At the same time, this could be the story of a boomerang effect striking today's dominant players, eventually leading to new block building and a deeply fragmented and polarised political and knowledge landscape. Transferring knowledge between the block division lines is seen as unpatriotic.

Were status-quo critical, societal groups to gain significant political influence, this could lead to diminishing societal trust and growing criticism towards the current system for appropriation of knowledge, including the IP system, causing its gradual erosion. Only a few patents would be granted and enforcement would be avoided where possible. But is the predictable fallback to secrecy going to reinforce knowledge generation and sharing?

Finally, the logic of technology (doability) is not necessarily the one of capital (accumulation). In a world where techno-science, driven by techno-politicians and futurists, takes the lead, too much control could be seen as a break rather than an accelerator. A split of the patent system across industrial sectors could thus occur. IP rights in several technological sectors would lose their most powerful weapon, the monopoly right, enforcement in its classic sense would become obsolete there.

These stories are not EPO's visions or strategies, since the actors and their acts are largely beyond anyone's reach. Further, as analytical abstractions (in reality all four driving mega-forces are working in parallel) they are not meant to be exact forecasts or previsions. However, this approach is proving useful, because it helps many persons and institutions around the world to better understand the extremely complex processes happening already now, in front of our eyes. The scenarios achieve also to extend the reach of the public debate beyond the "Market Rules" horizons, which have been usually taken for granted, and hint at inherent dangers looming large in all examined configurations.

However, there is one more aspect that is fundamental and critical. It concerns a specific type of science, assuming that it still exists: the one that is not confined within the logic of technology or of the market, but is linked to an emancipa-

tive process, commonly viewed as the search for 'truth'²¹. Further, it is almost instinctively understandable that by remaining within an unchallengeable context and respecting a priori fixed boundaries one cannot produce this type of science. Some philosophers go so far to argue that this process is a 'truth procedure' which cannot take place outside the broader process of political emancipation²², 'art' being another such a process²³. And like all emancipative processes, this is a double-edged sword.

I know well that the most beautiful only exists in dreams; however, it is about the paths that lead to the dreams...

Ludwig Hohl²⁴

Discontinuity, Paradox & Precision

Current models of reflection with regard to an increasingly questionable access to knowledge

The revolution in information and computer technologies created an overabundance of information available, whereas operational knowledge is at risk to be reduced to mere key words. Consequently, new relationships within this context are bound to develop a kind of meta-language of easily recognizable terms, which do not address a deepening of understanding, but are rather directed to instant and readily adaptable or feasible recognition. Creative Intelligence and intuition become superfluous, attentive perception is reduced to simple 'Gestalt' verification and memory becomes obsolete, for it will be out-sourced from human capacity to self-administered and monitored databases.

The acceleration of history corresponds to a vast multiplication of events presenting us with the problem of excess: "What is new is that we need to give a meaning to the present overabundance of events. For it is our need to understand the whole of the present that makes it difficult for us to give meaning to the recent past."²⁵ Disengagement from the past by excess of present meaning corresponds

21. See 'Managing Science: Is the Cudos still in Place?', K. Karachalios, *Biotechnology Journal*, 2008, 3, 306-310.

22. E.g. Badiou, Castoriades, Rancières.

23. Partenheimer Jürgen «Das freiwillige Exil der Unabhängigkeit - Über den Unterschied zwischen Branding und Werk, Kommerz und Freiheit» Edition Copernicus, Nov. 2007.

24. Ludwig Hohl, X, *Traum und Träume*, No.19, Die Notizen, Frankfurt, 1984, S.671.

25. Marc Augé, *Non-Places*, London-New York, 1995, p.28 ff.

to a situation Augé defines as “super modernity - to express its essential quality: excess.”²⁶

Memory and oblivion, historical awareness as reflective and constructive references for the here and now are at stake, as described by the Swiss philosopher Ludwig Hohl: “The possible choices of a version of the present will become apparent in whichever past it chooses to contemplate. And the first question is, whether it contemplates any past at all. If it contemplates nothing, then it is itself nothing”²⁷.

Science and art pertain to creative research and visionary invention in an absolute and at times in a labyrinthine, quixotical sense. Science per se as well as art practice search for quintessential answers, which are generally not determined by pragmatic interest but by investigative curiosity, the quest for understanding, for the new, the unknown. They are not necessarily linked to the common search for ‘truth’ either; their goal is rather oriented towards independent reflective analysis, which may result in creative inventions to be applied or to stay independent. This approach is marked by profound experience and identity rather than strategic interest in commercial lifestyle and its exploitation.

In regards to ‘super modernity’, the common search for truth as an idealistic approach has been confronted with a critical connotation of pragmatic codification (abuse), which Michel Foucault already pointed out in the late seventies: “We live in a society, which by and large is on a ‘truth-trip’, generating and introducing a discourse with truth capacity and, once accepted as such, endorsed with a specific power at command. The appointment of ‘truthful’ discourse (which by the way is due to constant change) is one of the basic problems of the Western world. The history of ‘truth’ i.e. power, which is the proper quality of a discourse acknowledged as true, is still waiting to be written.”²⁸

On Art

The artist goes on a journey, leaves the pier for the ocean into the expanse of the sea, “to where no one can go any further” as Rilke²⁹ put it. Going beyond the boundaries discloses a necessary sanctuary for the artist – crossing the narrow gates as a sign of independence. The setting out on a journey is the declared approach toward the unfamiliar, toward the uncertainty, as Maurice Blanchot characterized it: “The more the world asserts itself as the future and the ‘high noon’,

26. Marc Auge, *ibid.* p. 29.

27. Ludwig Hohl, *Die Notizen*, Frankfurt, 1984, S.

28. Michel Foucault, *Short Cuts*, Berlin, 2001, p.93.

29. Rainer Maria Rilke, *Briefe*, Frankfurt, 1985.

where everything has value, where the whole is made complete under the rule of man and for his use, the more it seems as though art must descend to the point at which nothing yet has any meaning”.³⁰ That sounds very much like Rilke and at the same time it leads to the contents and works of “Pier and Ocean”, Mondrian’s theme of exiles during his time in New York.

In a conversation on art, Samuel Beckett³¹ describes the essential state of uncertainty, that state in which nothing as yet makes sense, and uses the expression “unavailable terms”. He declares this “absence of terms”, as he also calls it, as the true territory of art, for it has no system of references, no conventional “aesthetic automatisms!” Absence of terms in the digital era – how terrible! And yet – what a gift! “At once powerful and powerless, powerful precisely through one’s powerlessness” (Jacques Derrida³².) Thus, free from all aesthetic automatisms, the artist in his work can draw on all the boundless possibilities of failure, and only in so doing make what is in fact a beginning. The identity of this experience expresses nothing other than following one’s vision, one’s imagination, and trusting in it.

The market

The available commercial structure of a hegemonic globalization that operates in a capitalist way, confronts the essential “unavailable terms” of productively creative independence. The economic, capitalist behavior of the West and increasingly also that of eastern society, absorbs culture in the productive, commercialized development of controlled systems and values it according to the principles of supply and demand. This evaluation however, destroys the essential ‘neutrality’ of culture placing it under the ‘presence of terms’ in the unequivocal authority of verifiable rules.

The high-speed course of this controlled system into which art is assimilated and which changes art dramatically, is discussed by means of the concept ‘cultural turn’: “If we wish to understand our contemporary society, we must first understand how culture was colonized by capital and what kind of devastating effects such colonization has on the politics and goals of emancipation (intellectual property). Global capitalism is cannibalizing culture, all culture worldwide, and is thereby threatening the basic foundations of society, such that it is causing the

30. Maurice Blanchot, *Das Unzerstörbare*, München/Wien, 1991, p.78.

31. Samuel Beckett, *Three Dialogues*, New York, 2006, p.560.

32. Jacques Derrida, *Die Wahrheit in der Malerei*, Wien, 1992.

disintegration of the cultural riches of our world in an increasingly intensive and mad economic instrumentalisation".³³

The absurdity of the cultural turn in the art world is a good example of this instrumentalisation. In the general 'branding' interests of the market, art auction houses worldwide are rising to the status of a new, open and interactive gallery system, in which a bourgeoisie motivated by social prestige rather than passion for art has taken over the bidding. Anyone can access information on the Internet as to 'winners' and 'losers' at any time. Points are distributed; brochures of the 'winners' are printed and sent as future investment fund incentives. Value manipulates desire! Record values devour artists; the hunt for the highest figure characterizes the unmistakably absurd quality of the trend.

Casino moguls and software cowboys buy into this and pin their modern icons in the cabins of their yachts, or hide them away from the eyes of the true admirers as they increase in value in the dark rooms of their safes. In the Bermuda triangle of art biennials, art auctions and art fairs, the little vessels race from one island to another, stocking up with art, just as in former times the Dutch fell victim to the tulip bulb frenzy in a period of socio-market-influenced hysteria. The outcome is history, the stock market crashed.

The economy is proving itself to be a dangerous temptation in the territory of culture. It is lowering culture to the level of speculative goods in the exchange of commercial interests and reveals the decadence of the cultural turn. "With what is each instance of decadence identified?" asked Nietzsche in his cultural writings and he answered "the fact that life is no longer lived to the full", the fact that it is no longer a matter of the "depths" of the life that is lived, as Joyce put it. It's obviously not about a life that is lived at all, but about strategies that submit to the market within recognized principles. Pragmatism versus Idealism. The guise of things.

Strategies of subversion

Discontinuity as a reflective attitude of contemplation. A sudden or anticipated standstill, "time-out" as an important corrective, which pertains to the slowing down of the state of affairs that leads to contemplation, re-evaluation and distance as a free decisive choice or by forced circumstance (current financial crisis). The artist/scientist should take that distance and it is through this distance the she intervenes.

33. Laymert Garcia dos Santos, "High-tech plundering, biodiversity and cultural erosion: The case of Brazil".

Paradox. It pertains to surprising situations or is a logical demonstration of absurdities. A true statement that leads to contradiction heightens attention and causes reflection and offers brisk surprise, triggers unconventional solutions.

Precision in order to ensure clarity, conceptual awareness, a focused mind, anticipation and perseverance.

Classic yet Contemporary Gender Norm: is ICT an amplifier of gender bias?

Ryoko Asai*

Centre for Business Information Ethics,
Meiji University

Abstract

This study analyses the social impact of online communication from a gender perspective. Since the mid-1990s, Internet business has flourished around the world. Nowadays, many companies come to focus attention on effectiveness of communication between diverse stakeholders with information and communication technology (ICT). Rapid development and permeation of the Internet allowed people to use the interactive media for non-business purposes in a convenient and comfortable manner. Based on this circumstance, many companies have sought to build a new business model utilizing online human networks generated by Internet users. Individual users have acquired simplified ways to communicate online with their friends and, even, the general public called UGM/CGM which include social network services (SNSs) and blogs. What is communicated using CGM sometimes contains important information for business like consumer needs, tastes and trends as well as consumers' evaluation of the quality of products and services companies provide, which affect behavior of other consumers. Consequently, many companies have formulated new marketing strategies to handle the change of consumer behavior through CGM and utilize the media as thrust of their economic performance. However, information communicated through the Internet seems sometimes to have such huge power that anyone cannot control the social influence of it. In other words, websites are full of information related to gender, sometimes more directly than in the real world. On the other hand, it is difficult for many people to find and realize gender issues behind own communication through the Internet. Because both information senders and receivers usually communicate each other regarding own social norm and social experiences in gendered societies as quite natural. This study has an overview of the current situation in online communication, and describes characteristics of information through the Internet. In addition, this study examines the social effect of information generated from general users in the perspective of gender.

* *Ryoko Asai* is a Research Fellow at the Centre for Business Information Ethics at Meiji University in Tokyo, Japan. Its focus is on gender and computing, and gender and information communication technology in the context of working life.

Keywords: UGM/CGM; marketing strategy; gender; online communication; gender norm

Introduction

Since the mid-1990s, Internet business has flourished around the world. In 2009, Japan holds about 90 million Internet users, including those who connect to the Internet with their mobiles, and the penetration rate of the Internet has exceeded 75.3 percent of the Japanese population (Ministry of Internal Affairs and Communications 2008). Though penetration rate of the Internet maintains an upward trend in every country, Japan is one of the highest rated countries. Many companies come to focus attention on effectiveness of communication between diverse stakeholders with information and communication technology (ICT). Nowadays, about 85 percent of Japanese companies have their own websites, about 30 percent of them run an advertisement on the Internet, and one in ten companies has started by setting up own blogs or social networking services (SNSs) sites for their customers. The Internet use by companies has become steadily established in our societies. Also, rapid development and permeation of the Internet allowed people to use the interactive media for non-business purposes in a convenient and comfortable manner. In contemporary Japan, a large number of people join in online games or SNSs. As of the end of March 2006, about 7 million people were registered as social network service members, and 8.68 million were registered as bloggers. In addition, enrollment in online gaming has recently increased to 28 million, with 1.75 million of these gamers estimated to be members of fee-based game sites.

The internet makes interactive and many-to-many communication possible for users. And the deployment of ICT allows people to increase their opportunities to make use of ICT as a means of communication on a daily basis. The development of ICT provides simplified information transmission capacity such as blog and SNS, and improves utility value of information by finding and sorting data in a manner that is appropriate and effective. In other words, blogs, SNSs, word of mouth sites and video-sharing sites may have an extensive impact on business reliability or reputation, and eventually on business performance to a greater or lesser extent (Toyama, Kishi, and Murata 2008). General internet users are actualized as potential and unignorable customers for many companies. Against this underground, many companies seek to develop new business models that utilize networking created by internet or internet users. Websites for personal use, which are full of information submitted by people all over the world, are referred to as User Generated Media (UGM) or Consumer Generated Media (CGM). UGM/CGM contains various kinds of media contents which are publicly available and created by end-users. The scale of market produced by UGM/CGM have

been growing rapidly with popularization of blog and SNS that involve RSS (RDF Site Summary) in its own tool, and RSS enables sending users information in an appropriate and efficient manner. In recent years, consumers have become increasingly concerned about the safety of products, and also have easily searched information through the Internet. Therefore companies inevitably pay a lot of attention to Internet and its users.

Today many companies, especially big companies, embark on some aggressive strategies to respond to change of consumers' behaviour through CGM, and introduce CGM into their existing marketing strategies. For example, in the US, the Coca-Cola Company has already started the own CGM website ("my coke") which enables internet users to post 45 seconds user's video related to Coca-Cola products and also has opened its own virtual space ("CC metro") to communicate users since 2006. This video-sharing site has run a contest for users posted videos and the winner has determined by a vote of its users. This kind of sites and events apparently seems to plan to grab users' attention and gain in popularity, in a sense it is true, but in fact the company is able to manage information for the company, products and customers' responses via with own websites. This content was originally designed to introduce aggressive internet users into the company's marketing strategy taking a hint from user's sharing video ("Diet Coke + Mentos") posted on torrent sites. In terms of interactive communication between companies and internet users (potential customers), this UGM/CGM marketing strategy, sometimes recognized as "Web 2.0" marketing, has some efficacy in sharing and using information among them. However, information might inevitably spread beyond particular companies and use and the effects are being felt throughout the entire society. This study offers an overview of the current situation in UGM/CGM, and describes characteristics of information through online communication. In addition, this study examines the social effect of information generated from general users in the perspective of gender.

Gender Issues on UGM/CGM

Sharing "Interesting" Information through the Internet

Nowadays, many companies come to focus attention on effectiveness of communication between diverse stakeholders with ICT. Therefore, many companies have sought to build a new business model using online human networks generated by Internet users. Individual users have acquired simplified ways to communicate online with their friends and, even, the general public called consumer generated media (CGM) which include SNSs and blogs. What is communicated using CGM sometimes contains important information for business-like consumer needs, tastes and trends as well as consumers' evaluation of the qual-

ity of products and services companies provide, which affect behaviour of other consumers. Companies have already become aware that communication between consumers using CGM may have a significant influence over companies' economic performance as well as trust and reputation. Consequently, many companies have formulated new marketing strategies to handle the change of consumer behaviour through CGM and use the media as thrust of their economic performance. According to the recent survey conducted by the Ministry of Economy, Trade and Industry in Japan, the size of CGM-driven economic effect, including affiliate advertisings and online word of mouth/mouse, was estimated about JPY 2 trillion in 2007, and this size was expected to expand in the future (Ministry of Economy, Trade and Industry 2008). As CGM websites evolve, some consumers who was one-time consisted an audience of unilateral communication media, turned out to be information transmitters on online bilateral media and have great influence over others' consumption behaviour.

Another recent survey conducted by the Japanese government shows the most popular purposes of using the Internet among non-business Internet users are browsing websites and blogs of companies and public organizations (56.8 percent), and sending or receiving e-mails (54.5 percent). And more than half of the internet users as a whole have experience of shopping and electronic financial transactions on the Internet. Therefore many companies trying to take every opportunity have promoted the use of the Internet rapidly. Now, 99 percent of Japanese companies have used the Internet to a greater or lesser extent (Ministry of Internal Affairs and Communications, 2009). In addition, to process information from consumers more efficiently many companies have set up communicable websites within cyberspace, such as blogs, SNSs, or virtual contents. Even if a company is not interested in the Internet or CGM, consumers positively use the Internet to choose appropriate products. This indicates that power relationships between consumers and companies in terms of consumption are gradually changing. Hence, companies have to intensify taking effect of CGM into account. Under the present circumstances of proceeding diversification and multichannel service of Media, not everyone watches the same TV program at the same time like they used to. In addition, the next generation network is developing. Therefore we can easily expect to transform companies' marketing fields from mass media into the cyberspace via the Internet. In other words, UGM/CGM will become much more vigorous and meaningful for companies and consumers than it has been for quite a while. However, information communicated via CGM seems sometimes to have such huge power that the social influence this power has cannot be controlled.

Striking Metaphors for “Wii” and “PLAYSTATION 3”

When Nintendo released their new game console the Wii™ in 2006, a user-posted video which showed functional differences between Wii and Sony PLAYSTATION3™ on the most famous torrent site YouTube™ (Figure 1). This has attracted attention of a large number of people; the video logged more than three million hits and put over about thirty thousand users’ comments in one and a-half years. Some of its’ viewers had claimed that the video was somewhat offensive or harmful information to people underage, and the video was eliminated from the site by the site administrator. Nevertheless, this video has been posted on YouTube again and other video-sharing websites by its users. What is the controversial and popular video? This video allows us to lean gender bias on the Internet from its metaphors and the way of expression.



Figure 1. User posted video “Wii vs. PLAYSTATION3” Figure 2. Official Wii-fit promotional video

Generally, companies pay the closest attention to their advertisements in order to receive great public attention. Especially, they take particular care in relation to violent and sexual scenes in their promotional video, posters and leaflets. For example, official Wii-fit promotional video made by Nintendo shows that men and women, young and old pleasantly play with Wii-fit (Figure 2). The official video gives people some ideas that the new game console Wii-fit aims at everyone without regard to sex or age and it also helps everyone succeed in maintaining healthy. However, the user’s video above Wii instance explains their differences from PLAYSTATION3 by making an analogy to two types of women (Figure 1). Based on their specification, Wii is symbolized by an active, glamorous and blonde woman who is pictured right. On the other hand, PLAYSTATION3 is represented as an owlsh, stout and brunet woman. That is to say, both consoles have been expressed by stereotyped and gendered images of women. By contrast, in the real world, these very images are recognised as gender bias to be eliminated as soon as possible. Though the video never asks its viewers which is acceptable to play with, many of them seem to prefer the Wii woman to the PLAYSTATION3

woman according to the comment field. Many comments from the torrent site's users watching this video were favourably-disposed toward it, such as "it's so cool..." or "Wii girl is hot!" and so on. In addition, there are a number of sexually-oriented roughhousing and reviews in the comment field. Some the comments complimented expressiveness of the video creator. This goes far beyond viewers comments. In Japan, a newsletter published in a famous Japanese MBA school have offered a broad introduction of this video as a good education material of comparative advertising. It is interesting to note that this video was created by game users from the point of consumers buying game consoles, neither by Nintendo nor by Sony Computer Entertainment.

Through the Internet, this information of new game console has immediately spread all over the world and made many users take notice. Namely, information generated from consumers spread quickly across the world sometimes through online communication such as blogs and SNSs, and sometimes through interaction between users and media with online retrieval functions. In many marketing studies, the Internet tends, to be recognized as so a powerful and epoch-making communication tool that changes conventional consumption behaviour. These studies emphasize that the characteristics of the Internet, such as accumulateness of data, immediacy and simultaneity, make possible to stabilize and reinforce the relationship between consumers and companies (Ishii and Atsumi 2002). This indication is based on positive outlook on dual characteristics of ICT: the use of ICT changes existing structures greatly, whereas ICT maintains and strengthens existing frameworks. However, information generated from users as well as consumers sometimes have a serious negative impact on a company's performance.

Social Risks of Information Generated from Consumers

Intensifying communication through the Internet has been changing power relationships between a sender and a receiver of information. That is to say, the primacy of a sender over a receiver is gradually disappearing. This is so because a receiver is easily able to transform into a sender with tools providing Internet access. Today, as the way of transmission of information simplifies, ICT could provide everyone with an opportunity to make many-to-many communication (Ikeda 2000). In other words, online communication without information control by mass media, companies and public organisations makes possible for individual users to send to other users direct messages, and create new fashion independent of mass media perceived as the powerful dispatcher of fashion. So far, as a logical consequence, many companies come to step up efforts to introduce new measure of online communication into their own marketing formula in order to get an edge on the severe competition. Eventually, the user's Wii video case is one of the

most successful cases in which CGM exert an influence on product image held by people.

On the other hand, without any check and control by companies and public bodies, sometimes users' direct information might have negative influence on other users' consumption behaviour. Therefore companies need to prepare to deal with the risk of dysfunctional communication such as false accusation, defamation or harmful rumour, whenever they try to involve CGM in their own business activities. However, why were many users so much drawn to the Wii video posted by users? Actually, we have to be aware of social risks behind the background where many people shared and resonated interests in or attentions to the video with others, such as "cool", "funny", "nice commercial" and so on. In other words, there are social risks that online communities as well as sources to send information could become much more functional as social entities providing social norms to people. This kind of social risks relating to social norms appears prominently in dissemination of information such as the Wii video. Today, with deeply penetration of the Internet and the increasing number of online communication users, people cannot help being affected by information in the cyberspace, sometimes by interests and normative consciousness derived from online communication. On the other hand, information in the cyberspace is generated on the basis of information in the real space. Now ICT as a new communication technology has served to blur boundaries used to be drawn sharply between the real space and the cyberspace.

Value and Norm provided by online communication

"Interesting" Information Based on Social Norm

Every information receivers not always comprehend and interpret sending information in the same way. According to Ikeda (2000), sending message itself has no meaning, and the meaning of the message are provided by information receivers. In order to perceive and construe the message, receivers extract "information as *Vorstellung*" from their own existing knowledge, experiences and reasoning capability. And receivers finally find meaning of the message through the extraction process. As compared with communication in the real space, online communication is established with many people, including anonymous person and users who are just browsing. Therefore, meanings of online message become diversified depending on the number of users those have experience, normative consciousness and social value in their own way.

Nevertheless, in the Wii video case the video became quickly popular across borders. In contrast to the remarkable implementation of gender equality policies in the real space, almost of its viewers' reactions for the video appeared to be

positive ones. Behind this background, it would appear that people gathering the particular online community share the social code and context which serve as a foundation for communication among them. In other words, such positive reactions for the Wii video mean that its viewers share gender norm and gender role necessary to interpret the meaning of the video message. The user's Wii video expressed game consoles by making an analogy to two characteristic women. The video creator never asked its viewers which was acceptable to play with, and seemed to leave a decision of which playmates was better up to the viewers. However, almost all of the viewers would construe the meaning of the video based on gender image sustained over long periods in gendered societies intentionally or unintentionally. Therefore, in societies where the Wii video posted by users is recognised as a good metaphor and did become popular among people, the video using gender image of women would whet consumers' appetite for playing with "an active, glamorous and blonde woman (Wii)", and also would provide gender norm with viewers through browsers, sometimes without their realizing it. "Interesting" information is created using the social norm and social value of people.

Gender Issues behind ICT

In the case of strengthening persuasive power of the message using by stereotyped gender image of women, the message tends to easily stick out in people's mind, and exert an influence on transforming people's social norm to a greater or lesser extent (Kendall 2000). Therefore, online communication using gender image or gender norm through the Internet would maintain and strengthen existing gender order both in the cyberspace and in the real space. This aspect also points out the ICT's dual characters. One is that utilisation of ICT brings about a major change in existing frameworks of marketing, society and so on. The other is that use of ICT maintains and, even, strengthens existing frameworks. As is well known, this trend in the cyberspace is opposed to the global trend away from gender discrimination and gender inequalities.

To be fettered by conventional gender roles and particular gender images could convey the wrong message to people at times. Of course many gender studies have developed this kind of argument. In these studies, researchers have pointed up inconsistencies in not only gender discrimination between men and women, but also inequalities among same sexes. In addition, based on this argument, some researchers have amplified on gender and ICT. In other words, ICT which make possible to communicate online has originally presented some gender problems. For example, most sophisticated computer users and many programmers or system administrators and generally people working with computers, are men (NRI 1999; Kendall 2000). Additionally, there are not only gender issues but also

racial, economic and age problems. In gendered societies, ICT would contribute to the gendered division of labour and it is assigned gendered symbolic values (Lohan 2001). And given the ICT's dual characters, the gendered division and gendered symbolic values would socially maintain and strengthen through the use of ICT. Of course, online communication using ICT inevitably faces to gender problem in the developing process and also in the uses of it.

However, it is difficult for many people to find and realize gender issues behind their own communication through the Internet. Because both information senders and receivers usually communicate with each other regarding their own social norm and social experiences in gendered societies as quite natural. Nowadays, as communication globally spreads on the Internet, gender as well as power relation among people needs to be recognized as the social risk for individuals, organizations and societies. On the other hand, when we try to eliminate one gender issue aggressively such as sex-conscious or pornography websites, we will face problems relating to social power relationships such as violation of freedom of expression and infringement on freedom of speech. Thus we have to address social risk assessment appropriately and ethically.

Conclusion

Interpretation of information is different to different people. Though a message in itself may have no meaning, information recipients will understand its meaning based on their own knowledge, experiences and reasoning abilities. When many people have a positive impression of the Wii video, such as "interesting" or "it is so understandable", they might share social contexts such as social roles or social norms which are prerequisites to communication. In other words, people who show positive reactions to the video may have similar gender images, and in some situations, they can force others to feel the same way they do. Thus, online communities can function as "social reality" imprinting people with specific norms, and have some kind of influence on people's formation of gender norms.

In the real world, gender norm has been a cause of various inequalities or disadvantages for women. Therefore many countries have implemented policies for building a gender-equal society over a long time. However, in the cyberspace, these policies seem to function less well. The origin of this problem can be found in the early days of computer technology. Namely, ICT has been positioned as a masculine technology since the gendered societies were commonplace. Research and development of ICT have involved few women. Even now, female technical experts are still a minority as female Internet users (Asai and Murata 2009). Under this "masculine" situation, communication naturally assumes a gendered character. Based on ICT's dual characters, online communication contains some

risks which maintain and strengthen gender inequalities. Moreover, among people who share similar norms or experiences, it is difficult to them to find their own gender bias.

In the real world, many countries articulated a vision for realising gender-equal societies in which unjustifiable gender norm including gender bias is eliminated. Actually, People never pointedly talk about appearances of women on many public occasions. They know that those engage in who have this kind of talk are considered to ignore social justice, and might be accused of being segregationist. On the other hand, in cyberspace, unjustifiable gender norm which is considered to be eliminated in the real space, has existed and, what is even worse, has been strengthened over and over again. Therefore, we need to find and realize gender issues behind ICT and communication using it, and also need to evaluate gender as social risk appropriately and ethically.

REFERENCES

Adam, A. (2008) The Gender Agenda in Computer Ethics, in Himma, K. E. and H. Tavani eds., *The Handbook of Information and Computer Ethics*, John Wiley & Sons, Hoboken, NJ, pp. 589-619.

Asai, R. (2007) 'NEKAMA: Living Another Gender on the Net', ETHICOMP2007, March, Tokyo.

Asai, R. and Kiyoshi M. (2009) 'Gender Issues in the Japanese Information Industry: Female Workers in the Masculine Workplace', UKAIS Workshop presentation material, May, Leicester, UK.

BB Serve (2005) Questionnaire for Online Game User of on-line game information site" 4Gamer.net", online at http://www.bb-serve.com/news/article/bbs_20050228.html, accessed 15.April.2009.

Cassell, J. and H. Jenkins M. eds. (2000) *From Barbie to Mortal Kombat*, The MIT Press, Cambridge, Massachusetts.

Cerulo, K., (1997) Identity Construction: New Issues, New directions, *Annual Review of Sociology*, 23:385-409.

Cherny, L., and Reba Weise E. eds. (1996) *wired_women*, Seal Press, Washington.

DiMaggio, P., Hargittai, E., Russell N. and Robinson, J.P., (2001), Social Implications of the Internet, *Annual Review of Sociology*, 27:307-336.

ECR Report (2006) *The Results of Research Online Game Market and Online Game Companies/Users 2004-2010* (In Japanese), online at <http://www.ec-r.co.jp/press/pdf/20060629.pdf> accessed 15.April.2009.

Friedman, E. J., (2005) "The Reality of Virtual Reality: The Internet and Gender Equality Advocacy in Latin America", *Latin American Politics and Society*, Vol.47, No. 3 (Autumn,2005), pp.1-34.

Green, E. and Adam A., (2001) *Virtual Gender*, Routledge, London.

Green, E., (2001) "Technology, Leisure and Everyday Practices", in Green, E. and A. Adam eds., *Virtual Gender*, Routledge, London, pp.173-188.

Ikeda, K. I., (2000) *Communication*, University of Tokyo Press.

Ishii, Junzo and Naotake Atsumi eds. (2002) *Marketing in the Internet Society*, Yuhikaku.

Kendall, L. (2000) "OH NO! I'M A NERD! Hegemonic Masculinity on an Online Forum", *Gender & Society*, Vol.14 No. 2, pp. 256-274.

Lohan, M. (2001) Men, Masculinities and 'Mundane' Technologies, in Green, E. and A. Adam eds., *Virtual Gender*, Routledge, London, pp.189-205.

Ministry of Economy, Trade and Industry (2008) "Annual report on the ICT use in Japan 2009", online at <http://www.meti.go.jp/press/20080818002/20080818002.html>, accessed 15 April 2009.

Ministry of Internal Affairs and Communications (2006) *WHITE PAPER Information and Communications in Japan*, National Printing Bureau.

Ministry of Internal Affairs and Communications (2008) *WHITE PAPER Information and Communications in Japan*, National Printing Bureau.

Nakamura, M. (2001) *Words and Gender*, Keisoshobo.

Nomura Research Institute (NRI) (1999) *Realities of the Internet users 1998*, online at <http://www.nri.co.jp/news/1999/990225/990225.pdf>, accessed 15 April 2009.

Online Game Forum (2006) *Online Game Market Statistics Research Report-2006*, online at <http://www.onlinegameforum.org/gaiyou2006.pdf> accessed 15 April 2009

Riesman, D. (1969) *The Lonely Crowd*, Yale University Press.

Toyama, A., Mariko K. and Kiyoshi M. (2008) *Information Management* (new edition), Yuhikaku.

Turkle, S., (1995) *Life on the Screen*, Simon & Schuster.

Turkle, S., (1999) Cyberspace and Identity, *Contemporary Sociology*, 28(6): 643-648.

USC Annenberg School Center for the Digital Future (2006) The 2007 Digital Future Report "Online World as Important to Internet Users as Real World?", online at <http://www.digitalcenter.org/pdf/2007-Digital-Future-Report-Press-Release-112906.pdf>, accessed 15.01.2007

Wajcman, J. (2004) *Techno Feminism*, Polity Press, Cambridge.

Wellman, B. and Hampton K., (1999) Living Networked On and Off, *Contemporary Sociology*, 28(6): 648-654.

Wolmark, J. (1999) *Cybersexualities*, Edinburgh University Press, Edinburgh.

Yamagishi, T. (1998) *Construction of Trust*, University of Tokyo Press.

Figures

Figure 1. User posted video "Wii vs. PLAYSTATION3" online at <http://www.youtube.com/watch?v=x7PhJp3ciRQ>, accessed 15 April 2009.

Figure 2. Official Wii-fit promotional video online at http://us.wii.com/viewer_tvcm_usa.jsp?vid=5, accessed 15 April 2009.

Ethics at the Crossroads of Bioinformatics and Nanotechnology

Athanasios Alexiou*

PhD candidate, Ionian University,
Department of Informatics

&

Panayiotis Vlamos**

Assistant Professor, Ionian University,
Department of Informatics

Abstract

In our days, bioinformatics is additionally concerned with whole organism data, acting as Humanistic Science and outbidding in a way the basic principles and methods which arise from the well-known Central Dogma of Genetics. On the other hand, the challenge of constructing nanodevices of biological interest by imitating the operations of cells and other biological systems seems more realistic through the successful efforts in the synthesis and manufacturing of nanoscale materials, using building blocks. While the plethora of bionanomedical applications like micro-electromechanical systems (MEMS) and 'lab-on-a-chip' devices is enriched and combined with the possibilities of bioinformatics, the variability in the ideological use of such concepts is associated with bioethical issues and several legal aspects. The convergence of bioethics and computer ethics, attempts to illustrate and approach problems, occurring by the fusion of human and machine. But how feasible is to achieve and control privacy, on the attempt of developing new techniques with real time access, to genetic information or human neurons? Several social and ethical aspects have to be addressed, for example the

* *Athanasios Alexiou* is a PhD student on Algorithms in Bioinformatics, in the Department of Informatics of Ionian University, with Diploma in Mathematics from the University of Athens. His main research interests are in the area of Algorithms and AI tools in informatics and Mathematical modeling in Education with the use of new technologies.

** *Dr. Panayiotis Vlamos* is an Assistant Professor in the Department of Informatics of Ionian University, with a Diploma in Mathematics from the University of Athens and a Ph.D. in Mathematics from the Department of Mathematics of National Technical University of Athens. His research interests are Mathematical Modeling applied for image restoration purposes, Mathematical Modeling applied to Biological, Physical and Chemical Engineering problems, Discrete Mathematics, Mathematical modeling in Education with the use of new technologies and Bioinformatics.

ones concerning clinical issues, privacy, confidentiality in medical diagnosis, directed individualized treatment or even more subjects of criminality and immorality, human dignity and justice or even industrial cost.

Keywords: Bioinformatics, Genetic, Nanoscience, Bioethics, Convergent Technologies, Micro-Electromechanical Systems

Introduction

Aldo Leopold, states that 'a thing is right when it tends to preserve the integrity, stability and beauty of the biotic community while it is wrong when it tends otherwise' (Leopold, 1987). The new realizations of science put the modern person in front of dilemmas, as far as the moral and the legal, the socially acceptable and the pioneering dangerous, the naturally valorized and the technologically innovative. These dilemmas are concerned diachronic and 'tragically' partially, if we consider Sofocleia tragedy, the 'Antigone' and the conflict between the legally imposed and morally deliberate.

In reality however, these dilemmas are conflict of values, and humans treat the values always, either by serving them consciously or deliberately alleging them. The converging of science and technology in several levels through their realizations is used to profit individuals, under the condition that humans serve always-high values and humanitarian ideal.

The world of moral debt is, as it should be, the world of science, serving fields sensitive to the human biological identity and uniqueness. The human effort of improving performance and correct natural imperfections of its biological existence in nano-level resembles with need to find parallel worlds (Kaku, 2004) in macro-scale and culturally promoted by establishing itself as the primarily dominant entity on the planet.

Speaking about the 'leading-edge' of knowledge sounds fine until we remember that it is also on the brink of ignorance (Holdsworth, 1995).

Bioinformatics, as it is known today is a phenomenon of the 1990s. There is a temptation to define bioinformatics as the use of electronic computing to analyze and interpret gene-sequences which have been yielded by gene-sequencing research in the laboratory.

Bioinformatics need software tools and special Biological Databases to compare DNA and protein sequences, search for coding regions in DNA sequences, and predict the multi level structures of molecules.

For example the capability of folding proteins constitutes one of the most important natural operations, where each kind of protein has a particular 3-D structure,

which is also determined by the sequence of amino-acids in its chain. During the examination of a protein's structure, various levels of self-organization can be distinguished. This could actually be impossible to achieve without the existence of Databases like the Protein Data Bank (PDB). The specific Database provides a variety of tools and resources and users can perform simple and advanced searches based on annotations relating to sequence, structure and function. These molecules are visualized, downloaded, and analyzed by users ranging from students to specialized scientists.

On the other hand nanotechnology is having a great impact on the fields of biology, biotechnology and medicine. This area of nanotechnology is generally referred to as nanomedicine, and sometimes widely called bionanotechnology (Vo-Dinh, 2007; Niemeyer, Mirkin, 2004).

The transition to the era of nanotechnology, if the claimants are right, will be a technological revolution even greater than the information technology and biotechnology revolutions (Drexler, 1992).

'In short, replicating assemblers will copy themselves by the ton, then make other products such as computers, rocket engines, chairs, and so forth. They will make disassemblers able to break down rock or supply new material. They will make solar collectors to supply energy. Though tiny, they will build big. Teams of nanomachines in nature build whales, and seeds replicate machinery and organise atoms into vast structures of cellulose, building redwood trees. There is nothing too startling about growing a rocket engine in a specially prepared vat. Indeed, foresters given suitable assembler 'seeds' could grow spaceships from soil, air and sunligh' (Drexler, 1992).

Undoubtedly several issues are also related to advanced topics and applications of nanotechnology (molecular analysis, design and manufacturing), such as human immortality and democracy, artificial intelligence on nanomachines, therapeutic limitations e.tc. As Feynman said, 'there's plenty of room at the bottom'.

Bioinformatics and Genetic Information

It all began when Robert Hooke in 1665, discovered that organisms are composed of individual compartments called cells. In the next years the study of cells became the study of what life is made of, contributing in the development of Biology. We can perceive the cell as a complex mechanical system with many moving parts, which are born, they eat, replicate and die, containing all the information and machinery required to collect and manufacture its component (Jones, Pevzner, 2004).

Additionally, many remarkable scientific experiments of Molecular Biology led to discoveries about the genetic material (genome, gene, genotype, phenotype, nucleic acid, proteins e.tc.), like the double helical structure of the DNA molecule (Watson, Crick, 1953) or the genetic makeup of the human species through the Human Genome Project (1990-2003) changing completely the traditional medical care in all levels. Darwin's theory (1859) that all living things have involved through a process of incremental change over millions of years began to seem more realistic, through the knowledge of the genetic information and the raise of Bioinformatics. The well-known 'Central Dogma of Genetics' (Crick, 1958) and the observations of the various similarities and variations between humans helped define more accurately the term evolution among different species, but mostly among all human individuals.

Hopefully, no scientific criteria have been established for race, ethnicity or for division of human ethnicity (Human Genome Project-ELSI 2003). On the contrary several mutations have been identified, responsible for many genetic diseases.

In early days bioinformatics is concerned with whole organism data, especially human physiological variables. More specifically, it includes the process and methods applied to the upgrading of the information content of biological measurements namely the utilization of sequence, expression, proteomic and physiological data to identify characteristic patterns of disease and elucidate mechanisms of gene regulation, signal transduction, flux control and overall cell physiology (Stephanopoulos, 2003).

Therefore, as Bioinformatics deals with biotechnology, computer technology and also life sciences, the ethics emerging from this scientific field has to be an amalgam of the two major strands of applied ethics: computer ethic and bioethics (Hongladarom, 2006).

On the other hand the parallel innovating structure of the so called 'convergent technologies', referring to the NBIC tools and including nanoscience and nanotechnology, biotechnology, biomedicine and genetic engineering, information technology and cognitive science, seems to remove any barrier in scientific and technological achievement (Roco, Bainbridge, 2002).

The ability to operate at the scale of telomeres makes it possible to extend or shorten the life of a cell (Leutwyler, 1998; McKibben, 2003). The nanodevices which can repair cells, promise great improvements in longevity and quality of life, involving radical modifications of the human genome and leading to the old but diachronic issue of human immortality (Drexler, 1986). Such changes pose a threat to the stability of the basic principles of Molecular Biology, as they are expressed through the cell's cycle.

Bioinformatics nowadays entails the creation and advancement of databases, algorithms, computational and statistical techniques, and theory to solve formal and practical problems arising from the management and analysis of biological data. Over the past few decades rapid developments in genomic and other molecular research technologies have been combined to produce a tremendous amount of information related to molecular biology. Common activities in Bioinformatics include mapping and analyzing DNA and protein sequences, aligning different DNA and protein sequences to compare them and creating and viewing 3-D models of protein structures. Bioinformatics is that branch of life science, which deals with the study of application of information technology in the field of molecular biology.

The primary goal of bioinformatics is to increase our understanding of biological processes. What sets it apart from other approaches, however, is its focus on developing and applying computationally intensive techniques (e.g., data mining and machine learning algorithms) to achieve this goal. Major research efforts in the field include sequence alignment, gene finding, genome assembly, protein structure alignment, protein structure prediction, prediction of gene expression and protein-protein interactions, and the modeling of evolution.

Genomics helps identify genes responsible for illness. 'All of our drug therapy today is directed at about 500 targets—molecular sites in cells where drugs effect their actions and there will be 5,000 to 10,000 targets identified through genomics, which means there is a tremendous opportunity for new-drug discovery' (Stevens, 1997). Biology itself provides a fully worked out example of a functioning nanotechnology, with its molecular machines and precise, molecule by molecule, chemical syntheses. What is a bacterium if not a self-replicating, nanoscale robot? Yet the engineering approach that radical technologists have proposed to make nanoscale robots is very different to the approach taken by life (Jones, 2004).

Substituting Nature via Nanotechnology

Giving a basic definition, Nanotechnology is the engineering of functional systems at the molecular scale. In a more efficient way, imagine a medical device that travels through the human body to seek out and destroy small clusters of cancerous cells before they can spread, or a box no larger than a sugar cube that contains the entire contents of the Library of US Congress, or materials much lighter than steel that possess ten times as much strength (U.S. National Science Foundation).

Theoretically, Nanotechnology is also defined as building things from the bottom up, with atomic precision, which are manufacturing simultaneously, something that Richard Feynman (Nobel Prize winner in physics) envisioned in 1953.

Nanoscience is the study of materials and associated physical, chemical, biophysical and biochemical phenomena on the scale of ~ 1 -100 nanometers. Also nanotechnology can be defined as being the term used to cover the design, construction and utilization of functional structures with at least one characteristic dimension measured in nanometres (Kelsall, Hamley, Geoghegan, 2005). Nanotechnology seems to be an umbrella term for a wide range of technologies, and, of particular significance, represents a convergence of quantum physics, molecular biology, computer science, chemistry, and engineering (Mehta, 2002), involving the study, control, manipulation and assembly of multifarious nanoscale components into materials, systems and devices to serve human interests and needs (Berne, 2004)

An enormous range of nanoproducts seems to make our life better, from aerospace technology, sports and agriculture to human brains and targeted drugs. Especially in the field of medical care and biomedicine, the convergence of nanotechnology and biotechnology with cognitive science began to produce new materials for improving human performance. Researchers looked toward biology as a guide to assemble nanostructures into functional devices, where only a small amount of subunits are required to produce a rich and diverse group of functional systems (Mardyani, Jiang, Lai, Zhang, Chan, 2004). It is believed that in the next two decades additive neurons will be manufactured that will help millions of people with damage in the brain. The ability to create therapeutic structures, channels, or diagnostic agents small enough to escape blood vessels and insert into specific types of cells, such as cancer cells, requires materials less than 20 nanometers in diameter (Kong et al., 2000).

Even in the case of quantum theory, a space time journey can be allowed only in microscopic level, via the development of nanotechnology and the possibility of copy and transport of genetic information (Hawking, 2002).

Symptoms of certain neurological diseases can be soothed with the help of cables, which are implanted deeply into the brain of patient. This method of treatment is named Deep Brain Stimulation (DBS) and is used mainly when the pharmaceutical solution no longer proves sufficient in patients with neurological diseases, specifically when the disease can be connected with concrete centers of brain. The treatment with DBS method is applied in various centers of brain. In all cases, one or more thin cables are placed deeply into the brain of patient, where some of them are connected with electrodes and the others are connected with a very small computer that sends electric signals in the electrodes. The microcomputer is usually placed under the skin, in the skull or in some other point in the upper part of patient's body.

Additionally, using the cochlear implant, destroyed acoustic cells in the internal ear are also replaced. The destruction of these cells leads to deafness; if however the brain is stimulated by the implant, the hearing impaired can acquire once again limited hearing. Microphones receive the sound from the environment and through a microcomputer the signals are transported wireless through the implant in internal ear, where electrodes irritate the acoustic nerve. This transmits the signals in the acoustic centre of brain (EuroHear Program, EC FP6).

Respectively, the sensory cells of eye might in the future be replaced, while experiments are in development for the creation of artificial eye. Microchips replace destroyed sensory cells of eye, changing the light in electric stimuli, via which the nervous cells, the retina and the optical nerve, are transmitted in the brain.

Several companies already circulate appliances, which read the electric excitations of brain, changing in this way human thought in data. Researchers have also proved that it is possible to remote-controlle individuals, without certain surgical intervention, through special earphones sending electric signals to the ear's centre of balance. The brain is tricked and believes that the body will fall if the person does not move immediately to the direction that the researchers dictate.

Producing sensors that could fit inside cells and monitored non-invasively would allow the continual evaluation of events in biological systems or humans. Sensing could also be used to measure small alterations in human function that are related to toxic elements in the environment or other types of abnormalities or degenerative diseases (Bainbridge, Roco, 2006).

Bioinspired materials may create difficulties for biologic systems and ecosystems, as their small size allows them to be easily internalized in organisms. These materials can mimic biologic molecules and disrupt their function, and there have also been problems with certain synthetic structures such as "buckyballs" and carbon nanotubes, which have been proved to have toxic effects on cells and animals (Oberdörster, 2005).

Many scientists have developed nanoscale materials for biological applications that failed because of their toxicity or because they were bioincompatible (West, Halas, 2003; Lam et al., 2004).

In most of the cases the relation of applied nanoscience with several environmental and social threats, where humans seems to 'play God' with natural processes, cause questions of social and environmental nature to arise and wake up fears of the past about who patents and controls nanotechnology and these new nanomaterials (Preston, 2005). How immoral and harmful for the human freedom can be the effort to force biology to do a better job than nature has done? In any case

every living organism in this planet has to manipulate nature in order to stay alive and also to obey the laws of nature.

Furthermore, when the clinical trials involve novel nanomaterials whose physiochemical properties are insufficiently studied, potential research subjects should be informed that unpredictable risks may arise during the trials (Resnik, Tinkle, 2007).

The use of certain kinds of nanomaterials, nanomedicines or nanodevices also raises fundamental questions about human enhancement and human nature (President's Council on Bioethics, 2003), about what are living and non-living and the definition of normal and human entity.

Nanotechnology raises the question of agency; a central proposal of the nanotechnological project is precisely to create self-replicating machines. (Horner, 2006)

Nevertheless the extreme possibility of nanomachines going out of control by using their power of reconstruction and self-replication might more likely to happen in terms of a terrorist attack, despite than a machinery revolution.

Emerging Issues in Bionanotechnology

Already, there are a few nanomedicine-related products on the market with numerous other potential applications under consideration and development (Vo-Dinh, 2007; Niemeyer and Mirkin, 2004; Kubik, 2005).

In vivo disease detection and monitoring using micro-electromechanical systems (MEMS) also appears to be making applications for creating "lab-on-a-chip" devices to detect cells, fluids or even molecules that predict or indicate disease even more probable (Craighead, 2006).

Lab-on-a-chip devices involve a combination of nanotechnology and microfluidics where multiple sample mixing, transport, integration, detection and data processing are all conducted on a single chip.

The use of MEMS chips and other devices for the purpose of diagnosing or monitoring healthy or diseased conditions is likely to raise grave questions about health information systems, privacy and confidentiality in our healthcare system (Allhoff, Lin, 2009). The manufacturing of devices able to provide real time processing of several blood levels, leads to a strong cost benefit for people with chronic diseases or organ transplant.

Nano-diagnosis could possible prevent illnesses or the impact of a disease and reduce the cost of drug discovery and development. Inexpensive and higher throughput DNA sequencers based on nanotechnology can reduce the time for both drug

discovery and diagnostics. It is clear that nanotechnology-related advances represent a great opportunity for the drug industry as a whole (Allhoff, Lin, 2009).

This would include targeted therapeutics that has greater benefit and fewer side effects, in part because of individualized approaches based on a patient's genetic susceptibility (Phillips et al., 2003).

On the contrary a few researchers have further warned that the volume of data pouring out of the nanomedicine diagnostic spigot may eventually overwhelm the ability of health information systems to evaluate it—making effective treatment impossible (Goldstein, 2005). Currently, however, most countries do not have a healthcare information system ready to handle the significant amounts of data that would be generated by nanomedicine diagnostic devices described above.

Although many of these disorders are better dealt with in the industrialized world, new types of infections, releases of chemicals into the environment, and the development of new radiation sources are unique problems for developed countries that could also be addressed by nanotechnology (Koifman, Koifman, 2003).

Degenerative diseases are an area in which nanotechnology may work wonders at improving human health. Interventions in this area may achieve the greatest overall benefit in developed and industrial societies because of the tremendous financial implications for these societies; which societies are responsible for the care of a large amount of the elderly population that suffers from degenerative diseases (Hammel, 2003).

In addition, one would hope that eventually nanomedications would lead to enhancement of function of human biologic systems, not only by involving regenerating function but also by improving it to prevent disease (Lehmann-Horn, Jurkat-Rott, 2003). Nanotechnology improvements in diagnostics can be subdivided according to the time to their impact, from short-term to long-term. Improvements in diagnostics over the short term would essentially involve improvements in current laboratory techniques that would allow measurement with greater sensitivity and specificities (Majumdar, 2002).

Mid-term advances in diagnostics will involve diagnostics that are integrated into biologic systems. This would include concepts such as sensors within humans' cells and bodies that would provide constant information on biologic function (Shim et al., 2003).

These sensors would allow the real time monitoring and management of humans in any environment through wireless networks and in the same way that mechanical systems such as automobiles or airplanes are maintained. Constant feedback from biologic data would permit the immediate correction of abnormalities in

an individual. This “early warning”-type approach could prevent disease before it starts and would be much more effective than the current technologies (Bainbridge, Roco, 2006).

Long-term concepts in diagnostics would build on a biologically integrated diagnostic approach and would directly couple sensor systems to treatment modalities (Anderson et al., 2000).

“Smart therapeutics” can be expected within the next decade for the specific delivery of drugs and genetic agents. This will provide a degree of specificity in the action of therapeutics that could prevent side-effects and improve efficacy (Patri et al., 2003). Long-term therapeutic developments would include nanosystems that totally replace, repair, and regenerate diseased tissue. This could involve the correction of developmental defects or the resolution of problems from disease and trauma (Bainbridge, Roco, 2006).

Ethical and Legal Aspects

Is it possible to define the degree of influence of human conscience, dignity, rights and fundamental freedom by merging human and machine? Is it possible to achieve and control confidentiality and privacy on data concerning i.e. human brain activity, without of course increasing tremendously the high quality treatment cost? Who can develop and participate in such scientific experiments, who will be the subject of the experiment and how can we make provision for individuals with special needs? Will a human implant be the way of releasing a ‘supposedly without logic’ criminal, like a new generation of drugs? It is important to note, that such social and ethical issues are not specific to nanotechnology alone; any modern technology is the product of a complex interplay between its designers and the larger society in which it develops (Pool, 2003).

The development of nanotechnology is moving very quickly, and without any clear public guidance or leadership as to the moral tenor of its purposes, directions and outcomes; where nanotechnology is leading and what impact it might have on humanity is anyone’s guess (Berne, 2004).

What appears to be missing at the present time is a clearly articulated prognosis of the potential global social benefits and harms that may develop from further scientific and technological advances in all of these areas (Sweeney, 2006)

The International Declaration on Human Genetic Data, in its 3rd article gives a definition of person’s identity, guideline the good practices in these scientific areas: ‘Each individual has a characteristic genetic make-up. Nevertheless, a person’s identity should not be reduced to genetic characteristics, since it involves complex educational, environmental and personal factors as well as emotional, social, spiritual and cultural bonds with others and implies a dimension of freedom’.

According to Humanism, human beings have the right and responsibility to give meaning and shape to their own lives, building a more humane society through an ethic based on human and other natural values in the spirit of reason and free inquiry through human capabilities.

The conflicts on bioethics issues also extend to another form of ethics, the environmental, where the argumentation included, concerns the violation of aspects that are theoretically exclusive competence of natural mechanisms. The overshooting of limits that nature determines, through the mechanisms of self-adjustment and genesis, is considered as infringement in general, of ecological balance. Therefore, through the tendency of physiocentrism, we can distinguish especially the biocentrism, where all the living beings have value and humanity is not the center of existence (Lanza, 2007). The moral person however guided from his evolving social behavior, can easily comprehend and be committed to laws and principles that a scientific field, such as Nanoscience sets as a precondition, in order to improve the structural elements of man's biological existence.

Which is mainly the problem therefore? Are there any laws and ethical aspects for the consequences of artificial implants in humans, or mainly the consequences from their non-application?

The EU proposed the 'Nanosciences, Nanotechnologies, Materials and new Production Technologies (NMP)' theme in the Seventh Framework Programme (FP7) spending 3.5 billion over its duration of 2007-2013. On 19 July 2007, the Commission announced a public consultation for developing a tool, the 'Code of Conduct', which will make it very simple to address the legitimate concerns that can arise regarding nanotechnologies (European Science and Research Commissioner Janez Potonik 2007). Through a set of seven general principles, EU aims to control future research development on every Member States, to ensure that nanotechnologies are developed in a safe manner: Nanotechnology should be comprehensible to the public, including safe research activities, activities which benefit society and environment, guided by the principles of free participation to all decision-making processes; they should respect the right of access to information, having the best scientific standards and encouraging creativity, flexibility and innovation with accountability to all the possible social, environmental and human health impacts of their work.

Another aspect of the applied research of nanoscience in medicine is the economical cost. The cost of care and also the expectations of citizens are increasing rapidly. Even if research is directed towards solutions economically accessible, the more effective nanotechnological solutions can possibly lead to the unequal distribution of medical care and prompting assurances companies to stop their attendance in the social health system (Solodoukhina, 2007).

Is it possible therefore to achieve balance between the objective of scientific progress and the guarantee of obvious social goods particularly in sensitive groups of population? In which way will it become beneficial for all humans and

more feasible to merge the increasing cost of new methods in health diagnosis and treatment, with the accessible and high level medical care?

The Universal Declarations on Bioethics and Human Rights, on the 15th article cite that scientific research and its applications should be shared with society as a whole and within the international community, in particular with developing countries giving access to quality health care.

An approach towards this problem would be the withholding of scientific research in a minimum degree, in frames where the retributive profit, mainly to the citizens, is in a certain way and also in economical terms, proved bigger than the cost (Stefannson, 2002). Of course this subject seems multidimensional, with several politics and social-economical aspects concerning many various institutions and recipients involved in this, and critical problems like the open dialogue and the objective information of public.

However it appears that freedom in science and also the right to dignity and the autonomous and independent mappings out to the individualized future are elements that balance the scale of human and lead our culture. Otherwise, humans would have as base of existence and development, the eugenics, modifying their genetic identity to that of a perfect machine and betray all the history of philosophical and intellectual fights.

Drexler's 'gray goo' scenario in which self-replicating machines could get out of control and wreak destruction on the biosphere, gives the possibility of molecular engineering and manufacturing being used as a means of political power: '... the threat of advanced technology in the hands of governments make one thing perfectly clear: we cannot afford to have an oppressive state take the lead in the coming breakthroughs'.

In fact, some do argue that there is nothing ethically novel about nanotechnology (Litton, 2007; Lewenstein, 2005; Grunwald, 2005). These observers dismiss nanotechnology as a factor that will generate no truly novel ethical and social issues (Allhoff, Lin, 2009).

Instead, they feel that nanotechnologies simply raise the same standard issues of research ethics, privacy and confidentiality, put at stake by all other kinds of medical research and development. While this may be true to a large extent, nanoethics may be viewed as a convergence of many areas of ethics—it adds a new dimension to current ethical debates (Allhoff, Lin, 2006). Those ethical codes and frameworks differ slightly from profession to profession. This means that the ethics of nanomedicine may have a slightly different set of core moral values or considerations than traditional medical applications due to the influence of other ethical frameworks and perspectives on the research and development of these interventions (Allhoff, Lin, 2009).

Nanotechnology will make most things far cheaper, lighter, smaller and more efficient – redundancy and reliability will become affordable and practical. (Horner, 2006).

Nanomedicine will allow us to understand down to the atomic and single-cell level how our bodies are performing at any given moment. For some, this information could be helpful, empowering or enlightening and may enhance human health. For others, it is likely that such information could result in fear, anxiety and other mental health issues. Therefore, a delicate balance may need to be established here between the information processed/disseminated versus the benefit to society and individual health (Allhoff, Lin, 2009). It is imperative that the research risks be clearly communicated to the subjects. In fact, to gain and maintain public support for nanomedicine generally, an honest and open discussion with the public regarding the ethical and social issues surrounding nanomedicine should be promptly undertaken (Mills, Fleddermann, 2005).

It is quite obvious that the development of novel therapies based on the convergent of Bioinformatics and Nanotechnology will arise several ethics principles about human rights which have to be followed: moral, political and religious issues but also individual privacy, human dignity, justice, and fair access to the knowledge of the diseases but further more to any possible beneficial therapy. Therefore, it is ethically essential that researchers inform potential research subjects in clinical trials of all details pertaining to the study (i.e., purpose, experiments, risks/benefits, alternatives, confidentiality protection, e.tc.) (Donaldson, 2006).

However, most of these therapies and benefits may be out of reach for many people of lower socioeconomic status or those who reside in developing countries, increasing also national and international inequalities, where the knowledge of the diseases and the treatments is probably a 'big secret'.

Ethical issues in Nanoscience reflect the character of the new science itself. No longer can such issues be dealt with in an isolated way by "ethicists" or by scientists alone. The issues are now too complex, and they require the full range of skills of those in the sciences and the humanities (Khushf, 2004).

Conclusion

Any kind of ethical considerations about novel adopted and developed technologies should be discussed and explained to the social target groups at early stage.

Therefore, it is ethically desirable that extensive short- and long-term studies be undertaken to determine whether nanomedicines will be more effective and safe for humans when compared to conventional drugs, due to the complexity of biologi-

cal systems, and therefore a lack of appropriate caution when using nanomaterials (West, Halas, 2003; Lam et al. 2004; Bainbridge, Roco, 2006; Allhoff, Lin, 2009).

It seems difficult at this time to predict whether the convergent bio-technologies will deliver a variety of improvements or even, a technological and healthcare revolution. The study of cost-benefit and cost-effectiveness is necessary in order to consider if bionanotechnology is unpredictable science or a predictable future technology (Horner 2006).

'If we could predict tomorrow's technology to any degree of accuracy, we would be able to predict how it would work. But if we knew today how it would work, we would be able to develop it now. It would be today's technology...So if technological developments are, in their nature, unpredictable, then there is a very clear sense in which the future will be radically unlike what anyone can foresee now with any degree of certainty or justification' (O' Hear, 1999).

Several scientific concepts of Nanoscience, should be introduced at all educational levels, from students to scientists and to non-technical audiences, that may decide the use of technology and its funding (Roco, 2003).

REFERENCES

Allhoff, F., P. Lin. (2006) What's so special about nanotechnology and nanoethics? *International Journal of Applied Philosophy* 20.2: 179–190.

Allhoff, F., P. Lin. (2009) *Emerging Issues in Nanomedicine and Ethics. Nanotechnology & Society: Current and Emerging Ethical Issues*, Springer Science+Business Media B.V. 2009

Anderson, S.A., Rader R.K., Westlin W.F., Null C., Jackson D., Lanza G.M., Wickline S.A. and Kotyk J.J. (2000) Magnetic resonance contrast enhancement of neovasculature with alpha(v)beta(3)- targeted nanoparticles. *Magnetic Resonance in Medicine* 44: 433–439.

Berne, R.W. (2004) Towards the conscientious development of ethical nanotechnology. *Science and Engineering Ethics*, 10: 627-638.

Bainbridge, W.S., Roco, M.C., (2006) *Managing Nano-Bio-Info-Cogno Innovations. Converging Technologies in Society*, Springer, 119-132.

Craighead, H. (2006) Future lab-on-a-chip technologies for interrogating individual molecules. *Nature* 442: 387–393.

Crick, F. (1958) Central Dogma of molecular biology. *Nature* 227, 561-563

Donaldson, K. (2006) Resolving the nanoparticles paradox. *Nanomedicine* 1: 229–234.

Drexler, E. (1986) *Engines of Creation*. Bantam. New York

Drexler, E. (1992) *Engines of Creation: the Coming Era of Nanotechnology*. Oxford University Press, Oxford

Goldstein, A.H. (2005) *Nanomedicine's Brave New World*, 28 November.

Grunwald, A. (2005) Nanotechnology—a new field of ethical inquiry? *Science and Engineering Ethics* 11:187–201.

Hammel, J. (2003) Technology and the environment: Supportive resource or barrier for people with developmental disabilities? *Nursing Clinics of North America* 38:331–349.

Hawking, S.W., Thorne, K.S., Novikov, I., Ferris, T., Lightman, A. (2002) *The future of Spacetime*. W.W. Norton, New York

Hongladarom, S. (2006) Ethics of Bioinformatics: A convergence between Bioethics and Computer Ethics. *Asian Biotechnology and Development Review*, Vol. 9, No.1, 37–44.

Holdsworth, D. (1995) Ethical decision-making in science and technology. In B. Almond, editors, *Introducing Applied Ethics*, pp. 130–147. Blackwell, Oxford

Horner, D.S. (2006) Anticipating ethical challenges: Is there a coming era of nanotechnology? *Ethics and Information Technology*, Springer, 7, 127–138

Jones, R. (2004) *Soft Machines: Nanotechnology and Life*. Routledge, London

Jones, N.C., Pevzner, P.A. (2004) *An Introduction to Bioinformatics Algorithms*. MIT Press

Kaku, M. (2004) *Parallel Worlds: The Science of Alternative Universes and Our Future in the Cosmos*. London. Allen Lane

Kelsall, R. W., Hamley, I. W. & Geoghegan, M. (eds) (2005) *Nanoscale science and technology*. John Wiley & Sons, West Sussex, UK.

Khushf, G. (2004) Systems theory and the ethics of human enhancement: a framework for NBIC convergence, in: Roco, M.C. & Montemagno, C.D. (eds), *The coevolution of human potential and converging technologies* (*Annals of The New York Academy of Sciences*, Volume 1013, May 2004), The New York Academy of Sciences, New York, 124–149.

Koifman, S., Koifman, R. J. (2003) Environment and cancer in Brazil: An overview from a public health perspective. *Mutation Research* 544: 305–311.

Kong, G., Braun, R.D., Dewhirst, M.W. (2000) Hyperthermia enables tumor-specific nanoparticle delivery: effect of particle size. *Cancer Research* 60: 4440–4445.

Kubik, T. et al. (2005) Nanotechnology on duty in medical applications. *Current Pharmaceutical Biotechnology* 6: 17–33.

- Lanza, R.** (2007) *A New Theory of the Universe*. Spring, The American Scholar
- Lam, C.W., James, J.T., McCluskey, R., Hunter, R.L.** (2004) Pulmonary toxicity of single-wall carbon nanotubes in mice 7 and 90 days after intratracheal instillation. *Toxicological Sciences* 77: 126–134.
- Lehmann-Horn, F., Jurkat-Rott, K.** (2003) Nanotechnology for neuronal ion channels. *Journal of Neurology, Neurosurgery, and Psychiatry* 74: 1466–1475.
- Leopold, A.** (1987) *A Sand County Almanac: And Sketches here and there*. Oxford University Press, New York
- Leutwyler, K.** (1998) Turning Back the Strands of Time, *Scientific American*, Febr.
- Lewenstein, B.V.** (2005) What counts as a ‘social and ethical issue’ in nanotechnology? *Hyle International Journal for Philosophy of Chemistry* 5: 5–18.
- Litton, P.** (2007) Nanoethics: What’s new? *Hastings Center Report* 37: 22–25.
- Majumdar, A.** (2002) Bioassays based on molecular nanomechanics. *Disease Markers* 18: 167–174.
- Mardyani, S., Jiang, W., Lai, J., Zhang, J., Chan, C. W.** (2004) *Biological Nanostructures and Applications of Nanostructures in Biology*. Springer US
- McKibben, B.** (2003) *Enough*. Henry Holt, New York
- Mehta, M.D.** (2002) Nanoscience and nanotechnology: assessing the nature of innovation in these fields. *Bulletin of Science, Technology & Society*, 22: 269–273.
- Mills, K., Fleddermann, C.** (2005) Getting the best from nanotechnology: Approaching social and ethical issues openly and proactively. *IEEE Technology and Society Magazine* 24.4:18–26.
- Niemeyer, C.M., Mirkin, C.A.** (2004) *Nanobiotechnology: Concepts, Applications and Perspectives*. Weinheim, Germany: Wiley-VCH.
- Oberdörster, G. et al.** (2005) Nanotoxicity: An emerging discipline evolving from studies of ultrafine particles. *Environmental Health Perspective* 113: 823–839.
- O’Hear, A.** (1999) *After Progress: finding the old way forward*. Bloomsbury, London
- Patri, A., Majoros, I., Baker, J.** (2003) Dendritic polymer macromolecular carriers for drug delivery; using nanotechnology for drug development and delivery. *Current Opinion in Chemical Biology* 6(4): 466–471.
- Phillips, K. A., Veenstra D., Van Bebber, S., Sakowski, J.** (2003) An introduction to cost-effectiveness and cost-benefit analysis of pharmacogenomics. *Pharmacogenomics* 4: 231–239.

Pool, R. (2003) How society shapes technology, in: Teich, A. H. (ed.), *Technology and the future* (9th edition), Wadsworth/Thomson Learning, Belmont, CA, 13-22.

Preston, C.J. (2005) The Promise and Threat of Nanotechnology. Can Environmental Ethics Guide US? *International Journal for Philosophy of Chemistry*, Vol. 11, No.1, 19-44.

Resnik, D.B., Tinkle, S.S. (2007) Ethical issues in clinical trials involving nanomedicine. *Contemporary Clinical Trials* 28.4: 433-441.

Roco, M.C., Bainbridge, W.S. (2002) Converging technologies for improving human performance. *Journal of Nanoparticle Research*, 4, 281-295.

Roco, M.C. (2003) Converging science and technology at the nanoscale: opportunities for education and training. *Nature Biotechnology*, 21: 1247-1249.

Shim, J., Bersano-Begey, T.F., Zhu, X., Tkaczyk, A.H., Linderman, J.J., Takayama, S. (2003) Micro- and nanotechnologies for studying cellular function. *Current Topics in Medicinal Chemistry* 3: 687-703.

Solodoukhina, D. (2007) Bioethics and legal aspects of potential health and environmental risks of nanotechnology. *Nanotechnology-Toxicological Issues and Environmental Safety*, Springer, 167-184

Steffansson, H. (2002) *Life Sciences in Transition*. EMBL Essays on Science and Society. JMB, Academic Press

Stephanopoulos, G. (2003) *Bioinformatics_Methods|principles|applications*. Lectures MIT

Stevens, T. (1997) 'The gene machine'. *Industry Week*, August 1997, pp. 168-175.

Sweeney, E.A. (2006) Social and Ethical Dimensions of Nanoscale Research. *Science and Engineering Ethics*, Vol. 12, Issue 3, 435-464

Vo-Dinh, T. (2007) *Nanotechnology in Biology and Medicine: Methods, Devices, and Applications*. Boca Raton, FL: CRC Press.

Watson, J.D., Crick, F. (1953) Molecular Structure of Nucleic Acids. *Nature* 171, 737-738

West, J. L., Halas, N. J. (2003) Engineered nanomaterials for biophotonics applications: improving sensing, imaging, and therapeutics. *Annual Review of Biomedical Engineering* 5: 285-292.

Modeling Aspects of Action Theory

Sabah Al-Fedaghi*

Computer Engineering Department, Kuwait University

Abstract

Issues related to the nature of action have been of central concern in philosophy. In computing, this topic has raised a great deal of interest, especially in such areas as human-computer interaction, computer games, and artificial intelligence. This paper proposes building a conceptual model that operates as a framework to specify action-oriented situations. The central idea is that actions are viewed as “things that flow,” i.e., received, processed, created, released, and communicated. Things that flow include information, (physical) actions, beliefs, and values.

Keywords: Action theory, Conceptual model, Flow, Control

Introduction

Issues related to the nature of action have been of central concern in philosophy. Questions have been debated over differences between what happens, called events, and what is done, called actions. Furthermore, philosophical discussions have emerged over related issues, such as purposeful behavior that distinguishes human activity from a low-level, animalistic type. Special attention has been focused on action with complex psychological composition that is directed at some object and involves goals and desires. Also of special significance are autonomous actions, in which humans have direct control over their own activity. Further, a great deal of attention has been paid to the explanation of these purposeful actions, which entail causes, desires, beliefs, e.tc.

In computing, the notion of action has raised a great deal of interest. It is described as “one of the most important concepts in computer science” (Tian and Shi, 1997). For example, some researchers in the field of human-computer inter-

* Sabah Al-Fedaghi holds an MS and Ph. D. degree in computer science from Northwestern University, Evanston, Illinois; and a BS in computer science from Arizona State University, Tempe, Arizona. He has published articles in journals and had contributed to conferences on topics in database systems, natural language processing, information systems, information privacy, information security, and information ethics. Dr. Al-Fedaghi is an Associate Professor in the Computer Engineering Department, Kuwait University, Kuwait. He previously worked as a programmer at the Kuwait Oil Company and headed the Electrical and Computer Engineering Department (1991-1994) and the Computer Engineering Department (2000-2007).

action concern themselves with understanding the interaction between humans and computer technology. Of special interest are theories in psychology that deal with the psychological mechanisms of the person (computer user).

The notion of action has also been applied in the area of computer games. In these games, the player actions aim at solving problems in order to progress in the game. Of special interest is the action theory idea of “agency,” which deals with how a person (the agent) decides to take intentional actions and how the agent executes them. Agency is described as “the power to take action” (Laurel, 1997).

Action has attracted researchers in artificial intelligence, where one of its aims is to construct intelligent computer-based agents. Knowledge representation plays an important role, especially in the area of developing representative actions of an autonomously acting entity. Agents are software entities that formulate decisions to take action according to the perceived situation without interference by a human user.

Incorporating the concept of action into computer science stipulates building up a “conceptual model” that operates as a framework to specify action-oriented situations, different participants, and their relationships. In this paper, we propose the incorporation of a flow-based model that is uniformly applied to actions in machines, humans, and organizations. The central idea is that actions are viewed as “things that flow,” i.e., received, processed, created, released, and communicated. Things that flow include information, (physical) actions, beliefs, and values.

Theories of action [(Dick and Dalmau (1999); Argyris and Schön (1974))] deal with the relationship between thoughts and actions. They are utilized in developing a general model for behavior. The proposed flow-based model can be used in combination with these theories to investigate the nature of thought/action relationships. Theories of action are suitable for the flow-based approach because actions engage other phenomena, such as beliefs, feelings, and values; all can be envisioned as flowthings. Flowthings (pieces of information, pieces of beliefs, pieces of actions, and pieces of values) flow among different states: received, processed, created, released, and communicated. Additionally, the flow in each of these spheres, i.e., information, beliefs, actions, and values, may trigger a flow in another sphere (e.g., information triggers physical actions).

Models of Action

Action is typically defined as subjective behavior or activity distinguishing from mere thought or from mechanical behaviors. “Causal models” of action view action as intentions that cause specific actions. Hacker’s action theory (Hacker, 1985) analyzes human activity as a goal-directed behavior. Its main components are (Giacoppo, 2001): “acts” regulated by intentions (goals), “actions,” the small-

est units of cognitive and sensory-motor processes that are oriented towards conscious goals, and “operations” as components of actions that have no independent goals. According to Giacompo (2001),

Action Theory deconstructs the process of translating an intention into an action... The information processes of the users “regulate” the types of actions selected... Incoming stimuli are consciously organized by the user and trigger certain actions when the criteria for those actions are met [*italics added*].

The notion of triggering is an important element in the flow-based model and will be described later.

To illustrate the type of models that involve action, consider Dick and Dalmau’s (1999) model that draws a picture of inter-person interaction, as shown in Figure 1:

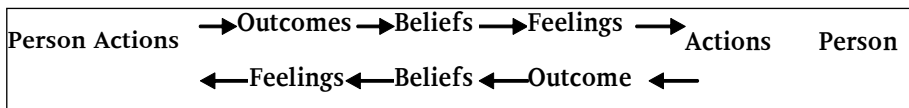


Figure 1. Inter-personal interaction (modified from (Dick and Dalmau, 1999).

Three features disturb the conceptual description in this figure as follows:

- It mixes ontologically different things. Actions, outcomes, and feelings are scattered (in the figure) without explicit acknowledgement of “typing” (e.g., actions on both ends are of the same ontological type).
- It mixes boundaries between actors and between actors and the outside world; i.e., are “feelings” inside persons and “outcomes” outside them?
- It lacks the specification of the types of relationships embedded in the arrows. “Outcomes ® beliefs” may indicate the creation of beliefs, or, alternatively, the reception of beliefs.

An alternative flow-based approach for the inter-personal interaction presented in Figure 1 is shown in Figure 2. It provides the general picture in the model that will be introduced in the next section.

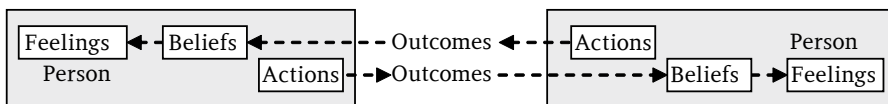


Figure 2. Re-drawing of figure 1.

In Figure 2, different “spheres” (persons and outside) are clearly distinguished. As will be described later, actions, beliefs (information or knowledge), and feel-

ings are “things that flow” in their own spheres and that have five states: received, processed, created, released, and communicated. The dotted arrows in Figure 2 indicate a “triggering” mechanism.

To further illustrate the type of models that involve action, consider the distinction between prior intention and intention in action (Searle, 1983). In this paper, we are not interested in the descriptive representation of actions; but, rather, in psychological analysis, such as the difference between the two types of intention mentioned above, e.g., the distinction between “conscious mental attitude leading” and “a real planning process.” Figure 3 shows a representation that describes action involving intention.

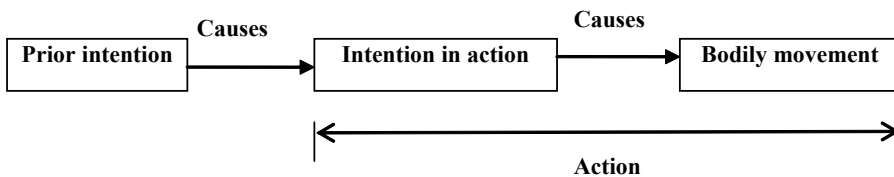


Figure 3. Prior intention, intention, and action (Searle, 1983, p. 94).

Conceptually, we can observe that this method of representation mixes different types of associations. While the “causes” arrow on the left indicates transformation from one type of intention to another, the right “causes” arrow signifies a completely different transformation from “intention” to “movement.” In analogy, consider the statement Ice causes water, which causes an electric circuit break. At least it seems that the two “causes” are completely different kinds. The conceptual picture ought to reflect that. The flowthing model categorizes intention in one sphere where intention is created and processed to generate a new intention. Intention may trigger (indicating a different type of flow) bodily movements that may be created, released, and transferred somewhere. Action, in this case, is a processed intention that that triggers the creation of movement.

Flowthing Model

This section summarizes previously published material that is presented herein, in order to make this paper self-contained (Al-Fedaghi, 2008). The new contribution in this paper is to consider action and related concepts as “things that flow” (called, for lack of a better term, flowthings) in order to apply the flowthing model (FM). We take the concept of flow as a foundation for conceptual modeling in contrast to modeling based on entities and relationships.

Notice that we are not interested in developing a metaphysical description of the discreteness of things in reality, such as actions. Simply, we assume that a “single

act” is what triggers a single piece of information or a single goal regardless of the complexity of these pieces.

In FM, the flow of flowthings indicates movement inside and between spheres. The sphere is the environment of the flow and includes five stages that may be sub-spheres with their own five-stage schema. The stages may be named differently; for example, in an information sphere, a stage may be called communication, while in action flow, the same stage is called transferring.

To illustrate the notion of flowthing, we will assume that the “thing that flows” is information. We use here the term information to refer to information and mis-information, as in the common usage that indicates a reporting statement, which can be either true or false. An information sphere denotes the information environment. The lifecycle of information is a sequence of states as it moves among stages of its lifecycle, as follows: (a) received; (b) processed (in a way that changes its form, but not content); (c) released; (d) communicated to another sphere; and (e) created (i.e., a new piece of information).

These five states of information are the main stages of the stream of flow, as illustrated in Figure 4. Each stage may include sub-stages, such as storage and usage.

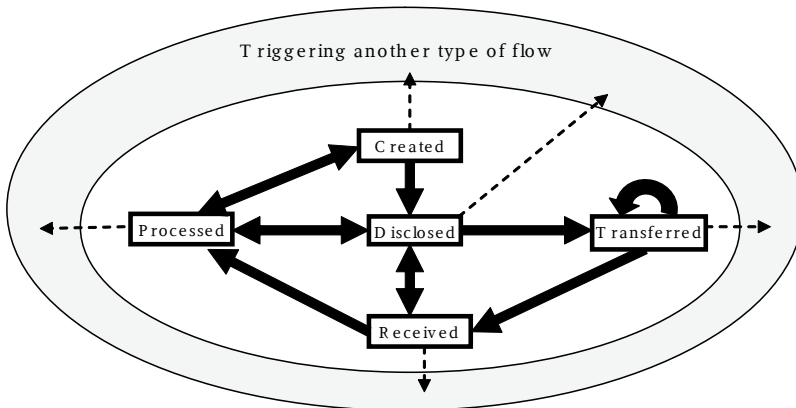


Figure 4. Information states in FM with the possibility of triggering another type of flow.

The “processing stage,” in FM, refers to any type of transformation of information without creating new information. The “disclosing stage” refers to releasing information to be transmitted outside the system. It is possible that the channel is down or busy. In case that it is not possible to make the channel functional again, then it is possible to return the tagged-out information to the processing stage or the receiving stage. This is analogous to passengers (flowthings) in the airport who have finished processing (e.g., passports) and are waiting to board the airplane.

General Model

The FM can be used to describe action and its related concepts. To illustrate this, we incorporate, in addition to concepts that have already been mentioned previously, “governing values.” Governing values “are goals we seek to satisfy, beliefs we seek to defend, values we seek to express... For example...to minimize expressing negative feelings; to be rational” (Dick and Dalmau, 1999) [*italics added*]. We propose that values are flowthings, just as are beliefs, feelings, and actions. They can be received (imported from others), created (self-imposed/created rules), processed (e.g., analyzed), released, and communicated. They are characterized by being voluntarily adopted. The values system has adopted/non-adopted values, such as true and false information.

Figure 5 shows the general concept of FM that includes actions, beliefs, feelings, and values. The flows of different flowthings are separated while permitting one flow to trigger another flow. The dotted arrows indicate the possibility that one flow triggers another flow.

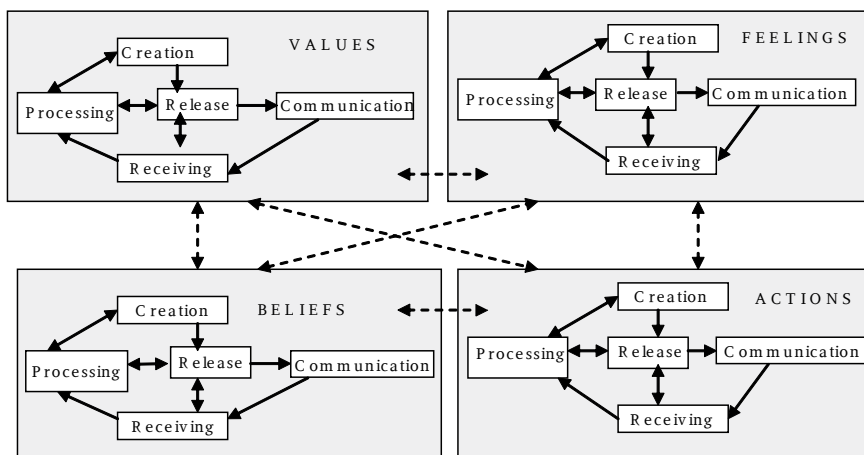


Figure 5. General conceptual model of actions, beliefs, feeling, and values.

We propose to generalize such a method of description to notions related to action. Action is typically defined as behavior with a goal. According to Li (1999), “an action may involve five aspects: perception, cognition, emotion, volition, and physical or motor performance.” Motivation refers to the “motives” that are related to the goal of determined behavior. Personal values are “standards or criteria held by people that effect the evaluative acts in which they are involved” (Rokeach, 1973). According to Nuttin (1984), “need” may be conceptualized as “a category of required relationships of the individual with his world.” Beliefs are

stated as propositions. Desires can be conceived as propositional desires (e.g., I will do something). Goal is defined in terms like end-states, outcomes, and desirable consequences (Heckhausen and Kuhl, 1985). These notions can be visualized as flowthings.

Consider again Searle's distinction between prior intention and intention in action, which was first modeled in Figure 3. The corresponding FM representation is shown in Figure 6.

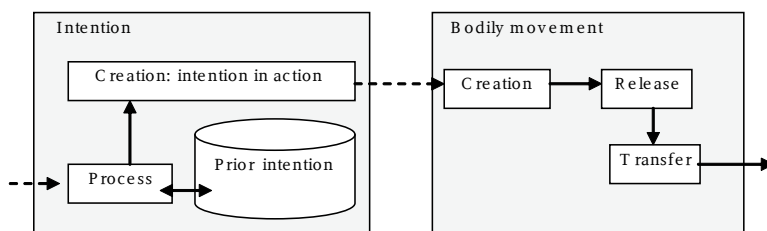


Figure 6. Representation of figure 3 in FM.

In this case we have two spheres: intentions and movements. We assume that prior intention is stored in the processing stage. The dotted arrow on the left indicates that some event triggered the retrieval of this prior intention to be further processed in order to create intention in action. This intention in action triggers the creation of bodily movement that is released and transferred to the outside. The resultant description provides a “purified” conceptualization of the semantics of the associations between different types of intention on one hand and bodily movements on the other. Additionally, it is more complete (e.g., the dotted arrow to the left) than the representation in Figure 3. Furthermore, this conceptualization separates the body sphere from the outside with introducing additional complexity.

Action, Intention, and Control

A variant of action theory, which is applied to behaviors that are under volitional control, is called reasoned behavior (Ajzen, 1985). In this conceptualization, intentions compete with each other; hence, the notion of control is introduced in order for the individual to execute the intended action (Li, 1999). Planned behavior is concerned with factors that influence the formation of intention and emphasizes the action plan. Later, we will use the concept of control and planned behavior in constructing a model of action. First, we examine the concept of intention that initiates the whole process that leads to bodily action.

Intention

As summarized by Li (1999), researchers have distinguished between “goal intention and implementation intention (behavioral intention).” Figure 7 shows a description of intentions that include plans, initiation, execution, and goals. Somehow, the place of control is missing in this picture. Nevertheless, it indicates that action involves planning and execution.

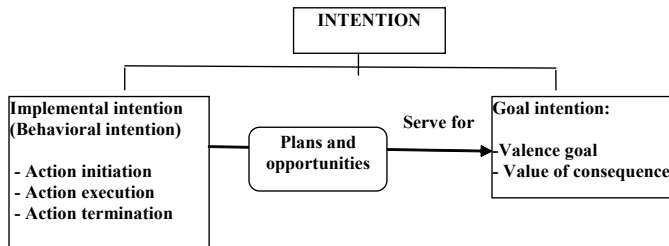


Figure 7. Goal intention and implementation intention (From (Li, 1999)).

Norman (1988) introduced seven stages of action related to the psychology of a person performing a task. As shown in Figure 8, the stages of execution are as follows:

- 1: Specification of the goal
- 2: Formulation of intention to do the action
- 3: Specification of a set of internal commands, an action sequence that can be performed to satisfy the intention
- 4: Execution of the action sequence

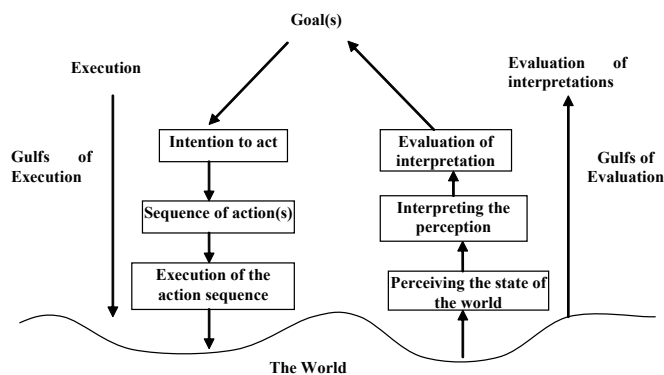


Figure 8. Norman's Action Cycle (From Human Factors Laboratory (2004) and Griffiths (2009)).

While this type of conceptualization seems to be successful in related studies, FM may provide a tool to build a scheme that complements this description and offers more details of the processes involved. However, before utilizing FM to represent the stages of action, it is necessary to understand the notion of control. Control was mentioned previously when reviewing Ajzen's (1985) volitional control where it is observed that, intentions compete with each other. Hence control is introduced in order for the individual to execute the intended action (Li, 1999). However, it is not clear how control fits into the sequence of stages as described, for example, by Norman's seven stages of action.

Control Process

Analyzing the previous descriptions that involve action shows that they include the following:

- 1: A pre-initiation presence determination of what to do next
- 2: An initiation of a non-static development of events
- 3: A flow that specifies movements and advances
- 4: An execution of something that results in an outcome

These characteristics point to a meta-system that initiates the process and controls its enactment. Control is usually visualized as a process with input and output. Also, control is described as "the process of monitoring activities to ensure that they are being accomplished as planned and correcting any significant deviation" (Robbins and Coulter, 2001).

In UML, control is visualized in the context of the general notion of behavior as follows.

Behavior models, in general, determine when other behaviors should start and what their inputs are... In particular, the UML 2 activity models follow traditional control and data flow approaches by initiating sub-behaviors according to when others finish and when inputs are available. It is typical for users of control and data flow to visualize runtime effect by following lines in a diagram from earlier to later end points and to imagine control and data moving along the lines. Consequently a token flow semantics inspired by Petri nets is most intuitive for these users, where "token" is just a general term for control and data values (Bock, 2003) (*italics added*).

Notice how the "flow" is qualified by "control and data" and then is connected to "token flow" in Petri nets. "Control flow," or "flow of control," is typically described as the order in which statements (of an imperative program), processes, operations, e.tc. are executed. But, does "control" flow?

Little has been written about the concept of flow, as discussed in this paper. According to Casni (2005):

The word ‘flow’ sprang up as the word fluxus in Latin, long before many of us can remember. Its root definition has remained intact, with the primary meaning “to move in a (steady) stream.” The cognitive image of a liquid is therefore fused into every metaphor using flow.

Casni (2005) distinguishes between things that flow (we call these flowthings) and their river beds (we call them flowstreams):

The sidewalks and isles are the customer’s equivalent to a river of asphalt – or dare I imply riverbed of asphalt. Because of their decidedly stationary and laterally limiting quality, these forms of indoor and outdoor flooring could not be the liquid centre of the metaphor flow, but are rather a channel for which something else of liquid quality will travel [underlining added].

Then, Casni (2005) asks, “What is the riverbed of a cash flow? Similarly, we ask, what is the flowstream of control? We suggest that, conceptually, control is the flowstream describing “control flow.” For example, the instructions (statements) of a computer program flow in and out of the stream-of-control unit. Assuming a sequential order of instruction/statements, Figure 9 shows the flow in the control sphere where instructions flow along the stages of the FM model of the control, one after the other. The control sphere refers to the flow of instructions in the FM representation of the control unit.



Figure 9. Statements' flow in the control sphere.

Control here acts as a machine that is fed with instructions/statements that queue to enter the control stream and exit. For example, if the program statements are given by the sequence s_1, s_2, \dots, s_n , then the flow can be described as:

- Transfer and receive s_1 in the control stream
- Process s_1 , where s_1 is activated
- Release and depart s_1
- Transfer and receive s_2
- Process s_2
- Release and depart s_2
- E.tc.

The flow, in this example, is the “pure” flow of statements in the control sphere.

Modeling Action Process

To achieve modeling of the phenomena that begins with intention and ends with a physical action, we use FM with mental and physical spheres as shown in Figure 10.

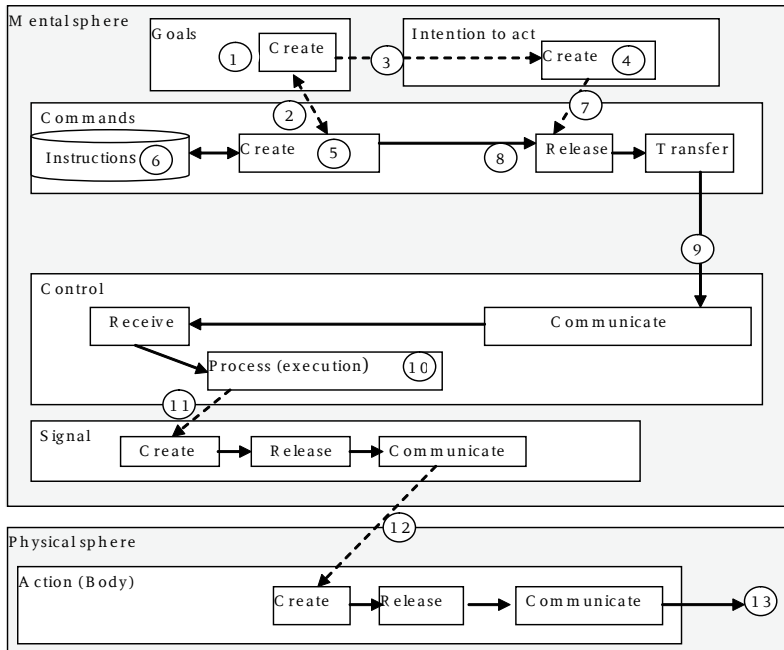


Figure 10. Process involved in creating an action.

In the mental sphere we have, first, the goal sphere, which, following Norman’s seven stages model, involves creating the goal (circle 1 in the figure above). In general, the goal may be received when outsiders influence the involved agent and insert in him or her the desire to achieve the goal. Figure 10 does not show this situation; but, certainly, it is possible to add it to the figure. The goal sphere interacts with other spheres in two ways:

1: Creating a goal is concurrent with triggering the creation of required steps of commands necessary to achieve that goal (circle 2). That is reflected in Figure 10 by the creation of these commands and storing them (circles 5 and 6 in the figure). This amounts to conceiving an ordered set of instructions to do an action.

The instruction set is created and stored, ready to be implemented. Picking a berry from a bush and eating it, for instance, involves moving the hand, grasping the berry, and pulling the hand back toward the mouth. Still, such a plan needs a decision to implement the instructions (implemental intention). Sometimes, these steps reveal that it is not feasible to accomplish the goal; thus, another goal is created. In Figure 10, to indicate this possibility, the triggering (dotted arrow) between goals and command spheres is bi-directional.

2: Creating a feasible goal triggers the “intention to act” (circle 3), i.e., a decision to execute the sequence of commands necessary to reach the goal. Thus, the process starts at point 3 (circle 4) when a decision is made to pick the berry.

The creation of that decision (circle 4) triggers (circle 7) the flow of instructions (release – circle 8) to the control sphere to be executed (circle 9). In the control sphere, each incoming instruction is processed (executed – circle 10). Execution means triggering the creation of a signal (point 11) that is transmitted to the hand (physical sphere). This triggers the creation of the required bodily movement (circle 12). The execution of the instruction cumulates in the actual movement of the hand (circle 13). This process is repeated for all instructions one at a time.

The evaluation side in Norman’s seven stages model contains three stages: perceiving the state of the world, interpreting the perception, and evaluation of the interpretation. This can be modeled as shown in Figure 11, using FM representation. To simplify the figure, the release and transfer/communicate stages of the FM model are not included.

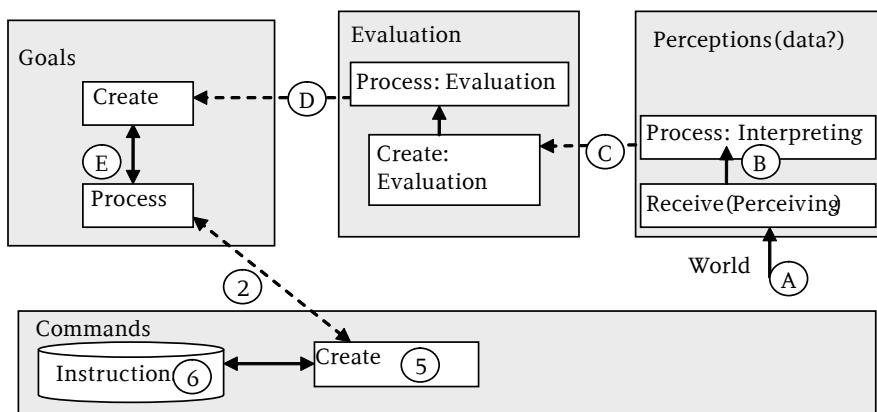


Figure 11. Stages of evaluation of action.

First, data is received from the field (state of the world – circle A in the figure). It is processed through interpretation (circle B). Perceiving and interpretation are two stages in the flow of data. Interpretation triggers (circle C) the creation of evaluation (metadata, e.g., evaluation reports). Evaluation is also processed to trigger the creation of goals. Goals are processed (circle D), and triggers a plan of realization are shown in the figure through circles 2, 5, and 6. These circles and the commands sphere are repeated from the previous figure, Figure 10.

This FM-based description provides a complete map of the sequence of flows involved in action. It clarifies the control notion and ties goals with control. For example, upon creation of a goal, it is processed and the sequence of commands to realize it is created and evaluated for feasibility.

Conclusion

This paper proposes to incorporate a flow-based model into the description of action. The model includes five stages of flow. We claim that action and all related concepts can be visualized as flowthings. Flowthings are things that are received, processed, created, released, and transferred.

We claim that this novel method of description provides a clearer and more complete specification of activities that are related to action. Beside the descriptive aspect, this paper does not claim any new contribution to the domain of action theory itself.

REFERENCES

- Al-Fedaghi, S.** (2008) "Systems of Things That Flow." The 52nd Annual Meeting of the International Society for the Systems Sciences, University of Wisconsin, Madison, USA, July 13–18, 2008. <http://journals.iss.org/index.php/proceedings52nd/article/viewFile/1001/354>.
- Ajzen, I.** (1985) From intentions to actions: A theory of planned behavior, in *Action-control: From cognition to behavior* (ed. J. Kuhl & J. Beckman), Springer, Heidelberg.
- Argyris, C. and Schön, D.** (1974) *Theory in practice: increasing professional effectiveness*. Jossey-Bass, San Francisco.
- Bock, C.** (2003) UML 2 Activity and Action Models, in *Journal of Object Technology*, 2(4), 43–53. http://www.jot.fm/issues/issue_2003_07/column3.
- Casni, J.D.** (2005), "Flow" Hits Its Peak. <http://metaphorobservatory.blogspot.com/2005/11/flow-hits-its-peak.html>.

Dick, B. and Dalmau, T. (1999) Values in action: applying the ideas of Argyris and Schön (2nd ed), Interchange, Chapel Hill, Qld.

Giacoppo, A. S. (2001). The Role of Theory in HCI, <http://www.otal.umd.edu/hci-rm/theory.html>.

Griffiths, R. (2009) Norman's Gulfs of Execution and Evaluation. <http://www.it.bton.ac.uk/staff/rmg/teaching/notes/NormanGulfs.html>.

Hacker, W. (1985) Activity: A fruitful concept in industrial psychology. In Goal directed behavior: The concept of action in psychology (ed. M. Frese and J. Sabini), Erlbaum, Hillsdale, New Jersey.

Heckhausen, H. and Kuhl, J. (1985) From wishes to action: The dead end and short cut on the long way to action. In Goal directed behavior: The concept of action in psychology (ed. M. Frese & J. Sabini), Erlbaum, Hillsdale, NJ.

Human Factors Laboratory (2004) National Institute for Aviation Research at Wichita State University, Norman's Action Cycle. <http://www.niar.twsu.edu/humanfactors/toolbox/Norman.htm>.

Li, L. (1999) Action Theory and Cognitive Psychology in Industrial Design: User Models and User Interfaces [Ph.D. thesis], <http://www.hbk-bs.de/bibliothek/ediss/data/19990630a/19990630a.pdf>.

Laurel, B. (1997) Computers as Theatre Reading, Addison-Wesley Publishing Company, Menlo Park, New York.

Norman, D.A. (1988) Psychology of Everyday Action. In The Design of Everyday Things, Basic Books, New York.

Nuttin, J. (1984) Motivation, planning, and action: A relational theory of behavior dynamics. LEA, Hillsdale, NJ.

Tian, Q. and Shi, Z. (1997) Model-theoretical foundation of action and progression, Science in China, 40(4).

Robbins, R. and Coulter, K. (2001) Management. Prentice Hall.

Rokeach, M. (1973) The nature of human values. Free Press, New York.

Searle, J. R. (1983) Intentionality: An essay in the philosophy of mind. NY: Cambridge University Press.

A frame for ethical evaluation of (information) technologies

Josep M. Basart*

Universitat Autònoma de Barcelona

Montse Serra**

Universitat Oberta de Catalunya

Computer Ethics (CE) can be included in a broader subject, namely Information Technologies Ethics (ITE), or even in a Global Information Ethics (Bynum and Moor, 2000, 275), in which the word technology indicates that computers belong to a very sophisticated field. Thus, the broadest area would be Technology Ethics (TE). Undoubtedly, many books and articles have been recently published about TE (Goujon and Dubreuil, 2001); in fact, this is the case even when we restrict ourselves to ITE—a good proof of its present importance is shown in (Tavani, 2008). Now, although we all could easily agree on what the word ethics means in these contexts, the same is not true when the word technology is considered. Especially when technology is approached not from a technical point of view, but from an ethical one.

Technology is a concept, a general label used to refer to a large number of quite different technologies. The point is, in our world we do not deal with abstract technology, but with particular technologies in particular contexts. Therefore, all general discourses about TE have the same initial—and often ignored—difficulty, namely what specific technology is analyzed and what its circumstances are. Without concreting some initial criteria of analysis, all ethical evaluations of technologies become very difficult and lack realism.

* *Josep Basart* Born in 1960 in Girona (Catalonia, Spain). Graduate in Informatics (1984), PhD in Informatics (1988), Graduate in Philosophy (1998), always at the Autonomous University of Barcelona. Full Professor at the Engineering School of the Autonomous University of Barcelona since 1990. His current research interests are: applied ethics, engineering ethics and philosophy of technology.

** *Montse Serra* - Computer Sciences Engineer by Universitat Autònoma de Barcelona (UAB)
- Master in Digital Communication by Universitat Autònoma de Barcelona (UAB)
- Master in Computer Sciences Security by Universitat Politècnica de Catalunya (UPC)
- Teacher of Computer Sciences, Multimedia and Telecommunication Studies of Universitat Oberta de Catalunya (UOC)

Some of the most recent works about ethics in technology (Arnold and Pearce, 2008; Verbeek, 2006) take into account how morality is present (“materialized”) in each technology, and how technologies moralize their surroundings. These works apply some concepts and ideas developed by M. Akrich, B. Latour and C. Venn (Akrich, 1992; Latour, 1992/1994/1999; Latour and Venn, 2002); for instance, the “script” concept. The script describes the many roles that technology can play in every context. Each technology brings a script inside itself with all the instructions (actions) to follow by the actors involved in it. It means, “Technologies are able to evoke certain kinds of behaviour” (Verbeek, 2006). Therefore, scripts are a way technology uses to influence human beings’ behaviour. In accordance with this approach, technologies act by themselves. They are not just tools and means at our service, they negotiate with us what is going to be and how it will be done. Like human beings, they are responsible actors mediating in human affairs. For this reason, to add technology as any other actor, becomes necessary because is not only a set of means used to do specific tasks, but also it is another active participant of a social and cognitive context. Technology and human beings cannot be moved away from each other, both entities shape a unit thanks to the aligned relationship established among them (Basart, 2008). Consequently, ethical criticism of technologies should not be limited neither to the accurate design of technologies nor to the proper use of them. A more complex approach is required.

In particular, designers can anticipate in how users will interact with a given technology. One way to carry out this anticipation is by using the ability to see another’s situation as if it were one’s own. With the help of the designer’s moral imagination (Coeckelbergh, 2006) it is quite possible to predict many implications and consequences in this complex relationship. While working in a design, technicians and engineers are adding prescriptions, therefore, they are delegating responsibility to technology. Thus, while technology needs to be morally evaluated, a responsibility in the designers’ work appears. The new perception where technology plays a mediating role in the actions of users impregnates the designer’s work with a revised moral dimension. “The fact that technologies always mediate human actions charges designers with the responsibility to anticipate these mediating roles” (Verbeek, 2006).

Some precedents

First systematic works in this field date back, at least, to the eighties (Kling, 1980). In his “Declaration of Empowerment” address to users in human-computer interaction, Shneiderman proposed (Shneiderman, 1990) to elaborate on the Social Impact Statement “[...] at the start of every human-computer interaction project”.

The framework proposed by Bologna (Bologna, 1991) focuses on some areas of ethical and legal concern, mainly from an instrumental approach to the computers (software applications, hardware, infrastructures...). In this sense his study goes very deep, but also the dynamic relationship between human beings, environment, and technology is missed.

Huff and Martin (Huff and Martin, 1995) presented the Project Impact CS “[...] for integrating social impact and ethics into the computer science curriculum”. This project recognizes that from the perspective of computer science, every ethical issue about a technology is located at a particular level of the social analysis. The framework provides the sort of comprehensive, conceptual overview that the field of computer science had been lacking until then. This framework considers nine ethical issues present in seven levels of social analysis (individuals, groups, organizations, cultures, institutions, nations and global). It takes into account that the analysis of any ethical issue needs to go in parallel with both social and technical analysis. The framework distinguishes two kinds of responsibilities: professional responsibility and individual responsibility. In some situations, it is quite easy to make such a distinction, but this is not always the case. In fact, which responsibility is paramount when complex ethical dilemmas arise in the engineering practice? What should be decided when contradictions appear? For this reason, in our proposal there is no clear-cut separation between a person as individual and the same person as a professional.

Gotterbarn and Rogerson (Gotterbarn and Rogerson, 2005) presented a Software Development Impact Statement that “[...] improves and expands risk perception” and “[...] should reduce the dangers of a narrow focus on quantitative risks” in software development. This work considers that “[...] any software project goes through three distinct phases: an initial phase where the feasibility of the project is examined; a requirements phase that lays out the overall structure and function of the project; and a detailed phase that lays out the plans for building the software. Each of these phases has its peculiar risks.” It is necessary to include in the evaluation “[...] social, professional, and ethical risks that lead to software failure and extends the range of stakeholders considered in risk analysis”. They argue that in a process of software development, it is necessary to make an actors’ list identifying all the stakeholders implied in the project because limiting the consideration of them to developers and clients can be a cause of failure. Another type of failure arises when developers limit the scope of software risk analysis to technical and economical issues. Therefore, a complete software development process requires the identification of all relevant stakeholders and their roles and enlarging risk analysis to include social, political, and ethical issues.

A new frame

In our proposal, it is essential to bear in mind how human beings and technologies are always bound together in a dynamic relationship. Technologies never exist alone, but as a more or less ambiguous opportunity to open new possibilities often unknown to us at the beginning. They act both actively and passively: actively, imposing their requirements and restrictions and passively, when used for attaining certain human purposes. Having this point of view in mind, it is possible to assess the ethical quality of a particular technology by considering to what extent this technology is in accordance with an established list of relevant ethical factors. The list is fixed in the sense that it does not change from technology to technology, but it is also open and may —should— be revised according to improvements in our ability to adjust our ethical assessments. The result of such an evaluation shows what conditions impose, what values promote, and what purposes seek the technology under scrutiny; in short, what its script is. By proceeding in this way, ethical evaluations may result in being clearer, more realistic, systematic and fairer —because the questions are always the same for every technology, whatever its particular characteristics are. The current list of eleven ethical factors is as follows:

- (1) Does it answer or help to answer to an existing demand?
- (2) Does it make understanding and cooperation easier among people trying to manage a problem?
- (3) Does it point towards the characteristics and demands of its users? Is it flexible, easy to adapt to changes in its environment and to new requirements?
- (4) Does it promote users' autonomy?
- (5) In what ways does it increase the welfare of its users?
- (6) Does it respect valuable social practices and universal human rights?
- (7) Does it keep working under human control?
- (8) Can it be integrated in other well-founded technologies? Is it easy to update and maintain?
- (9) Is it constrained to heavy demand of resources when it is manufactured or used?
- (10) What are the undesired effects on people or the environment?
- (11) To what degree is it recyclable and/or reusable?

It is necessary to specify that here the autonomy to be considered in factor (4) is that showed by Shneiderman (Shneiderman, 1990) “[...] users want to be empowered by technology to be able to apply their knowledge and experience to

make judgments that lead to improved job performance and greater personal satisfaction". From an ethical perspective, this is always a central issue in the relationship between individuals and technology.

Now, due to the importance of the codes of ethics in the engineering profession (Lynch and Kline, 2000), and being aware of the need to frame the basic values, ethical principles and standards through them, it is very suitable to show how our eleven ethical factors presented here are according to the spirit of some of the main codes of engineering ethics. In the following, three well-known codes of engineering ethics have been revised: Association for Computing Machinery (ACM), Institute of Electrical and Electronics Engineers (IEEE), and National Society of Professional Engineers (NSPE).

The relevant statements of these codes with regard to our proposal are the following: ACM 1 Contribute to society and human well-being. 1.4 Be fair and take action not to discriminate. 2.1 Strive to achieve the highest quality, effectiveness and dignity in both the process and products of professional work. 2.7 Improve public understanding of computing and its consequences. 3.5 Articulate and support policies that protect the dignity of users and others affected by a computing system.

IEEE 1 To accept responsibility in making decisions consistent with the safety, health and welfare of the public, and to disclose promptly factors that might endanger the public or the environment. 5 To improve the understanding of technology, its appropriate application, and practical consequences. 8 To treat fairly all persons regardless of such factors as race, religion, gender, disability, age, or nation origin.

NSPE I.1 Hold paramount the safety, health, and welfare of the public. III.2.a Engineers are encouraged to participate in civic affairs [...] III.2.b Engineers shall not complete, sign, or seal plans and/or specifications that are not in conformity with applicable engineering standards. III.2.d Engineers are encouraged to adhere to the principles of sustainable development in order to protect the environment for future generations.

The following table summarizes the results of the comparisons:

	ACM	IEEE	NSPE
(1) Does it answer or help to answer to an existing demand?	2.1	*	III.2.b
(2) Does it make understanding and cooperation easier among people trying to manage a problem?	*	*	*

(3) Does it point towards the characteristics and demands of its users? Is it flexible, easy to adapt to changes in its environment and to new requirements?	2.1	*	III.2.a
(4) Does it promote user's autonomy?	3.5	*	*
(5) In what ways does it increase the welfare of its users?	1	1	I.1
(6) Does it respect valuable social practices and universal human rights?	1.4	8	III.2.a
(7) Does it keep working under human control?	2.7	*	*
(8) Can it be integrated in other well-founded technologies? Is it easy to update and maintain?	2.1	*	III.2.b
(9) Is it constrained to heavy demand of resources when it is manufactured or used?	*	*	III.2.d
(10) What are the undesired effects on people or the environment?	2.1	5	III.2.d
(11) To what degree is it recyclable and/or reusable?	*	*	*

(An asterisk means that there is no obvious correspondence between the ethical factor and the concerns shown in the professional code.)

It is significant that two of the eleven ethical factors (numbers (2) and (11)) receive no support from any of the three ethical codes. Moreover, three of the factors (numbers (4), (7) and (9)) receive some support from just one of the codes. We consider that this is not a bad news for the frame, quite the contrary. In addition, it shows an important characteristic of many professional codes. Both assertions are closely related and need to be justified.

To begin with, it is convenient to distribute these five ethical factors in two groups. The first group contains numbers (9) and (11), both of them pointing to environmental concerns. The second group contains the other three factors: numbers (2), (4) and (7). These are related to —let us say— users' advantages concerns. On the one hand, until recently environmental concerns had not been introduced in professional codes. This was considered, and still is, a huge social and political problem lying outside of the ethical professional context. However, as the NSPE code recognizes in III.2.d, there is also a clear responsibility in the work of engineers, both to the environment and to the future generations. Many professional decisions, practices, and attitudes have an influence on the problems of waste, squandering of resources, and contamination. Probably, in the near future, many professional codes will include this worry among the most relevant. On the other hand, with regard to people, professional codes in engineering are directed towards not to hurt anyone. Of course, this is a paramount concern, but

just a negative one. It is also very important to think with a positive will how to benefit people through the professional activity. This new commitment means that perhaps professional ethics deserves a still more serious consideration.

An application to IT

Now it is possible to illustrate how the analysis of these questions inside the IT environment shows many of the key ethical points to consider in these technologies. As an example, the exploration could proceed as follows.

- (1) Is this technology necessary or is it just another model or version of something already existing? Is it just business in the cycle of production and consumption? Does it contribute to solve a real social necessity? Novelties seem to be indispensable in the market; "New!" appears to be the best advertising, but newer is not necessarily synonymous with better.
- (2) Does this technology facilitate and enhance communication among people? From an ethical analysis, communication is not just transmission of information. In a communication environment, quality (what and how it is said) is often more important than quantity (how many bits are sent).
- (3) To what extent does the design of these systems and devices take into account the physical and psychological structure of human beings? Do people adapt to the machine rather than the reverse?
- (4) Does this technology produce an empowerment of the users' capabilities with regard to working, communicating or managing information?
- (5) Are there clear benefits for people, or are we talking of benefits for systems and machines? The latter does not always imply the former.
- (6) To what extent are trust, privacy, reliability and security protected?
- (7) The high level of automation that computers have introduced should not be an excuse to elude accountability and responsibility. Nevertheless, this is possible only when there is someone making a final decision whenever it is necessary.
- (8) IT opens a vast and not yet explored horizon of new possibilities. However, its success often depends on its capacity to operate well with other digital technologies. The procedures to maintain and to update these technologies are not just technical possibilities that can be more or less fulfilled. The ethical claim is that these procedures must exist and have to be as effective as possible.

- (9) Again, consumption of resources is not just an economical matter, but also an ethical one. This is because today more than ever, we all share a world of scarce resources where saving (rational and reasonable spending) is not optional.
- (10) Disposal of electronic waste damages the environment, while the accumulation of radiation emissions from electronic products may be potentially dangerous in the long term. In addition, an important point to analyze is how information transmission devices have an effect on human communication and human relation patterns.

Finally, (11) Recycling is also necessary to reduce the huge economical and ecological impact of IT as much as possible. As an example, restricting ourselves to domestic goods, we can find personal computers, laptops, cellular phones, TV sets, video cameras, video games, digital players and so on.

REFERENCES

- M. Akrich**, "The De-Description of Technological Objects". In *Shaping Technology-Building Society. Studies in Sociotechnical Change*. W. Bijker and J. Law (eds.), pp. 205-224. Cambridge, MA: MIT Press, 1992.
- M. Arnold and C. Pearce**, "Is Technology Innocent? Holding Technologies to Moral Account", *IEEE Technology and Society Magazine*, vol. 27, no. 2, pp. 44-50, 2008.
- J. M. Basart**, "Ethics Applied to Technologies — Is All Well?", *IEEE Technology and Society Magazine*, vol. 27, no. 4, pp. 45-49, 2008.
- J. Bologna**, "A Framework for the Ethical Analysis of Information Technologies" *Computers & Security*, vol. 10, pp. 303-307, 1991.
- T. W. Bynum and J. H. Moor** (eds.), *The Digital Phoenix. How Computers Are Changing Philosophy*. Oxford: Blackwell, 2000.
- M. Coeckelbergh**, "Regulation or Responsibility? Autonomy, Moral Imagination, and Engineering", *Science, Technology, & Human Values*, vol. 31, no. 3, pp. 237-260, 2006.
- D. Gotterbarn and S. Rogerson**, "Responsible risk analysis for software development: Creating the Software Development Impact Statement", *Communications of the Association for Information Systems*, vol. 15, pp. 730-750, 2005.
- P. Goujon and B. Hériard Dubreuil** (eds.), *Technology and Ethics. A European Quest for Responsible Engineering*. Leuven: Peeters, 2001.

C. Huff and D. Martin, "Computing Consequences: A Framework for Teaching Ethical Computing", *Communications of the ACM*, vol. 38, no. 12, pp. 75-84, 1995.

R. Kling, "Social Analyses of Computing: Theoretical Perspectives in Recent Empirical Research", *ACM Computing Surveys*, vol. 12, no. 1, pp. 61-110, 1980.

B. Latour, "Where are the missing masses? The Sociology of a Few Mundane Artifacts". In *Shaping Technology-Building Society. Studies in Sociotechnical Change*. W. Bijker and J. Law (eds.), pp. 225-259. Cambridge, MA: MIT Press, 1992.

B. Latour, "On technical mediation — Philosophy, sociology, genealogy", *Common Knowledge*, vol. 3, no. 2, pp. 29-64, 1994.

B. Latour, *Pandora's Hope: Essays on the Reality of Science Studies*, Cambridge, MA: Harvard University Press, 1999.

B. Latour and C. Venn, "Morality and Technology. The End of the Means", *Theory, Culture & Society*, vol. 19, no. 5/6, pp. 247-260, 2002.

T. Lynch and R. Kline, "Engineering Practice and Engineering Ethics", *Science, Technology, & Human Values*, vol. 25, no. 2, pp. 195-225, 2000.

B. Shneiderman, "Human Values and the Future of Technology: A Declaration of Empowerment", *ACM SIGCAS Conference: Computers and the Quality of Life*, Sept. 1990.

H. T. Tavani, "ICT ethics bibliography 2006-2008: A select list of recent books", *Ethics and Information Technology*, vol. 10, no. 1, pp. 85-88, 2008.

P.-P. Verbeek, "Materializing Morality. Design Ethics and Technological Mediation", *Science, Technology, & Human Values*, vol. 31, no. 3, pp. 361-380, 2006.

Notes

<http://www.acm.org/about/code-of-ethics>

<http://www.ieee.org/web/aboutus/ethics/code.html>

<http://www.nspe.org/Ethics/codeofEthics/index.html>

The Greek scale of attitudes towards unethical behaviors on the Internet

George Briskolas*

Psychologist

&

Petros L. Roussos**

Department of Psychology, School of Philosophy,
University of Athens, University Campus

Abstract

The purpose of the present study was to develop an attitudes scale towards unethical behavior over the web for the Greek population. 124 items were constructed and the initial questionnaire was administered to 375 computer users who were asked to use a 5-point Likert response scale to indicate their level of agreement or disagreement with each of these items. 24 items were retained and the final version of the scale that was developed had three subscales: distribution of intellectual property, Internet safety and hacking.

Keywords: Internet ethics; Attitudes; Measurement scale

Introduction

During the last 25 years the Internet has become an important part of our lives, changing the way that computer users handle information (Cooper, 2004; Copeland, 2004). The most important change concerns the access to an infinite amount of information. In addition, Internet offers an alternative way of human interaction. There are three distinct characteristics of the Internet: the potential

* George Briskolas is a psychologist (University of Athens, Greece). His master's degree is in cognitive science (University of Athens). He is, also, trained in cognitive-behavioral psychotherapy. His research interests include the effects of computer technology (especially Internet) on human behavior.

** Petros Roussos is a lecturer of Cognitive Psychology at the University of Athens, Greece. Since 1994 he has taught cognitive psychology, research methods and statistics in psychology at the University of Crete (Greece), the University of the Aegean (Greece), the University of Athens and many other colleges and higher education institutions. His research interests include the effects of ICT on students' cognitive abilities, applications of ICT in education, HCI with emphasis on teaching and learning, systems and educational software ergonomics, statistical reasoning.

of global communication among many users at the same time, the potential of anonymity and the potential of infinite reproducibility of information (Johnson, 2000, 2005; Stamatellos, 2007).

The major developments caused by the Internet are related to a number of different aspects of daily life and also affect users' ethical behaviour when interacting with a computer (Ess, 2008; Floridi, 2008; Kizza, 2003). For example, the 'cookies' that can track the preferences of Internet users or monitor the computer of an employee is only one of the emerging ethical issues related to computer and Internet use (Kierkegaard, 2005). The related literature refers to many issues such as: intellectual property rights, hacking, privacy of personal information, perversion, anonymity, cybercrime, ethical responsibility for computer professionals (Cooper, 2004; Ellerman, 1998; King & King, 2000; Mason, 1986; Raymond, 1996; Spafford, 1992; Tavani, 1999). Computers have created new species of problems in areas such as: speed/reflex, storage/privacy, identity theft, internationality, copying/stealing, pornography, gambling, stalking, gender, race and social class, selling private data, and opt-in versus opt-out for solicitation. However, the list does not exhaust the possibilities for new species of problems (Barger, 2008). Given these concerns, it is of crucial importance to understand what might affect an individual user's ethical behavior and intentions.

The first issue raised has to do with the different way that technology, and specifically computers, affects ethics, so that they should be studied as a separate part of ethical behavior. According to Johnson (2004, 2005), there are at least two arguments pointing the need for a separate study of computer ethics and both derive from the philosophy of 'action theories' (Hornsby, 1980; Searle, 1983). It should be noted here that in 'action theories' responsibility of moral agents has an enriched meaning that is going further from the simple responsibility which derives from the voluntary intended behaviors of a person. In 'action theories' the causes of an action are defined by the intentionality of internal mental states such as 'intendings', desires and beliefs (Johnson & Powers, 2005). The first argument refers to the fact that technology offers to people the possibility to do things that were impossible to do in the past (Gottterbarn, 1992; Johnson, 2004; Moor, 1985). This is also called by Moore (2008) the "informationalization" of a task. For example, the control of traffic lights from computers helps people regulate the street traffic, or the advances in medical technology, like an fMRI, offer the capability to monitor the organs of a patient. As for the computer technology and in extend the Internet, someone can also easily track the new ways they have offered us to perform our daily actions (Johnson, 2004). For example, computers and the Internet have facilitated an increasingly speedy form of communication.

Another significant difference in the area of computer ethics is the fact that it is very easy to any computer user to harm other people by simply releasing a virus from his computer (Johnson, 2004; Mitcham, 2004). This situation provides to the human moral agent a new perspective, a new tool to cause many problems just by clicking few buttons without being necessary to have any special knowledge regarding computers. So, from the moment that human action changes, it is of big importance to see how technology defines ethics or ethics defines and, in extend, shapes the steps of progress of technology.

However, there is not a universal agreement that computer technology generates wholly new ethical problems. Johnson (2000), for example, argued that computing technology could alter old ethical problems in interesting and important ways and thereby “give them a new twist.”

Moor (1985) defined computer ethics as ‘...a field concerned with ‘policy vacuums’ and ‘conceptual muddles’ ...’. Moor (1985) also refers to two significant stages. The first stage concerns the ‘technological introduction’, that has already occurred after the Second World War, and the stage of ‘technological permeation’ of every aspect of our daily life. As a result of this, Moor (1985) suggests that fundamental concepts, such as ‘money’, will eventually alter. According to Bynum’s (2001) definition, computer ethics ‘... identifies and analyzes the impacts of information technology on such human values health, wealth, work...’

There are three important issues that have attracted most of the attention on computers ethics and on which the rest of this section will focus: intellectual property, privacy, and hacking (Mason, 1986).

Intellectual Property

Intellectual property is considered -in general terms- the nonphysical property (Kimppa, 2005b). Specifically, it is the product of cognitive processes whose value is based upon some idea or collection of ideas (Moore, 1997, 2008). In contrast with the ownership of a physical property, intellectual property can be considered as a public good (George, 2008; Kimppa, 2005b; Spinello & Tavani, 2005). There are two distinct characteristics for public goods: firstly, a material good cannot be possessed by two people, whereas a public good can be owned by many persons, e.g. everyone can possess a copy of the same book. Secondly, a public good can be used by everyone.

The problem with intellectual property is to define it precisely, and also define the legitimate rights of its owner (Akester, 2004; Carlisle, 1999; Kimppa, 2005a). It is somehow difficult to ‘possess’ an idea or a concept. The practical way to copyright intellectual property is to express it in a ‘physical’ form and of

course be the first doing it, e.g. to publish a book or produce a music cd (George, 2008; Himma, 2007).

In Locke's (1988) theory for property the main argument about -physical- property is that every person has natural rights over the products of his labour. This argument is based on two aspects. The first is that a person is the owner of his body and in extend owns the work produced from his labour. The second aspect is that the right to have his own property comes as a result of his hard labour (Kramer, 1997; Lampert, 1997; Moore, 1997; Simmons, 1992). Therefore, according to Lockean theory, the right to property is inviolable as the human right to freedom. The physical property as a natural right for humans must be protected. The limits of the personal property are regulated by the level of personal labour made by the person. If someone, for example, tills a piece of land he is the natural owner of the land and its products. So the logic that underlies here is that the possession of something has to do with the effectiveness and the purpose of the labour (Ashcraft, 1992; Kramer, 1997; Zack, 1992). Stealing is considered any act that has to do not only with the property of someone, but also with his products (Ashcraft, 1992; Scanlan, 2005). Another important point to Lockean theory is the moral view of the degree of natural goods that someone can possess, known as the 'Lockean Proviso' (Oksanen, 1997). The idea here is that people should not be insatiable with the exploitation of the earth and they should take as much goods as they need. The main argument against Lockean theory is that it only refers to physical property and not to intellectual property, and it is insufficient to cover the wide area and the problems that concern intellectual property (Kimmipa, 2005a; Spinello & Tavani, 2005). One of the problems related to intellectual property is that in physical property there can be only one owner of each good, something that does not apply to the case of intellectual property. The 'Lockean Proviso' for wise use of natural resources does not apply to intellectual property, because the production of intellectual goods can be infinite (Kimmipa, 2005b).

According to Hegel's theory, there are unbreakable bonds between property and the human personality (Becker, 1997). The underlying hypothesis here is that the development to self-actualization can be achieved by personal expression to external objects. Therefore, the person must control natural resources and possess goods, tangible and intangible (Moore, 2008). The action to external objects is essential to Hegel's theory because without property self-expression does not exist and without self-expression there is no personal freedom. In other words, property is considered to be a natural right that leads to freedom by leading the person to objectify and externalize his personality. As a result, the theory applies both to natural and intellectual property. A problem to this theory is the difficulty of finding an objective way to measure and quantify the self-expression that could be used as a base for the assignment of intellectual rights (Lampert, 1997; Spinello & Tavani, 2005).

The criterion used in utilitarian theories is the moral evaluation of any human action according to its social utility (Goodin, 1995; Howard-Snyder, 1994). As a result, the intellectual property is defined by the degree to which a social good can offer to the society. The general idea behind utilitarian theories is that people need to acquire and use goods in order to reach a level of happiness and fulfill their wishes. Because of the insecurity that derives from personal possession, it is essential to protect possession and its use, and control its products. So the security of the goods resides in a system that defines property rights.

A version of the utilitarian theories is the incentive theories (Adams, 1976; Moore, 2008). There are four basic elements in these theories: the first element is that society should create such constitutions that would, or expect to, lead to the maximization of overall social utility. The second element is that giving rights to authors and inventors over their works is very useful and incentive to future production of intellectual works. The third element is the amelioration of social prosperity as result of motivating the creation and production of intellectual works. Therefore, as an overall result it is useful the adoption of a system of intellectual property.

In general, the basic disadvantage with utilitarian theories is the lack of empirical data which will correlate the need of constitutions for intellectual works with social prosperity (Hooker, Mason & Miller, 2000; Scheffler, 1982; Spinello & Tavani, 2005). Also, it is quite difficult to foresee the limits of protection of intellectual property, so that it could provide incentive for the production of other social goods, or which would the effects of creating such a system to the creators and to the public.

Privacy

The world of computer technology that permits the creation of enormous databases with personal information, poses a new dimension to the big issue of privacy. In the past the collection of personal information was a government affair. Nowadays, computer technology and the Internet can provide access to personal information, legally or illegally, virtually to anyone.

One of the most important issues in computer ethics is that of privacy. The two basic problems with privacy are related to its conceptual framework and its value (Gritzalis & Lambrinouidakis, 2008; Tavani, 1999; Van den Hoven, 2008). In the effort to define the notion 'privacy' there are two main views concerning whether we have to do with an independent notion or a notion that is part to other concepts like 'freedom' or 'property'. There are three distinct characteristics of privacy: personal identity, autonomy and social relationships (Kierkegaard, 2005; Kizza, 2003; Nissenbaum, 1998). The above three elements have to do with the way that a person develops his personality according to his independent personal

characteristics, how he shares that information in the social web he lives and how he interacts with other society members. As a result, Internet privacy deals with the degree that a person can give his personal information with safety, e.g. buy something online or simply communicate with someone (Kizza, 2003).

There are four distinct kinds of privacy proposed by Floridi (Tavani, 2008). The first, psychical-accessibility privacy, refers to the degree that someone can be, by choice, bodily intact from the others. The second, decisional privacy, refers to the freedom a person has to make his own choices on important matters, for example on education and work, excluding others from his decisions. The third, psychological/mental privacy, refers to the person's capability to remain intact from psychological interference from others. The fourth, informational privacy, refers to the control of access in personal information, for example information about the personal lifestyle of someone or his medical history. In general, informational privacy affects four broad categories: consumer privacy, medical privacy, employee privacy and location privacy (Tavani, 2008).

Cyber-crime and hacking

The Internet has brought many and quite important changes on our lives affecting us not only in a positive way, but also in a harmful way, which is called cybercrime. Hacking (as it is widely used) or cracking is called *-grosso modo-* the illegal internet activity (Crowell, Narvaez & Gomberg, 2005; King & King, 2000; Raymond, 1996). It usually refers to the intrusion and unauthorized access via Internet to other users' computers. Cybercrime activities focus mainly to economic crimes (e.g. stealing credit card numbers). Another common form of cybercrime is the release of a virus that can crash or destroy the data of every computer that has infected (Johnson, 2004; King & King, 2000). In conceptual level a question to be answered is whether there is correlation between common crimes, as theft or sexual harassment, and illegal activities with Internet activities (Johnson, 2004). For example, the illegal download of a music cd should be considered as the same crime with the theft of things from a house? There are two basic points to these kinds of question (Johnson, 2000, 2004; Kimppa, 2005a, b). The first is that in cases like the above the hacker does not deprive from the owner the right of use of his possession. The second is the difficulty of finding the hacker because of the online anonymity.

Except from the mentioned forms of cyber attacks, some other kinds of cyber attacks pose new ethical issues, and in final analysis new dilemmas whether they are illegal or not. Three main areas of cyber attacks are to be mentioned here (Denning, 2008; Freeman & Peace, 2005; Kimppa, 2005a; Spafford, 1992). The first area covers the cyber attacks in the interests of national security. The second area covers the cyber attacks with political or social motivation, and is referred with the term 'hacktivism' and sometimes as 'cyberterrorism' if the result of the

attack is harmful for civilians. The third area covers cyber attacks for security reasons or cyber defence in order to find the flaws of a system.

Aims of the present study

A review of the roots and current direction of computer ethics suggests that its critical potential is yet to be realized. Ethics, as a philosophical subject, attends to theories of morality and, hence, it is normative rather than descriptive. The position taken here is that ethics, particularly an applied ethics, such as computer ethics can be a potentially potent political force as ethical debate feeds into policy and, ultimately, into legislation. This signals a need to find explanations for such behaviour, otherwise policy and legislation designed to regulate human conduct are unlikely to prove effective. Central to this paper is the concern that researchers have overlooked the area of users' attitudes towards unethical behavior over the Web. Relatively little empirical research has been conducted during recent years in order to increase our understanding over this topic.

In order to encourage high quality research, enable integration and consistency across research studies, and increase understanding of users' perceptions and attitudes towards unethical behaviors on the Internet, there is a need for valid and reliable multiple-item measures for this construct. A fast, effective measure of Internet users' attitudes is crucial to the study of the extent and the manner in which we use the Internet. Ideally, this instrument should also be short, efficient and easy to administer to a wide range of people. The purpose of this paper was to develop a Greek measure of Internet users' attitudes and beliefs; one that will be useful with members of the general population.

Method

Participants

The scale was administered to a large convenience sample of 375 participants who were selected from a number of cities in many different parts of Greece. The majority (67.7%) was females and all were computer users (191 of them had over five years of Internet experience and only 35 had an experience shorter than a year). Most of them were undergraduate and postgraduate university students (83.7%) and the age range was between 18 to 63 years (mean = 31.3, SD = 9.3 years).

Instrument development

The first step was to draw the main outlines from the current literature of computer ethics and especially the issues concerning internet usage (Bynum, 2001; Johnson, 2004; Johnson & Powers, 2005). Four main issues derived: illegal distribution of copyrighted material, unethical behaviors concerning e-mail usage, hacking and unethical behaviors through chat-rooms. The second step was to create a pool of positive and negative statements (items) related to these issues. The items were then reviewed, revised, edited and 124 of them were finally put in the questionnaire in random order. A 5-point Likert response scale that ranged from 'strongly disagree' to 'strongly agree' was used to indicate the level of agreement or disagreement with each of the items. Items were screened for their tendency to elicit extreme responses, items being excluded if they produced mean responses of more than four or less than two on the 5-point Likert-type scale employed. Twelve questions were excluded on these grounds. A principal components factor analysis was used on the data for item analysis (Jackson, 1991; Jolliffe, 2002). The factor load was set to be over 0.30.

Results

Principal component factor analysis was carried out for the 124 items covered in the scale. Kaiser-Meyer-Olkin (KMO) value was 0.87 and the Bartlett's test of sphericity was $\chi^2(6216)=21316.84$, $p<.001$, showing that the principal component analysis was appropriate for the analysis. Communalities were all above 0.560 and intercorrelations were in normal levels varying between 0.30 and 0.80. As for the internal reliability, Cronbach's α was 0.88. Three factors were extracted explaining 27.05% of the total variance. Twenty four items (8 for each of the three factors) were retained. The retained items, the corresponding factors, factor loadings and descriptive statistics for the three extracted factors are displayed in Table 1.

The first factor was 'Distribution of intellectual property' (Mean = 29.25, s.d. = 5.8), the second factor was 'Internet safety' (Mean = 27.94, s.d. = 4.5), and the third factor was 'Hacking' (Mean = 21.63, s.d. = 4.9).

The first factor explained 17.8% of the total variance and its Cronbach's α was 0.88. The second factor explained 5.4% of the total variance ($\alpha=0.70$) and the third explained 3.8% of the total variance ($\alpha=0.72$).

Table 1.

Retained items and factor loadings of the Greek scale of attitudes towards unethical behaviors on the Internet

Items		Factor 1	Factor 2	Factor 3
Factor 1: Distribution of intellectual property ($\alpha = 0.88$, Mean = 29.25, s.d. = 5.8)				
60	I believe that the free distribution of software over the web is not a serious offense	0.764		
52	The free distribution of movies, TV shows e.tc. over the web does not annoy me	0.730		
100	The free distribution of music files over the web is justified when it is for personal use only	0.723		
104	The free distribution of e-books or journals over the web is justified when it is for personal use only	0.714		
55	I believe that the free distribution of e-books or journals over the web should be allowed	0.691		
88	The free distribution of software over the web is justified by the high market prices	0.644		
62	I do not consider myself a concealer when I download and use intellectual property (for example, music, movies, software)	0.622		
40	The fact that music stars gain high financial rewards along with the high prices of music CDs make me mad, so I prefer to download music for free	0.591		
Factor 2: Internet safety ($\alpha = 0.70$, Mean = 27.94, s.d. = 4.5)				
106	I would never provide my credit card information over the web in case it was stolen		0.526	

82	I avoid using electronic services (e.g., e-tax, e-banking, e-shopping) because I am afraid that hackers may steal my personal info		0.485	
114	Many people use chat rooms to behave in an indecent way		0.466	
90	E-shopping is dangerous		0.434	
93	The Internet enables an international organization of criminal transactions		0.398	
15	Email messages from usually contain viruses		0.373	
28	Many people use chat rooms to do punishable acts		0.362	
25	It is easy for someone to access my computer files over the web		0.349	
Factor 3: Hacking ($\alpha = 0.72$, Mean = 21.63, s.d. = 4.9)				
13	Hackers help in improving the Internet (for example, they discover vulnerable spots)			0.530
42	Hacking may also have honorable intentions			0.515
115	Hackers help in improving computer software (for example, they expose security problems in antivirus software)			0.458
3	I would approve hacking a webpage in case the intervention was not malicious			0.451
51	Hacking is right when is directed against large companies			0.383
39	Hackers are the Internet 'rebels'			0.373
30	I would approve hacking a webpage in case I had a financial interest			0.311
97	I would approve hacking a webpage in case this was done for national security reasons			0.302

Discussion

The Greek scale of attitudes towards unethical behaviors on the Internet is a Likert scale devised to measure computer users' beliefs and attitudes. The present paper focused on the construction and reliability analysis of the instrument, but it is clear that further studies are necessary in order to test its validity and the effects of various independent variables on the construction of these beliefs.

Three dimensions were determined as a result of the factor analysis carried out. The first factor (distribution of intellectual property) is related to one of the most complex issues faced in the modern society. Information is very hard to safeguard and hard to keep to one's self (Mason, 1986). Just a few years ago it was Napster that started a revolution by enabling the distribution of almost every type of file, mostly of mp3 music files (Mortensen, 2005). Nowadays, the number of those users who illegally download music, software, movies, e-books, e.tc. without any reservation has grown significantly (Cooper, 2004; Craig, Burnett & Honick, 2005; Im & Van Epps, 1992; Kimppa, 2005b). The fact that most of the participants in the present study did not seem to consider this behaviour as unethical needs further study.

The second factor was named as 'Internet safety' since it included items related directly with the safety of a number of activities over the web. A number of important issues related to websites used for social networks (e.g., Facebook, Hi5, Myspace) have recently attracted the users' interest. For example, over the last months a long discussion has taken place over Facebook's decision to sell the personal data of its members to a number of companies for commercial reasons (Tavani, 2008). Nowadays a huge amount of personal information – such as passwords, social security numbers, driver's license numbers, credit card numbers, e.tc. are available on the Web. These can be bought and used by thieves to provide authentication for criminal transactions. Sometimes personal information is obtained by 'phishing' and sometimes, of course, the information is simply discovered off-line.

Finally, the third dimension, hacking, included items related to the most frequent hacker arguments that maintain that computer break-ins are ethical. Specifically, the reasons hackers use as excuses for computer break-ins are to expose security problems, to check others' computers to protect them against misuse of their data by the governments or big corporations, and so on. Interestingly, the participants of the present study seemed to disagree with these statements.

As with any empirical study, limitations do exist and further questions remain. The lack of similar studies and psychometric tools in the literature poses a problem when testing the validity of the scale. One further limitation of such a scale

is the fact that due to the constant changes and advances on the Internet it is time constrained. Therefore, a scale that today measures users' attitudes towards a number of behaviors on the Internet, in few years or even months it might be insufficient.

The next step of this study should work on providing more strong evidence about the validity of the scale. Researchers are strongly encouraged to use this scale to further examine its reliability and validity. Future research should continue to test the possible relationships between users' beliefs and attitudes towards unethical behaviors on the Internet and variables such as users' personality characteristics, computer experience and self-efficacy, sex and age. In addition, further research should seek to investigate the stability of users' attitudes toward these behaviors over time, and develop normative data for different occupational, educational, and socioeconomic groups. Also, the relationship between attitudes and behavior in a true work environment would be of interest. Finally, a comparative study of these questions on an international basis under the light of factors such as society and culture would further aid in the understanding of computer ethics in today's electronic society.

REFERENCES

- Adams, R.** (1976) Motive Utilitarianism. *Journal of Philosophy*, 73, 467-81.
- Akester, P.** (2004) Authorship and authenticity in cyberspace. *Computer Law & Security Report*, 20, 6, 436-444.
- Ashcraft, R.** (1992) The politics of Locke's two treatises of government, in John Locke's "Two treatises on government": new interpretations (ed. E. Harpham), Allen Press, Lawrence.
- Barger, R. N.** (2008) *Computer ethics: a case-based approach*. Cambridge University Press, Cambridge, UK.
- Becker, L.** (1977) *Property rights: philosophical foundations*. Routledge, London.
- Bynum, T.** (2001) Computer ethics: its birth and its future. *Ethics and Information Technology*, 3, 109-112.
- Carlisle, J.** (1999) Ethical considerations of the software-dependent organization. *Journal of Systems and Software*, 44, 251-255.
- Cooper, W.** (2004) Internet culture, in *Philosophy of computing and information* (ed. L. Floridi), Blackwell Publishing, Malden/Oxford.
- Copeland, J.** (2004) Computation, in *Philosophy of computing and information* (ed. L. Floridi), Blackwell Publishing, Malden/Oxford.

Crowell, C., Narvaez, D. and Gomberg, A. (2005) Moral psychology and information ethics: psychological distance and the components of moral behaviour in a digital world, in *Information Ethics: privacy and intellectual property* (eds. L. Freeman & G. Peace), Information Science Public, Herhsey/London.

Denning, D. (2008) The ethics of cyber conflict, in *The handbook of information and computer ethics* (eds. K. Himma and H. Tavani), Wiley & Sons, New Jersey.

Ellerman, E. (1998) The Internet in context, in *Psychology and the Internet* (ed. J. Gackenbach), Academic Press, California.

Ess, C. (2008) Culture and global networks: hope for a global ethics, in *Information technology and moral philosophy* (eds. J. Van den Hoven & J. Weckert), Cambridge University Press, Cambridge.

Floridi, L. (2008) Information ethics: its nature and scope, in J. Van den Hoven & J. Weckert. (Eds.), *Information technology and moral philosophy* (40-65). Cambridge: Cambridge University Press.

Freeman, L. and Peace, G. (2005) Revisiting Mason: the last 18 years and onward, in L. Freeman & G. Peace (Eds.), *Information Ethics: privacy and intellectual property* (1-18). Herhsey/London: Information Science Public.

George, A. (2008) Intellectual property and the Internet, in *Internet technologies and applications* (eds. M. Freire & M. Pereira), Information Science Reference, Hershey.

Goodin, R. (1995) *Utilitarianism as a Public Philosophy*. Cambridge University Press, New York.

Gotterbarn, D. (1992) The use and abuse of computer ethics. *Journal of Systems and Software*, 17, 75-80.

Craig, P., Burnett, M. and Honick, R. (2005) *Software piracy exposed*. Syngress Publishing, Rockland.

Gritzalis, S. and Lambrinoudakis, C. (2008) Privacy in the digital world, in *Internet technologies and applications* (eds. M. Freire & M. Pereira), Information Science Reference, Hershey.

Himma, K. (2007) Justifying intellectual property protection: why the interests of content creators usually win over everyone else's, in *Information technology and social justice* (eds. E. Rooksby & J. Weckert), Information Science Publishing, Hershey.

Hooker, B., Mason, E. and Miller, D. (2000) *Morality, rules, and consequences*. Edinburgh University Press, Edinburgh.

Hornsby, J. (1980) *Actions*. Routledge & Kegan Paul, London.

Howard-Snyder, F. (1994) The Heart of Consequentialism. *Philosophical Studies*, 76, 107-29.

Im, J. and Van Epps., P. (1992) Software piracy and software security measures in business schools. *Information & Management*, 23, 193-203.

Jackson, E. (1991) *A user's guide to principal components*. John Wiley & Sons, New York.

Johnson, D. (2000) *Computer ethics* (3rd ed.). Prentice Hall, New Jersey.

Johnson, D. (2004) *Computer ethics*, in *Philosophy of computing and information* (ed. L. Floridi), Blackwell Publishing, Malden/Oxford.

Johnson, D. (2005) *Computer Ethics*, in *Encyclopedia of science, technology and ethics*, Thompson Gale, Farmington Hills.

Johnson, D. and Powers, T. (2005) *Ethics and technology: a program for future research*. In C. Mitcham (Eds.), *Encyclopedia of science, technology and ethics* (vol. 1) (xxvii-xxxv). Thompson Gale, Farmington Hills.

Jolliffe, I. (2002) *Principal component analysis* (2nd ed.). Springer, New York.

Kierkegaard, S. (2005) How the cookies (almost) crumbled: privacy and lobbyism. *Computer Law & Security Report*, 21, 310-322.

Kimppa, K. (2005a) Intellectual property rights in software – Justifiable from a liberalist position? Free software foundation's position in comparison to John Locke's concept of property, in *Intellectual property rights in a networked world: theory and practice* (eds. R. Spinello & H. Tavani), Information Science Publishing, Hershey.

Kimppa, K. (2005b) Intellectual property rights – or rights to the immaterial – in digitally distributable media gone all wrong, in *Information Ethics: privacy and intellectual property* (eds. L. Freeman & G. Peace). Information Science Public, Herhsey/London.

King, W. and King, D. (2000) Hacking 101: does it appear we are training the cyberspace criminals of the future? *International Information & Library Review*, 32, 463-472.

Kierkegaard, S. (2005) How the cookies (almost) crumbled: privacy and lobbyism. *Computer Law & Security Report*, 21, 310-322.

Kizza, J. (2003) *Ethical and social issues in the information age* (2nd Ed.). Springer-Verlag, New York.

Kramer, M. (1997) *John Locke and the origins of private property: philosophical explorations of individualism, community, and equality*. Cambridge University Press, New York.

Lampert, J. (1997) Locke, Fichte, and Hegel on the right to property, in *Hegel and the tradition: essays in honour of H.S. Harris* Toronto (eds. M. Baur & J. Russon), University of Toronto Press, Toronto.

Locke, J. (1988) *Two Treatises of Government* (ed. P. Laslett). Cambridge University Press, Cambridge.

Mason, R. (1986) Four ethical issues of the information age. *MIS Quarterly*, 10, 1, 5-11.

Mitcham, C. (2004) *Philosophy of information technology*, in *Philosophy of computing and information* (ed. L. Floridi), Blackwell Publishing, Malden/Oxford.

Moor, J. (1985) What is computer ethics? *Metaphilosophy*, 16, 4, 266-275.

Moore, A. (1997) A Lockean theory of intellectual property. *Hamline Law Review*, 21, 65-108.

Moore, A. (2008) Personality-based, rule-utilitarian, and Lockean justifications of intellectual property, in *The handbook of information and computer ethics* (eds. K. Himma and H. Tavani), Wiley & Sons, New Jersey.

Mortensen, M. (2005) Would be pirates: webcasters, intellectual property and ethics, in *Intellectual property rights in a networked world: theory and practice* (eds. R. Spinello & H. Tavani), Information Science Publishing, Hershey.

Nissenbaum, H. (1998) Protecting privacy in an information age: the problem of privacy in public. *Law and Philosophy*, 17, 559-596.

Oksanen, M. (1997) The Lockean provisos and the privatisation of nature, in *Justice, property and the environment: social and legal perspectives* (eds. T. Hayward & J. O'Neill). Ashgate, Brookfield.

Raymond, E. (1996) *The new hacker's dictionary* (3rd ed.). The MIT Press, Cambridge, Massachusetts.

Scanlan, M., (2005) Locke and intellectual property rights, in *Intellectual property rights in a networked world: theory and practice* (eds. R. Spinello & H. Tavani), Information Science Publishing, Hershey.

Scheffler, S. (1982) *The rejection of consequentialism*. Clarendon Press, Oxford.

- Searle, J.** (1983) *Intentionality*. Cambridge University Press, Cambridge.
- Simmons, A.** (1992) *The Lockean theory of rights*. Princeton University Press, New Jersey.
- Spafford, E.** (1992) Are computer hacker break-ins ethical? *Journal of Systems and Software*, 17, 41-47.
- Spinello, R. and Tavani, H.** (2005) Intellectual property rights: from theory to practical implementation, in *Intellectual property rights in a networked world: theory and practice* (eds. R. Spinello & H. Tavani), Information Science Publishing, Hershey.
- Stamatellos, G.** (2007) *Computer ethics: a global perspective*. Jones & Bartlett Publishers, Sudbury.
- Tavani, H.** (1999) Privacy On-Line. *Computers and Society*, 29, 4, 11-19.
- Tavani, H.** (2008) Informational privacy: concepts, theories and controversies, in *The handbook of information and computer ethics* (eds. K. Himma and H. Tavani), Wiley & Sons, New Jersey.
- Van Den Hoven, J.** (2008) Moral methodology and information technology, in *The handbook of information and computer ethics* (eds. K. Himma and H. Tavani), Wiley & Sons, New Jersey.
- Zack, N.** (1992) Locke's identity meaning of ownership. *Locke Newsletter*, 23, 105-113.

Computer-mediated Friendship: Illustrating Three Tasks for a Computer Ethics of the Good

Adam Briggie*
University of Twente

Abstract

I use computer mediated friendship to illustrate three tasks for a computer ethics of the good. First, it must combine the strengths of philosophical and empirical research. This is necessary to avoid conclusions based on values assumptions or uninformed analyses. Second, it must make some effort at generalization. This is necessary because an ethics of the good is holistic. Third, it must openly confront incommensurability. This is necessary for at least clarifying if not transcending the “it depends” conclusion.

Keywords: Good life; Friendship; Computer-mediated communication; New media; Online relationships

Introduction

Life is increasingly mediated through electronic networks, keypads, cameras, and screens. In what ways is this development improving or diminishing the quality of life? In order to address this question, Philip Brey (2007) has proposed broadening the agenda of computer ethics to include the implications of new media for the good life. In this paper, I use computer-mediated friendship to illustrate three tasks that such a computer ethics of the good must accomplish if it is to advance critical insight. First, it must combine the strengths of philosophical and empirical research. This is necessary to avoid conclusions based on values assumptions or uninformed analyses. Second, it must make some effort at generalization. This is necessary because an ethics of the good is holistic: judging whether this or that practice is good requires an understanding of how it fits into and reshapes the entire pattern of society and rhythm of life. Third, it must openly confront incom-

* *Adam Briggie* is an Assistant Professor in the Department of Philosophy and Religion Studies at the University of North Texas. His research interests lie at the intersection of science, technology, ethics, and policy. He is author of the forthcoming book *The Kass Bioethics Council: Public Philosophy and the Politics of Biotechnology* (University of Notre Dame Press).

mensurability. This is necessary for at least clarifying if not transcending the “it depends” conclusion.

Computer ethics has been subject to several interpretations (see e.g. Johnson 2000). And it has undergone previous expansions to include not just systematic study of the impacts of computers on society, but also the impacts of society on computers (e.g., Nissenbaum 1998). Yet, with some exceptions (e.g. Bynum 2006; Johnstone 2007), the field largely adopts an “ethics of the right” focused on right and wrong conduct in the design and use of information technology. Though important, this focus is limited. It does not conceptualize or evaluate how new media are qualitatively altering the way people live. Perhaps the most significant implications of new media in terms of quality of life stem from the increasing mediation of friendship and other personal relationships (Briggle 2008a; Briggle 2008b; Briggle and Spence forthcoming).

Computer ethics is primarily an applied ethical field. Thus, the emphasis in this paper will be on the practical topic of computer mediated friendship. However, I structure my analysis in such a way as to demonstrate how progress in an applied computer ethics of the good depends on these three meta-philosophical tasks. Thus, the paper aims to contribute to the development of a new approach within computer ethics both by examining a specific topic and by demonstrating how such applied analyses require certain meta-philosophical approaches.

Computer Ethics of the Good and Friendship

Computer ethics has traditionally focused on rights, risks, and justice. Classic examples are privacy rights, the risks posed by sexual predators online, and the injustices associated with the digital divide. But this is not the entirety of either philosophical or computer ethics. It is, rather, a “rump morality,” or morality understood as principles for what we owe to each other. This “ethics of the right” does not fully encompass “morality in the wider sense,” which asks questions about what kinds of lives are good or bad for people to lead and what kinds of societies foster human flourishing. This wider sense could be called an “ethics of the good,” where “good” is our most general term of positive evaluation. Theories of the good, then, specify what sorts of things in life are good and therefore worth striving for.

Deontologists and utilitarians argue about whether the right should be prior to the good in moral reasoning or vice versa. A computer ethics of the good, as a general project or specific assessment, need not assert its priority to the traditional emphasis on the right. It is, rather, a complementary focus—one that will allow us to have a “big tent” computer ethics that may parallel related developments in bioethics (Asch 2005). Bioethics has made a similar expansion beyond

the realm of obligation (e.g., respecting autonomy) to include consideration of flourishing. Broadly, this entails assessing biomedical technologies in terms of their implications for the meaning of being human. More specifically, it involves assessments of happiness, procreation and the relationship between generations, the significance of embodiment, the meaning of excellent performances, and the processes of aging and dying all in the context of technological developments.

Similarly, there are several issues raised by information and communications technologies that do not fit squarely within an ethics of the right. For example, Albert Borgmann (1999) articulated the goodness of reality and the impoverished nature of virtuality—a position directly opposed to that of Philip Zhai (1998). Hubert Dreyfus (2001) critiqued the quality of computer-mediated or distance education. And Sherry Turkle (2005) celebrated the multiplication of the self made possible through cyberspace and computer simulation. In this way, an inchoate computer ethics of the good already exists. Thus, Brey's proposal to broaden the agenda of computer ethics is a call to build from and formalize such pioneering works.

One promising territory for this expansion is the computer mediation of friendship and other personal relationships. New media have long been designed and utilized for interpersonal communication. Internet designers and users, for example, had an early focus on research, but this was quickly surpassed in the 1970s and 1980s by a new emphasis on communication. As the networking capabilities of computers increased, many countries and, later, private Internet Service Providers started their own networks. By the mid-1990s, dial-up connections were bringing the World Wide Web into an increasing number of homes and businesses.

Online relationships—from communities of shared interests to intimacy between two people—became more prevalent. Malcolm Parks and Lynne Roberts (1998), for example, found that nearly 25% of their respondents using MOOs (multi-user, text-based environments) formed online romantic relationships. Parks and Kory Floyd (1996) found that roughly two-thirds of newsgroup users developed personal relationships. The later developments of Web 2.0 and 3.D signify in part the growing use of the Internet for personal relationships and the enhanced affordances offered by such content generation and delivery systems as Second Life and Facebook. It is now simply common experience for most people in the developed world to have some degree of computer mediation in their personal relationships. This is especially true for youth (Ito et al. 2008).

Friendship is an appropriate topic for an ethics of the good, because it has long been understood as something that is good for humans. Aristotle argued that "friendship is a thing most necessary to life, since without friends no one would

choose to live, though possessed of all other advantages” (Nic. VIII, 1155a). Of course, some will feel drawn toward friendship more than others, and friendship exists within a plurality of goods that demands limits and tradeoffs. But in general, friendship, like wisdom, is a human good worthy of striving for. Being social creatures does not just mean that we are in need of others to sustain a functional community. Ants, elephants, and other social animals share this need. It also signifies our need to find meaning in life and the special role that personal ties with particular individuals play in fulfilling this need (Frankfurt 2004). Most people would find life impoverished if they lost their friends. For most people, such personal relations are ultimately what make life worthwhile. Friends not only support us, they also help us along the way toward meaningfulness, self-awareness, and self-improvement (LaFollette 1996).

This is not to deny that an ethics of the right has some place in discussing friendship. Indeed, our privileged knowledge about our friends and our position to act for their well-being creates duties to help them even when it would be easier or more pleasant not to speak-up. But to speak of friendship solely in terms of rights and correlative duties is to distort the phenomenon under consideration. Imagine that Smith visits you in the hospital and you enjoy a pleasant conversation just when you are in need of good company (Stocker 1976). You tell Smith that you appreciate his friendship, but he demurs—he confesses that he is merely doing his duty. He is not visiting you because he wants to or because he likes you, but because it is his duty to “do the right thing.” He sees you as a bearer of a rights claim to companionship that outweighs his desire to stay at home watching television. His visit suddenly seems cold and calculating, and it loses value for you, because what matters is not the act but the moral motivation behind the act or the spirit in which it is done.

Thus, conceptualizing friendship solely in terms of right action from some impartial moral standpoint is inadequate. Friendship is about unique personal bonds and cannot be understood absent the desires motivating those bonds and the values they serve. In short, friendship cannot be understood without an account of how it is experienced as good or as contributing to a flourishing life.

Such accounts are given by theories of the good life, and different theories will characterize friendship in different ways. I have so far implicitly appealed to a more “objective” theory of the good. According to such accounts, friendship will be considered as something that is good for the kinds of creatures that we are—it constitutes one part of a fully flourishing human life. There are also more “subjective” theories of the good that would value friendship in terms of its ability to satisfy the desires of the friends involved. Though friendship is often enjoyable and would be empty without some pleasure, it strikes me that a desire-satis-

factionist theory of the good will have trouble explaining why friendships often endure through times of hardship and struggle. It may be possible, however, to argue that friendships satisfy a desire for meaningfulness, which is a desire that by definition is more enduring than the fickleness that the term “desire” often conjures to mind.

If friendship is central to philosophical accounts of the good life, then it stands to reason that the computer mediation of friendship is a suitable topic for a computer ethics of the good. In what follows, I explore this topic in such a way that allows me to illustrate three tasks that any project within computer ethics of the good must accomplish if it is to provide the traditional benefits afforded by normative assessment. I take those benefits to be the articulation and evaluation of moral intuitions and the justification of moral judgments—or, in short, the advancement of critical insight into an issue of social and moral significance.

First Task: Explicitly Normative and Reality Based

A computer ethics of the good must combine philosophical evaluation with empirical research. If the former is missing, then it is not explicitly normative. If the latter is missing, then it is not based in reality. A look at the existent literature on computer-mediated friendship suggests room for improvement with regard to this task.

Explicit normativity means specifying what the good at stake is, justifying why it is good, and demonstrating how a particular technology or slice of material culture more broadly promotes or diminishes that good. Borgmann, for example, must be clear about what reality is, why real experiences are better than virtual ones, and how technology undermines reality. Dreyfus must be clear about what a good education consists in and how computer-mediation threatens it. And Turkle must specify why multiplying one’s sense of self is a positive development and how computers contribute to it. Such existent projects within a computer ethics of the good are at times insufficiently explicit about their norms.

To conduct explicitly normative inquiry into computer-mediated friendships, one must specify and relate (a) the nature and value of friendship; (b) the nature of mediating human relationships; and (c) the value implications of mediation in terms of the quality of those relationships (see Briggles 2008a). A conceptual framework for guiding such inquiry could take the following form: Friendship means A / thus the differences, B, introduced by mediation / entail values implications C.

The argument put forward by Dean Cocking and Steve Matthews (2000) in one of the rare philosophical assessments of online friendship fulfills each portion of

the framework and thus provides a good example of explicit normativity. They argue that:

Close friendship requires non-voluntary cues (A), which are filtered out in text-based Internet communication (B), thereby making wholly mediated close friendship impossible (C).

Note that the technologies involved will determine much of the value differences between mediated and non-mediated relating. For example, Cocking and Matthews limit their analysis to text-based communication, because web-cams convey non-voluntary cues.

Furthermore, being explicit about the implications of computer-mediation for friendship requires openly defending some account of the technological shaping of human activity (from determinism to neutrality). Some use the concept of a technological "script," which prescribes behavior to some extent, but also allows actors to adapt diverse interpretations and styles (Bijker and Law, 1992). Others use the concept of "value suitabilities." A given technology "is more suitable for certain activities and more readily supports certain values while rendering other activities and values more difficult to realize" (Friedman et al., 2006, p. 351). Cocking and Matthews argue that in this context of text-based communication, user-motivations and behaviors ultimately do not matter, because they cannot overcome the deterministic "structural barrier" of the medium that necessarily filters out non-voluntary cues.

This particular position can be criticized. But the important point here is that some account of the flexibility or rigidity of technology is essential for making a normative argument and thus should be explicitly defended. It makes the difference between, for example, a claim that the Internet cannot sustain close friendships (and thus should be avoided altogether) and a claim that the Internet is not likely to sustain close friendships (and thus should be used in specific ways).

Cocking and Matthews excel in terms of explicit normativity. Other studies do not fare so well. Many adopt evaluative standards based on an implicit conception of friendship without explaining or justifying the choice of those standards or that conception. For example, D. Chan and G. Cheng (2004) adopt a list of seven dimensions for evaluating the quality of online friendships. Yet the list is not grounded in any theoretical account of friendship and its value. One dimension is "interdependence," but they do not make it clear when or why more interdependence is good or bad (for other examples of this failure see Cummings et al. 2002; Cheng et al. 2006). If one is not clear about what friendship is, then there is no way to defend an evaluative claim about the implications of computer-mediation for friendship.

But these empirical and survey studies are naturally strong in the other part of this task, namely, they are reality-based. This is because they are concerned to discover actual user motivations and practices. This is where Cocking and Matthews do not fare so well. Their analysis pertains only to wholly mediated and text-based interactions. Yet, in most actual practices offline and online interactions mix and mediated interactions increasingly feature pictures, voice, and cameras in addition to text.

Restricting one's analysis to wholly-mediated, text-based friendships is a way to obtain conceptual clarity. But clarity is gained at the expense of reduced relevance to the real-world as most friendships are not wholly mediated and wholly text-based. The challenge, then, for a philosophical computer ethics of the good is to marry explicit normativity with more nuanced accounts of actual user practices in all their complexity. Cocking and Matthews could begin this task by arguing that mixed (offline-online) friendships or wholly mediated friendships that involve more than text are not impoverished, because they avoid the filtration of non-voluntary cues. But this would only be a beginning, because it identifies just one prohibitive aspect of mediating technology. There are many other relevant technological affordances and constraints that factor into the mediation of friendship (see Briggles 2008b).

Second Task: Save the Forest from the Trees

One response to the complexities introduced by the first task is to hone in on specific practices or platforms. This is the "empirical turn" strategy in the philosophy of technology, which emphasizes specific contexts and user behaviors. It is a welcome complement to the "classic" philosophy of technology, which tended to abstract from particulars in order to criticize or praise "modern technology" writ large.

The danger in focusing on specific cases (trees), however, is that we will lose sight of the larger pattern of human life in a new media age (the forest). Determining whether some particular practice or device is good depends on an understanding of how it fits into and partially shapes society, the whole of a person's biography, and the complex ordering of his or her needs and desires. An analysis focused solely on specific devices and practices is blind to the emergent properties that result from their combination.

Cautious generalization is thus called for. In another work, Borgmann (1984) provides an example of this practice and its value. He begins with case studies. For example, he considers the replacement of wood-burning heating stoves with furnaces controlled by a thermostat or the shift from family meals at the table to individual meals at the television. He then extracts from such examples a gen-

eral style of life that characterizes the contemporary world. He calls this inconspicuous pattern the “device paradigm.” This is a way of taking up with the world aimed at making goods (or “commodities”) technologically available without the imposition of burdens: “Something is available in this sense if it has been rendered instantaneous, ubiquitous, safe, and easy” (41). The central heating system makes warmth available in a way it was not prior to modern technology. Similarly, frozen food makes a meal available, a faucet makes water available, and a stereo makes music available. The machinery makes no demands on skill, strength, or attention. “Progress” is movement toward machinery that is less demanding and that is thus more hidden from the experience of consuming commodities.

At the level of the forest, then, Borgmann argues that there is one predominant lifestyle, namely the “commodious” or “disburdened” way of life. His generalization casts light on this underlying pattern of life. Analyses that focus just on the case-study level are unable to capture and evaluate such emergent properties of contemporary life. Yet these emergent phenomena are central to any technological ethics of the good, because they speak to the whole of one’s biography—a life as lived in the context of history and culture and as shaped by both short-term and long-term goals. Borgmann’s analysis allows us to understand contemporary existence as commodious, thereby making it possible for us to ask in what ways or to what extent a commodious life is a good life and to what extent we are free to choose another style of life. In this way, generalization establishes an encompassing perspective from which to assess the design and use of any specific technology.

The study of computer-mediated friendship should similarly strive for different types and levels of generalization. Case-studies of, for example, particular social networking sites are necessary for this task, but they are by definition insufficient for identifying general and emergent properties of life in a world where friendships are increasingly mediated.

One promising approach to such generalization would be to follow Borgmann in an attempt to identify a “grain” or core bundle of properties characterizing social media. In my current research, I am attempting to develop this approach in the context of youth culture. The idea is that speed, ubiquity, brevity, and multi-tasking form the grain of new social media in youth culture. This speaks to an underlying reality: social media have created new channels for interaction (new cyberspaces) without, of course, creating more time. Thus, relating has become cramped (LOL, OMG), accelerated, and dispersed into a diffusion of decentered activity. Of course, one can work against the grain just as Borgmann claims that one can work against the device paradigm or the rule of technology. But this

would be work and to initiate it would require some awareness of what may be lost by simply going with the grain.

I can flesh this point out a bit more. New media are shifting focus toward the production and consumption of multiple streams of micro-reports and away from the slow processing of an unfolding relational narrative. With the increased quantity of interaction has come a two-fold change in its type or quality. First, the length of an average correspondence is shrinking. There is likely bidirectional causality at work—the imperative to connect more frequently reduces the length of each connection and the ability to send multiple short messages fuels the imperative to interact more frequently. This trend is apparent in the rise of microblogging, status updates and comments on Facebook, and Twitter feeds. Many teens prefer the brief give and take of texts to the commitment and focus of attention demanded by talking either face-to-face or on the phone (Olson 2008). E-mail is out of fashion, let alone a lengthy handwritten letter.

Second, there is an increase in multi-tasking. Over a quarter of the time that young people are using media, they are using more than one medium at a time (Rideout et al. 2005). Texting, for example, unlike talking, allows one to simultaneously allocate portions of one's attention to other tasks, including other forms of interacting. The nature of "windows" in a screen culture encourages multi-tasking by allowing several applications to appear at the same time. What multi-tasking displaces is "focus," a Latin word that means "hearth," that center of a household that gathers the rhythm of family life around it. In a culture of overlapping screens and windows, there is often no focus where lines converge to produce a well-defined center. Interaction with a friend is another distraction in a world of distractions. Dalton Conley (2009) has discussed this continuous partial attention in terms of being perpetually "elsewhere." New media allow work to penetrate home life, vacation, and even commuting. Youth culture is in this sense preparing teens well for an adult life of interpenetrating spheres that constantly fragment and shunt one's attention elsewhere.

The general pattern of youth culture in a social media age, then, is one of distraction and fragmentation. So, the question becomes: how might this influence the quality of friendships and the quality of life more generally? Addressing this question requires guidance from the conceptual framework above. For example, if friendship requires time and focus for reflection (A) and mediation is in general dispersing focus and fragmenting attention (B), then social media may ironically undermine the goal of friendship by flooding our lives with too much interaction (C). The value of this kind of analysis is that it allows one to evaluate the quality of life in general under the influence of social media considered not as

isolated devices but as contributing to a newly emergent culture or pattern of life with distinctive rhythms, practices, and values.

Third Task: Confront Relativism

To illustrate the third task, I will continue with the example of youth friendship. Teens of the new media age—digital natives—are constantly tuned into their peers through social media. With their ubiquity and speed, new media technologies have given rise to what one major survey calls “always-on intimate communities” (Ito 2008). For example, one research group estimates that the average U.S. teenager generates between fifty and seventy text messages per day (see Olson 2008). To an unprecedented degree, teens are wired into the minutiae of one another’s lives.

Is this good or bad for those teens—for the quality of their relationships and lives? As my conceptual framework suggests, there are many concepts of friendship and each one provides its own normative standards for answering the question about the value of hyper-friendship or always-on intimacy (see Pakaluk 1991; Badhwar 1993). There is a strong temptation, then, to conclude that it depends on one’s understanding of friendship and its relation to well being. We seem stuck in a fundamental incommensurability reminiscent of the regress of justification in philosophical ethics—each theory presupposes basic principles that cannot themselves be justified.

This may indeed be an impasse, but this would only be so after the philosophic work has been done to translate muddled assumptions about friendship into clear conceptions. If differences remain after this work is done, then at least the debate will have been refined to give a clearer sense of the values dimensions at stake. For example, the rise of “always-on” friendships is likely to illicit contrasting intuitions—some seeing it as a renaissance of relationships and others seeing it as somehow detrimental to true friendship. When intuitions clash about the value of new technologically-mediated practices, a computer ethics of the good can help elevate the debate and set it into context.

In this case, the debate about always-on friendships may be fruitfully cast in terms of a deep division between the conceptions of friendship put forward by Plato and Aristotle. To paint the contrast starkly, where Plato saw friendship as a process of mutually shaping one another on the path toward virtue, Aristotle saw it as the accomplishment of those who are already virtuous. In the former conception, character is something that is formed along the way within the friendship. In the latter conception, character is a necessary condition for the formation of friendships.

Of course, this paints their differences too simplistically. Plato did not deny the importance of good character for initiating good friendships. And Aristotle focused on the development of character through habituation. But even such a cursory analysis already helps clarify the debate. Those optimistic about always-on friendships may be drawing from an inchoate Platonic conception. Their enthusiasm about social media may stem from their notion that more interaction means more opportunities for friends to mutually shape and improve one another. Those skeptical about hyper-relationships may be drawing from an Aristotelian perspective. Their doubts stem from a notion that socializing at some point crosses a threshold where it becomes a constant distracting task of managing and updating information. Friendship requires more than the increased opportunities afforded by social media for influencing one another. At least as important are the time and resources of character to interpret and steer those interactions in appropriate directions.

So, is always-on friendship a good or bad thing? It may be that “it depends” is the final answer. But through a computer ethics of the good, we can at least clarify what it all depends on. In this case, it depends on whether one sees more interaction with friends as necessarily and intrinsically positive. Alternatively, one may think of interaction as governed by a mean where at some point more becomes counterproductive either by displacing other goods or by diminishing the quality of friendship itself. As technology shapes the moral landscape, it is likely that positive assessments will come to prevail—humans have a way of getting used to and embracing their own inventions to the point that they become simply “natural.” But the ultimate task of a computer ethics of the good is not just to highlight this moral change, but to evaluate it. In this case, then, it must make clear what we stand to lose in unthinkingly embracing a world of constant interaction.

REFERENCES

- Asch, A.** 2005. “Big Tent Bioethics: Toward an Inclusive and Reasonable Bioethics.” *Hastings Center Report* 35, no. 6: 11-12.
- Badhwar, N.K.** (1993). *Friendship: A Philosophical Reader*. Ithaca, NY: Cornell University Press.
- Bijker, W. and Law J.,** eds. (1992). *Shaping Technology / Building Society*, MIT Press, Cambridge, MA.
- Brey, P.** (2007). “Theorizing the Cultural Quality of New Media.” *Techné. Research in Philosophy and Technology* 11 (1), 1-18.
- Briggle, A. and Spence E.** (forthcoming). “Cosmopolitan Friendship Online,” in *Friends and Foes* (Oxford: Cambridge Scholars Publishing).

Briggle, A. (2008a). "Love on the Internet: A Framework for Understanding Eros Online," *Journal of Information, Communication, and Ethics in Society*, vol. 6, no. 3, pp. 216-232.

Briggle, A. (2008b). "Real Friends: How the Internet Can Foster Friendship," *Ethics and Information Technology*, vol. 10, no. 1, pp. 71-79.

Borgmann, A. (1984). *Technology and the Character of Contemporary Life*. Chicago: Chicago University Press.

Borgmann, A. (1999). *Holding On to Reality: The Nature of Information at the Turn of the Millennium*, University Of Chicago Press, Chicago, IL.

Bynum, T. (2006), "Flourishing Ethics," *Ethics and Information Technology*, 8(4): 157-173.

Chan, D.K.S., Cheng, G.H.L. (2004). "A Comparison of Offline and Online Friendship Qualities at Different Stages of Relationship Development," *Journal of Social and Personal Relationships*, vol. 21, no. 3, pp. 305-320.

Cheng, G.H.L., Chan, D.K.S., Tong, P.Y. (2006). "Qualities of Online Friendships with Different Gender Compositions and Durations," In: *Cyberpsychology and Behavior*, vol. 9, no. 3, 14-21

Cocking, D. and Matthews S. (2000). "Unreal Friends." *Ethics and Information Technology*.

Conley, D. (2009). *Elsewhere, U.S.A.: How We Got from the Company Man, Family Dinners, and the Affluent Society to the Home Office, BlackBerry Moms, and Economic Anxiety*. New York: Pantheon.

Cummings, J.N., Butler B. and Kraut R. (2002), *The Quality of Online Social Relationships*, *Communications of the ACM*, Vol.45, No.7, 103-108

Dreyfus, H. (2001). *On the Internet*, Routledge, London.

Frankfurt, H. (2004). *The Reasons of Love*, Princeton University Press, Princeton, NJ.

Friedman, B., Kahn, P. Jr., and Borning, A. (2006). "Value sensitive design and information systems," Zhang, P. and Galletta, D. (Eds.), *Human-Computer Interaction and Management Information Systems: Foundations*, M.E. Sharpe, London, pp. 348-372.

Ito, M., et al. (2008). *Hanging Out, Messing Around, Geeking Out: Living and Learning with New Media*. Cambridge: MIT Press.

Johnson, D. (2000). *Computer Ethics*, 3rd ed. Upper Sadle River: Prentice Hall.

Johnstone, J. (2007). "Technology as empowerment: a capability approach to computer ethics." *Ethics and Information Technology* (2007) 9:73–87.

LaFollette, H. (1996). *Personal Relationships: Love, Identity, and Morality*, Blackwell, New York, NY.

Nissenbaum, H. (1998). "Values in the Design of Computer Systems," *Computers in Society*, 38-39.

Olson, S. (2008). "Cell Phone is Mom-avoidance Device for Teens," CNET News, July 14, available at: http://news.cnet.com/8301-1023_3-9991199-93.html?tag=nefd.top.

Pakaluk, M. (1991). *Other Selves: Philosophers on Friendship*. Indianapolis, IN: Hackett.

Parks, M.R. and Floyd, K. (1996). "Making friends in cyberspace," *Journal of Communication*, vol. 46, pp. 80–97.

Parks, M.R. and Roberts, L.D. (1998). "Making MOOsic: the development of personal relationships online and a comparison to their off-line counterparts," *Journal of Social and Personal Relationships*, vol. 15, pp. 517–537.

Rideout, V., Roberts D.F. and Foehr U.G. (2005). *Generation M: Media in the Lives of 8-18 Year Olds*. Menlo Park, CA: Kaiser Family Foundation, available at: <http://www.kff.org/entmedia/upload/Executive-Summary-Generation-M-Media-in-the-Lives-of-8-18-Year-olds.pdf>.

Stocker, M. (1976). "The Schizophrenia of Modern Ethical Theories," *Journal of Philosophy*, vol. 73, pp. 453-466.

Turkle, S. (2005). *The Second Self: Computers and the Human Spirit*. Cambridge, MA: MIT Press.

Zhai, P. (1998). *Get Real: A Philosophic Adventure in Virtual Reality*. Lanham, MD: Rowman and Littlefield.

Internet Research Ethics: Reports from US-Based Institutional Review Boards

Elizabeth A. Buchanan*
University of Wisconsin-Milwaukee

&

Charles M. Ess**
Drury University/Aarhus University

Abstract

The Internet has been used as a place for and site of an array of research activities. From online ethnographies to public data sets and online surveys, researchers and research regulators have struggled with an array of ethical issues around the conduct of online research. This paper presents a discussion and findings from Buchanan and Ess's study on US-based institutional review boards and the state of internet research ethics.

Keywords: Internet research ethics, human subjects, institutional review boards, online research, research ethics

Introduction

With the emergence of Internet use as a research locale and tool throughout the 1990s, researchers from disparate disciplines, ranging from the social sciences to humanities to the sciences, have found a new fertile ground for research op-

* *Elizabeth Buchanan*, Ph. D. is Director, Center for Information Policy Research, and Associate Professor, School of Information Studies, University of Wisconsin-Milwaukee. She is Co-Director of the International Society for Ethics and Information Technology; Chair, Ethics Working Group, Association of Internet Researchers, and Past Chair, Wisconsin Library Association Intellectual Freedom Round Table.

** *Pr. Charles Ess*, Professor, Department of Information and Media Studies, Aarhus University, Denmark. «Since 2003 Charles Ess has been a 'Distinguished Professor' at Drury University, US, and he is a leading and internationally recognized scholar within a variety of fields. He has made significant contributions within such fields as Internet research, with emphases on culturally-variable elements of Information and Communication Technology and Computer-Mediated Communication, ICTs and democratization, ICTs and religion, Internet research ethics as well as Information and Computing Ethics, with an emphasis on cultural perspectives».

portunities that differ greatly from their traditional biomedical counterparts. Populations, locales, and spaces that had no corresponding physical environment became a focal point, or site of research activity. Human subjects protections questions then began to arise, across disciplines and over time: How do researchers best protect their subjects' privacy - especially given how privacy online is oftentimes much more difficult to ensure than offline? How is informed consent obtained and sustained - especially in the light of a «click-through» habit that may incline subjects not to read the consent form carefully? What best practices for research with/on minors exist - e.g., as ensuring the real-world identity of the minor is much more difficult online, and as online anonymity seems to foster a greater «confessional» attitude that may reveal disturbing, potentially life-threatening information (Stern, 2004; Bober, 2004)? What are «harms» in an online environment? Is this really human subjects work (White, 2003)? What about the ethics of knowledge-reuse databases, and reusing data and field notes in online environments? How does the fluidity of hyperblogging as a research site confound the static notion of informed consent as it currently is enacted in research ethics reviews? And, ultimately, what are the ethical obligations of researchers conducting research online and are they somehow different from and/or the same as other forms of research ethics practices? These questions speak directly to ethics and values studies, examining normative issues in the conduct of research. As such, researchers across disciplines, and across borders are implicated in Internet research ethics (IRE).

Case studies are often a great way to understand the complexities of an issue, especially as they provide a way of teasing out the myriad concerns facing ethics boards. Consider these three examples:

1. A researcher plans to monitor blogs (selected according to popularity, with a statistical sample across a broad range) to see if the more popular blogs exhibit more racially-biased entries and responses. The researcher claims that this is not “human subjects research” because the blogs are public forums, much like newspapers, only privately published.
2. For my research, I'm studying the online presence of Israeli parliamentarians in social networks. I have a fictitious profile in FB (of course under the guidance of my supervisor) (which is base on a paraphrase of my own name and my own pic [sic] as a child and not a stolen identity!). I have invited Israeli politicians to be my friends so I can get updates about their use of the network. I do not write to them nor interact with them in any way apart this preliminary action. All I do is use my online presence to see what they are doing. All their postings and actions on the network are public and known to all. Are there any special ethical guidelines I should follow? I have to say I have detailed in my current occupation

that I'm a researcher who deals with the connections between internet and politics.

3. We had a researcher using the website "Gay Bombay" to study gay Indian men's attitudes, and the board was worried that since homosexuality is illegal in India, would participation get the respondents in trouble somehow?

With these as exemplars of the types of research ethics issues facing IRBs in their reviews of online research, we can briefly contextualize the history of IRE. Disparate disciplines throughout the 1990s began in piecemeal fashion looking at these ethical complexities and implications of conducting research online. Whether or not such research ethics guidelines as The Belmont Report and such federal human subjects protections as codified in the US Code of Federal Regulations "fit" or were applicable were at best uncertain. The debate began to take serious academic form when one of the first journal issues devoted entirely to Internet Research Ethics (IRE) appeared in 1996, in a special issue of *The Information Society*, followed by a workshop on IRE in 1999 funded by the National Science Foundation and the American Association for the Advancement of Science, overseen by Frankel and Siang. Further evidence of the recognition and development of IRE came through the release of the Association of Internet Researchers (AoIR) Ethics Working Group's report on Ethical Decision Making and Internet Research, chaired by Charles Ess, in 2002. Such consideration occurred among researchers, policy makers, and institutional review boards, which were seeing an increase in the number of Internet-based research protocols (Buchanan, 2003, 2004). Too, such prominent professional societies as the American Psychological Association convened a Board of Scientific Affairs Advisory Group on Conducting Research on the Internet, releasing a report in 2004 in *American Psychologist* (Kraut, et al). And finally, three books in the field of IRE were published between 2003 and 2004 (Buchanan; Johns, Chen, & Hall; Thorseth). The *International Journal of Internet Research Ethics* and the *International Journal of Internet Science* were both founded in 2006. These were all, indeed, important moments in the development of IRE as a discrete research phenomenon, and promoted serious consideration about the ethical implications of research in online or virtual environments.

While scattered literature has appeared across disciplines, IRBs and other policy makers have struggled to understand the multiple issues – some familiar, some novel – involving human subjects that take place online. Moreover, professional ethics workshops conducted by Buchanan and Ess (2003, 2004, 2005, 2006, 2007) have demonstrated, however, that while the AoIR guidelines are indeed useful in educating IRBs and resolving common issues – many more ethical problems are emerging as Internet venues change, technologies advance, researchers seek out new avenues of investigation, but, by contrast, human subjects protec-

tions models remain static. Anecdotal evidence reveals that IRBs generally do not know what such protections apply strictly to online research, and such boards often ignore the complexities of such research and thereby risk harming subjects while also violating federal regulations, or, they apply such restrictive models that inhibit researchers from pursuing important online endeavors (Buchanan and Ess, 2003, 2004, 2005; Ess, 2005). At the same time, however, compliance in the area of human subjects protections is of paramount concern to academic institutions:

Just about any study conducted at a university medical center, a hospital, a contract research organization, or elsewhere must now pass the muster of an Institutional Review Board (IRB). IRB approval is necessary before the project is begun, in some cases before investigators can even apply for funding. This is as true of social science projects in anthropology, economics, epidemiology, political science, and sociology as it is of clinical or experimental research in medicine and psychology. The aggregate activity conducted under human subjects protections is staggering: every year thousands of IRBs in the US examine over 20 000 research proposals, and hundreds of thousands of experimental subjects and patients are presented with their legal human subjects' rights and sign consent forms stating that they understand these rights as they participate... (Carpenter, 2005).

Prior to our study, no one has documented the actual number of Internet-based research protocols across disciplines and reviewed by IRBs, with the exception of Peden and Flashinski (2004), who examined the field of psychology and showed a dramatic lack of adherence to ethical guidelines in online psychology experiments. This emergent area thus raises particular concern, due to the lack of knowledge of the new vocabularies of IRE, new technologies employed for conducting research, new disciplines using online research, particularly from those in the sciences who traditionally are not well equipped in human subjects protections models, and new possibilities for noncompliance where researchers and IRBs are not well equipped to review IRE protocols.

For this study, Internet research is defined as "research that (a) utilizes the Internet to collect information, e.g., through online surveys; (b) studies about how people use the Internet, e.g., through collecting and examining activities on list-serves, web sites, blogs, or other online environments; and/or (c) datasets available online."

Methodology

In 2006, we applied for and received funding through the National Science Foundation's Ethics and Values in Science Directorate. To understand the state of the US institutional review board in both qualitative and quantitative ways, we developed a database of over 700 US-based institutional review boards, based on the Carnegie Classification system. The database was built in 2007. We included institutions of mixed status, including doctoral research, master's level, special institutions, such as tribal colleges, medical schools, and business schools. After receiving institutional review board approval, as this was considered exempt human subjects work under the US Code of Federal Regulations category 2, a mixed-method survey was developed and mailed to the participants, with a return of 334, a response rate of 44%. Data were collected from September 2007-May 2008, with three follow-ups sent to non-respondents. Given the exploratory nature of the work, this response was satisfactory. We made a conscious decision not to employ online survey tools, as we were questioning their very use in the survey, and there was a certain 'meta-ness' to their use. Thus, data were collected, compiled, and entered by hand in SPSS for analytical purposes. Qualitative data were coded and analyzed separately.

Selected Findings

Full findings are still under development (Buchanan and Ess, 2010), and additional reports are available in Buchanan (2009) and Buchanan and Hvizdak (2008, 2009).

Based on our preliminary work, however, we can point out, to begin with, that of the 334 respondents, nearly half found Internet research an area of concern or importance. There was a commonality among boards that said IRE was not an issue of concern—many admitted that they knew it would only be a matter of time before they were faced with such research, and expressed encouragement for the development of specific IRE guidelines.

Given that 62% did not have guidelines or checklists in place for reviewing Internet research-based protocols, that few boards were aware of extant guidelines such as the Association of Internet Researchers Ethical Decision Making document, that 74% did not provide specific training around Internet research issues, and that less than half (42%) felt the Office Human Research Protections or other regulatory documents were useful in Internet research reviews, we must consider on what bases are review boards making decisions around IRE? Moreover, our qualitative data provided indications that many boards were "unsure of who to ask," "we don't even know what questions to ask of the researcher," and, "we rely on the IT department to advise us on such IT related issues."

We further found that the majority of research reviewed by IRBs came from the social sciences, followed by medicine/health, with a small percentage coming from arts and humanities. Overwhelmingly, Internet-research protocols fell under the US CFR “exempt” category for research (see 46 CFR 101b).

We found that online surveys (see also Buchanan and Hvizdak, 2008, 2009) are the most frequently reviewed methodology used in online research, yet many (38%) institutional review boards do not review the privacy policies from such tools, and even fewer have institutional tools of their own, built with an awareness to research ethics issues, or value-sensitive design. When one considers that the default for many online survey tools is mandatory responses (as in, one can not skip questions), the fundamental issue of voluntariness is compromised. Furthermore, tracking IP addresses, third party access, auto-fill ins, public Internet terminals, and ownership of the data contribute to a situation where research subject/participants can be easily identified - contra the fundamental promise in research ethics to protect the identity, confidentiality, and anonymity of the persons involved as subjects. A further complication here is that In traditional research settings, the researcher assumes responsibility for protecting the participants’ identities, but in online research, he or she may not be solely responsible. Finally, the risks increase when certain types of data are being collected, such as medical information and thus, participants deserve greater protections (Svengingsson, 2004). Thus, online surveys are a good example where value-sensitive design has a significant role to play. Watt (2009), for example, recommends an open source project to develop an Internet survey research tool that would remove the reliance on commercial tools with their inherent dangers and would also give academic researchers a full set of questionnaire design tools that would allow them to create ways of eliciting information from research participants that go beyond the very elementary current norm. There are ways to explicitly deal with issues of data security, identity, and permission that could be made nearly invisible to the researcher (or at least a whole lot easier than they are now).

In terms of training or continuing education or both researchers and IRB members, our results indicate that for general research ethics, over 60% of respondents either require or encourage training. However, in regards to Internet research ethics, only 9% either require or recommend training. For IRB members themselves, nearly 30% require or recommend training on Internet research ethics. This leaves ample room for improvement. In addition to the clear need currently for additional training based on what is now an established body of literature on IRE, such training becomes even more pressing in light of the ever-changing nature of Internet research: as technologies change, the ways in which research is conducted also change - and hence our research ethics must continually respond and adapt. This leads us to the concept of transformative research: we

hope that our work in IRE pushes transformative research to a newfound model of transformative research ethics.

The major areas of concern cited by IRBs include:

1. Privacy
2. Data security
3. Appropriate models and modes of eliciting and maintaining consent
4. Sensitivity of data
5. Appropriate models for ensuring confidentiality and anonymity.

Discussion and Conclusions

Our preliminary findings make clear that IRBs do need concrete guidelines as the foundation of protocol review. The CFR, and related documents, do serve as a starting point, but are insufficient to deal with the emergent and quickly evolving technologies of the Internets. The areas of concern presented throughout this research, such as consent, privacy, protection of data, have been long-standing issues in research ethics, but the “greased” aspect of data, coupled with the idea of the “networked self” in research endeavors, present great tensions for researchers and research boards.

We hope that the intellectual merits of the project have related to the successful approach of using applied ethics to address pressing, real-world issues in human subjects protections, and continuing to develop practical resolutions not only to first-order ethical issues, but also to second-order issues of interdisciplinary collaborations.

In particular, our project integrated research and education by first advancing our understanding of both the significant ethical issues raised in Internet research and how far IRBs, as the primary ethical gatekeepers of research in the United States, understand and seek to resolve these issues as they emerge in research proposals from across the disciplines. Moreover, our survey of current IRB policies and procedures allowed us to highlight potential deficits and thereby greatest needs regarding resources necessary for ethically informed and responsible reflection, advice, and decision.

The next steps for continuing IRE scholarship involves a collaborative advisory board, comprised of researchers, ethicists, and research administrators from around the globe. From this project, we gained a greater appreciation of the deep complexities of Internet research, from the methodological, to the ethical, to the legal, to the practical. Happily, a global community of IRE scholarship is emerg-

ing, and a collaborative model of online sharing of research protocols, case studies, guidelines, best practices, and legitimate failures will facilitate greater collaboration, and, we are confident, lead to still more extensive philosophical reflection and scholarship. Moreover, it will continue to lead us to a space of global research collaboration and perhaps even to a place of shared norms and values in research conjoined with a pluralistic respect for disciplinary and cultural differences.

REFERENCES

American Association for the Advancement of Science (1999). Ethical and Legal Aspects of Human Subjects Research in Cyberspace. Available: <http://www.aaas.org/spp/sfrr/projects/intres/main.htm>.

Association of Internet Researchers Ethics Working Group, and Ess, C. (2002). Ethical Decision-Making and Internet Research: Recommendations from the AoIR Ethics Working Committee. Available: www.aoir.org/reports/ethics.pdf.

Buchanan, E. (2003). Internet Research Ethics: A Review of Issues. Presentation at the Association of Internet Researchers Annual Conference. Toronto, CA.

– (Ed.). (2004). Readings in Virtual Research Ethics: Issues and Controversies. Hershey: Idea Group.

– (2005). The IRB Review and Online Research. Presentation at the Association of Internet Researchers Annual Conference. Chicago, IL.

Buchanan, E. & Ess, C. (2003). Understanding Internet Research Ethics, A Workshop. Association of Internet Researchers Annual Conference. Toronto, CA.

– (2005). Internet Research Ethics Workshop. Association of Internet Researchers Annual Conference. Chicago, IL.

– (2008). Internet Research Ethics: The Field and Its Critical Issues. In *The Handbook of Information and Computer Ethics*. Edited by Herman Tavani and Kenneth Einar Himma. Boston: Wiley.

Buchanan, E. & Hvizdak, E., (2008). (Survey) Monkey Business. Presented at Association of Internet Researchers Ninth Annual Conference, Copenhagen, Denmark, 15 October.

– (2009). IRBs and Online Surveys: Ethical and Practical Considerations. *Journal of Empirical Research on Human Research Ethics (JERHRE)*. In Press.

Carpenter, D. (2005). Gatekeeping and the FDA's Role in Human Subjects Protection. American Medical Association, *Medicine and Society*. Available <http://www.ama-assn.org/ama/pub/category/13245.html>

Bober, M. (2004). Virtual Youth Research: An Exploration of Methodologies and Ethical Dilemmas from a British Perspective (pp.288-316). In Buchanan, E. (Ed.). Readings in Virtual Research Ethics: Issues and Controversies. Hershey: Idea Group.

Kraut, R., Olson, J., Banaji, M., Bruckman, A., Cohen, J., & Cooper, M. (2004). Psychological research online: Report of board of scientific affairs' advisory group on the conduct of research on the internet. *American Psychologist*, 59(4), 1-13.

Johns, M., Chen, S. L., & Hall, J. (Eds.) (2003). *Online Social Research: Methods, issues, and Ethics*. New York: Peter Lang.

The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1979). The Belmont Report. Available <http://ohsr.od.nih.gov/guidelines/belmont.html>.

Peden, B. and Flashinski, D. P. (2004). Virtual Research Ethics: A Content Analysis of Surveys and Experiments Online (pp. 1-26). In E. Buchanan (ed.), *Readings In Virtual Research Ethics: Issues And Controversies* Hershey, PA: Idea Group.

Stern, S. (2004). Studying Adolescents Online: A Consideration of Ethical Issues (pp. 274-287). In Buchanan, E. (Ed.). *Readings in Virtual Research Ethics: Issues and Controversies*. Hershey: Idea Group.

Sveningsson, M. (2004). Ethics in Internet Ethnography (pp. 45-61). In Buchanan, E. (Ed.). *Readings in Virtual Research Ethics: Issues and Controversies*. Hershey: Idea Group.

Thorseth, M. (2003). *Applied Ethics in Internet Research*. Trondheim, Norway: Programme for Applied Ethics, Norwegian University of Science and Technology: Trondheim.

Watt, J. (2009). Personal Communication.

White, M. (2003). Representations or People? Available http://www.nyu.edu/projects/nissenbaum/ethics_whi_full.html

Free software, economic “realities” and information justice

Samir Chopra* & Scott Dexter**

Department of Computer and Information Science
Brooklyn College

Abstract

Free and open source software (FOSS) is taking an increasingly significant role in our software infrastructure. Yet many questions still exist about whether a software economy based on FOSS would be viable. We argue that contemporary trends definitively demonstrate this viability. Claiming that an economy must be evaluated as much by the ends it brings about as by its size or vigor, we draw on widely accepted notions of redistributive justice to show the ethical superiority of a software economy based on FOSS.

Keywords: free software, information justice, redistribution, market

Introduction

The free and open source software (FOSS) phenomenon, as a novel technology for producing software, is the subject of numerous political, economic, and sociological studies, all reacting to the potential for radical change it embodies (Chopra and Dexter 2007). These studies focus mainly on three related claims. First, FOSS provides a social good proprietary software cannot; for example, FOSS may be the only viable source of software in developing nations, where programming talent is abundant but prices for proprietary software licenses are prohibitive (Leonard 2006). Second, FOSS challenges many central concepts of intellectual property: its novel copyright licensing schemes have prompted much debate about the foundations, both ethical and economic, of apparently well-established

* Samir Chopra is an Associate Professor in the Department of Computer and Information Science at Brooklyn College of the City University of New York. He is the co-author, with Scott Dexter, of *Decoding Liberation: The Promise of Free and Open Source Software* (Routledge Series on New Media and Cyberculture, 2007). His academic interests include: philosophical foundations of artificial intelligence, legal theory of artificial agents, philosophical implications of free software, politics and ethics of technology.

** Scott Dexter is a Professor in the Department of Computer and Information Science at Brooklyn College of the City University of New York. His current research interests include creativity in software development, software ethics, and the philosophical implications of free software.

notions such as property and ownership (Dixon 2003; St. Laurent 2004). Finally, FOSS is a threat to the corporate status quo. This facet of FOSS has been trumpeted vigorously by open source advocates, who argue open source software is a new and better way of doing business: one that should, as a result of free market competition, supplant much (though not all) of the proprietarily-licensed software produced and sold today (Kelty 2002).

The promise -or threat -of FOSS springs out of the four freedoms identified in the Free Software Definition (Free Software Foundation 1996). These are: the freedom to run the program, for any purpose (freedom 0); the freedom to study how the program works, and adapt it (freedom 1); the freedom to redistribute copies; the freedom to improve the program, and release modified versions (freedom 3). A world in which all distributed software has these freedoms attached would be one in which business models relying on the exclusivity and artificial scarcity permitted by intellectual property law would no longer be viable. The upending of power relations this would entail is foreseen by the proprietary software industry with much dread and vigorous rhetoric (Greene 2001; Bishop 2008).

Claims about FOSS's potential to reform the production and distribution of software are routinely countered by arguments that are, at bottom, variants of the questions, "How could FOSS be an economically viable alternative to the software industry?" and "Will FOSS programmers be compensated at the same levels as they are in the proprietary software industry"? (Aslett 2008) That is, skeptics claim the optimism of the FOSS community fails to engage the pragmatic and economic 'realities' of a software industry. This lack of engagement then justifies dismissing the FOSS community to the discursive fringe, condemned to this position as long it refuses to provide satisfactory responses to these queries.

These questions about the viability of FOSS are often asked in tandem—indeed, they are often treated as the same question—despite their significantly different implications (Stallman 2002). A demonstration that a programmer could make a living in a FOSS economy is not considered sufficient to prove the viability of such an economy. (Interestingly, it appears to be implicitly assumed workers in economically viable industries do make a living, despite some evidence to the contrary (Madrack 2008, p. 22-31)). So what is the real economic concern with FOSS? Must FOSS indeed provide a parallel alternative to the software industry, one that preserves its current power relations, profit margins, and significance in the global economy? Would it suffice simply to guarantee adequate compensation for one component of the software industry's workforce?

The implications of these questions are significantly different depending on the position of who is asking them. A particular corporate player in the software industry might ask, "Could a business model based on FOSS help us maintain our

current profit margins and market leadership?" The current economic significance of the software industry and the likely smaller market capitalization of an imagined FOSS-based industry may indicate FOSS is not an economically realistic philosophy for the software industry. But the question of whether FOSS is "realistic" may admit a different answer if this question is instead cast as, "Can the FOSS model underwrite an economy where profits are sufficient to encourage a healthy flow of new entrants and flourishing competition amongst old and new?" These competing questions have parallels from the perspective of the industry's workforce: The question, "Will FOSS programmers be compensated at the same levels as they are in the proprietary software industry?" may be alternatively framed as, "How could a FOSS programmer make a living?"

A challenging truth about FOSS is that it implies the disbanding of a particular class of business models. Predicting the impact of such a change is difficult; one possible result is a temporary shrinking of the software industry and, possibly, reductions in programmers' salaries. But such a change will certainly bring other changes which may be less tangible but are no less important. FOSS must be understood not only as yet another software development technique or business model but also as an instrument of social justice: it is not obligated to preserve the contours of the very social order it aims to overturn. Once we acknowledge FOSS cannot, and indeed should not, support the current order within the software industry, we can explore the contours of the new economy FOSS produces and reckon with its implications.

Our argument has two tiers. First, we unpack some claims that FOSS cannot support a viable software economy. We emphasize that a change in business models should not be understood as the decapitation of an economy, and suggest an economy of software distribution based on FOSS would be far from bereft of incentives, structures, and compensation. Second, we argue that FOSS must be seen not only as an economic engine but as a tool to bring about a form of social justice. By invoking the notion of information justice (understood as an equitable, fair, arrangement of the economic and social goods derived from information technology), we show FOSS is a social good in the redistributive sense; hence its economic role is an aspect of social justice which should be of concern to ethicists and technologists alike. We draw upon a tradition of redistributive arguments addressing injustices which arise from inequitable distribution of social goods. These arguments take the following general form: they identify a social good with near-universal value, then show this good is unjustly distributed, and finally show someone in the position of being able to address this injustice is morally obligated to do so. Such an argument results in the claim the current software economy should move to a FOSS model, a move that could be facilitated, for example, by a three-fold legal reform implicit in FOSS advocacy.

Economic Justice and Redistribution

As Amartya Sen often remarks, Adam Smith's "invisible hand" is not free of ethical constraints when it manipulates the supposedly free market: "[it requires] values other than pure profit seeking, and ... restraint and correction ... for preventing instability, inequity, and injustice." (Sen 2009, Sen 1991)

The political economy theorized by Adam Smith is best understood as a combination of political philosophy, concerned with social justice in general, and normative economics, concerned with resource allocation and policy design for a just society. In such a view, free markets and social justice are not opposed, though they may sometimes place different normative constraints on a decision-maker. Markets are tools that enable certain ends; evaluations of their desirability should be guided by their success in achieving them. Insofar as they promise liberty in the trade of goods and services and fair wages for skills, they are to be desired and promoted. But when the operations of the free market conflict with other socially desirable ends, then the ethical constraints Smith saw as necessary should engage: "Achievement of social justice is a higher value than the protection of free markets; markets are mere instruments to be evaluated by their effects." (Sunstein 1997, p. 9, cited in Butcher 2007)

Locke's theory of property, often touted as the classic justification of private property in the Anglo-American context (Macpherson, 1962), similarly includes provisions for the common welfare. For example, "Whatsoever then he removes out of the state that nature hath provided, ... it hath by this labour something annexed to it, that excludes the common right of other men: for this labour being the unquestionable property of the labourer, no man but he can have a right to what that is once joined to, at least where there is enough, and as good, left in common for others." (Locke 2002, Section 27, emphasis added.) As Simmons (1994, p. 293) interprets these provisos, "What must be protected from encroachment by the appropriations of others, ... is my rights of self-preservation and self-government. This is not identical to claiming that I must be left free to appropriate land or other natural resources....What must be guaranteed to each person is the opportunity of a living—a condition of nondependence, in which one is free to better oneself, govern one's own existence, and enjoy the goods God provided for all." This qualification of Locke, then, explicitly calibrates rights to private property so as not to engender excessive inequality.

Sen observes these important clauses have influenced modern economic thinking. Describing Alfred C. Pigou's work in *The Economics of Welfare*, Sen notes economic inequality emerges as a "major indicator for economic assessment and policy," where, given the pernicious effects of such inequality, the need to ameliorate the lot of those who suffer the most deprivation is urgent and suggests "the

role of supportive cooperation between business and government cannot stop only with mutually coordinated expansion of an economy" (Sen 2009).

One implication of these views of the market is that the health of an economy cannot be measured only by its gross domestic product (GDP), which is "merely a sum of national spending with no distinctions between transactions that add to well-being and those that diminish it" (Redefining Progress 2009) – as Robert F. Kennedy put it, the GDP "measures everything...except that which makes life worthwhile" (Uchitelle 2008). Alternative indices aim to emphasize (England and Harris 1998) other dimensions along which the health and viability of an economy may be assessed. Both the Genuine Progress Indicator (Redefining Progress 2009) and the Index of Sustainable Economic Welfare (Friends of the Earth 2009) explicitly account for the effects of factors such as unequal income distribution, environmental degradation, and domestic and volunteer labor. The Human Development Index, one of many composite indices compiled by the United Nations Development Programme similarly "measur[es] development by combining indicators of life expectancy, educational attainment and income into a composite human development index" (United Nations Development Programme 2009). Measures such as these are, in the context of the modern information-based economy, "significantly, though not solely, a function of access to information, knowledge, and information-embedded goods and services" (Benkler 2006, p. 310).

A contemporary economy in which the proprietary software sector has a significant role could be assessed as healthy by the GDP because of the relative economic health of the software-for-sale industries. But other indices could reveal a less rosy picture if this economy failed to provide adequate access to "information-embedded goods and services." One cause of such denial of access is the prohibitive prices of important (or essential) software packages. As Rishab Ghosh wrote in 2003, "The price of a typical, basic proprietary toolset required for any ICT infrastructure, Windows XP together with Office XP, is US\$560 in the U.S. This is over 2.5 months of GDP/capita in South Africa and over 16 months of GDP/capita in Vietnam. This is the equivalent of charging a single-user licence fee in the U.S. of US\$7,541 and US\$48,011 respectively" (Ghosh 2003).

The developed world is not immune to the impact of the exorbitant prices of proprietary software licenses. In 2001, Microsoft threatened to sue the Philadelphia public school system for copyright infringement, as several Microsoft products had been installed on more computers than the licensing terms permitted. Critics argued, "[N]o public school should be compelled to play by the rules of an ever-changing license system that treats cash-strapped educational institutions just as it does for-profit businesses. [Philadelphia's] schools and students are some of the

poorest in the country. At the end of June, city officials announced that without a massive influx of state or federal cash, the district won't be able to pay its 27,000 employees through the upcoming school year. 'It's kind of like AIDS in Africa and the drug companies,' [a veteran computer teacher said]. 'Can anyone expect a dying person to be concerned about the drug companies' profits?'" (Cave 2001).

Proprietary software also negatively impacts the pedagogical value of software; because of its closed nature, the software-literacy it permits is limited (Chopra and Dexter 2007). Thus, in countries with an inadequate informational technology infrastructure, to commit to proprietary software is to give up hope for self-sufficiency in maintaining and growing that infrastructure. The roots of FOSS, however, are deeply tangled with education, both informal and informal. Many FOSS projects have a pedagogical imperative, often beginning from an individual's desire to learn: "Linux was started by Linus Torvalds to learn about the 386 architecture, and later to learn more about operating systems. LyX was written as a college project. The Gimp was written because its creators wanted to learn how to do graphical programming, and Gtk+ was born out of it because they wanted to learn how to write a good toolkit. FOSS fosters education. For the persons contributing to it, and for the persons consuming it" (Tellis 2005). FOSS's presence in formal computer science education is increasing as well as an increasing number of educators exploit its use for teaching software engineering principles and tools, and for providing an under-the-hood look at complex, dynamic software projects.¹

Proprietary software and its IP regimes ensure a division of the world into a 'software global North' and a 'software global South'. In a world where economic wealth is increasingly underwritten by scientific and technical prowess which in turn is dependent on information technologies, any factor that limits access to these has a deleterious impact on a society's economic well-being. Without a modern information-rich economy, a country is destined to not acquire the economic power required to become information-rich either. There is thus a vicious cycle of dependence on the software infrastructure, the power to control which lies in the hands of a small group of information oligarchs. States such as Singapore actively plan to short-circuit dependence: providing tax breaks to companies that use FOSS rather than proprietary software is "a macroeconomic decision by the government to foster the development of its own domestic industry in a certain, planned direction. It is also a strategic political decision related to a desire for political autonomy" (Deek and McHugh 2008, p. 313).

1. See for instance, the Humanitarian FOSS Project (<http://www.hfoss.org>) and Teaching Open Source (http://teachingopensource.org/index.php/Main_Page).

Thus, an argument against a FOSS economy based solely on a predicted economic contraction is devoid of persuasive power on all but the narrowest measures of economic health. But the argument against FOSS seems to totter even on its economic foundations, for an alternative economically healthy software industry appears possible and, indeed, overwhelmingly likely.

Perspectives on the FOSS Economy

It is not uncommon for the “software industry” to be conflated with the “software-for-sale industry:” the New York Times description of the industry says it “consists of companies engaged in developing and marketing system and application software....includes developers of operating systems, word processors, spreadsheet applications, CAD and database engines....[and] excludes applications customized for specific tasks requiring continuous support from developers” (New York Times 2009). However, the software industry, whose top 500 companies brought in total revenues of \$451.8 billion worldwide in 2007 is more realistically understood to include sectors such as “system integration services/IT consulting ... enterprise application/data integration, storage management, vertical industry applications, e-learning, and outsourcing services” (Desmond 2008). From the perspective of software workers, even this characterization of the software industry misses important niches in the software job market, as many companies who have little direct interaction with the software market maintain their own in-house programming groups for providing custom application software which specifically furthers their business ends.

This software industry is dominated by “information-rich” proprietary vendors whose control of access to software resources is maintained by a trifecta of legal constructions: copyright, patents, and trade secret laws. The removal of these artificial barriers to movement of information would eliminate both the temptation and possibility of software hoarding. Such a move has economic implications for both proprietary software owners (typically corporations) and the software market.

The Corporate Perspective

Redistribution of resources on the scale associated with the move to a free software regime inevitably raises questions about the structure of the resultant software industry. Will the software industry collapse because there will no longer be sufficient incentive for people to write new code (or maintain and improve existing code)? Such a question presupposes the vast majority of the value of the software market inheres in proprietary software for sale, though the breadth of industry shows the reality is quite the opposite (Raymond 2002). Even in a

FOSS-dominated economy, NASA will still need to write custom software for the Space Shuttle, space probes, and engineering infrastructure, Chase Bank will still need to maintain and innovate its custom tools for financial analysis—that is, many industries and organizations will still have a need for in-house customized software, and there will still be a market for talented and qualified programmers. As computing increasingly penetrates other areas—the sciences are becoming increasingly data-driven (Kelty 2005), and governmental applications have recently proliferated (Deek and McHugh 2008)—new industries, and new markets for programmers, will only increase.

Will that narrow segment of the software industry which consists of companies relying largely on proprietary software sales for revenue evaporate in the new economy? It is reasonable to expect that if the false model of scarcity which props up the current market value of such companies is removed, their market value will decrease, at least in the short term. A company could respond to this development in many ways, as the current troubled economic times show us. One result might be lower salaries or selective layoffs in some segment of the work force. Another result could be a substantial retooling of the company to take advantage of new opportunities inherent in the new landscape. Indeed, some historically proprietary software companies such as SUN, Oracle, and IBM, have already begun to do this by using a variety of business strategies including IBM's approach of "using open source products to facilitate its own development and support[ing] open communities whose products add value to its proprietary software" (Capek et al 2005, in Deek and McHugh 2008, p. 333) That the non-profit Software Freedom Law Center has spun off a for-profit company dedicated to servicing for-profit free software clients suggests this segment of the software economy will only continue to grow (Software Freedom Law Center 2008).

Certainly, venture capital funding for business models based on proprietary notions of software would be much less forthcoming than in the past. But venture capital has already begun funding free software initiatives such MySQL (MySQL AB 2003) and JBoss (Montalbano 2004); this development extends across the open-source industry (Radcliffe 2008). Such a development may work to benefit the industry, their customers, and the code. As "open source strategy consultant" Andrew Aitken says, "Many of the open source software companies in North America have venture backing, so it means they have more time to bring in revenue from customers. They have more time to build their product" (Sayer 2008). In a world where free software companies compete for venture capital dollars, the openness of the code itself allows venture capital

decisions to rest as much on the technical quality of the project as on its predicted economic viability.

Even in a world without proprietary software licenses, venture capital funding for free software projects might only increase: quality software necessarily underwrites any modern technological economy. If software companies are placed in the position of having to compete for venture capital investments with other technologies like biotechnology or nanotechnology, then this competition will be on the basis of the quality of their products. Thus, innovation in code-writing will still have an economic dimension: the mere absence of proprietary software will not remove economic incentives for writing good code.

At the same time, FOSS may permit the creation of new models of entrepreneurship that are much less heavily reliant on venture capital funding: "Genuitec co-founder and Vice President of Technology Todd Williams says his company is able to keep prices affordable because it avoided using venture capital money, and because Genuitec itself is built completely on open source" (Gasperson 2007), which increases the company's autonomy and responsiveness to changing conditions while decreasing customers' costs. The reduced reliance on private funding may support software entrepreneurs in the developing world: "Free software also enables commercial software development with less capitalization. The availability of free-of-charge open source systems and tools dramatically reduces the cost of entry into the business of software development" (Deek and McHugh 2008, pp. 314-315).

In such a FOSS-dominated world, software users will have greater access to high-quality software, and qualified programmers will still make a living. The question of whether programmers' salaries will increase or decrease is difficult to answer, particularly in a global economy in which a wide array of factors beyond programmers' control interact in complex ways to determine earning potential. In-house software will continue to be valuable (and will likely continue to be a commanding portion of the software market), so we might expect salaries in that sector to remain the same. This, in turn, may provide upward pressure on programmers' salaries in other sectors, as all employers will compete for the same software workers. Given the presence of venture capital funding, self-employed free software programmers and entrepreneurs will become more significant figures in the software industry. The resulting industry will be one whose power relations will be determined not by a company's ability to hoard software but by its ability to attract and retain the best programmers.

The Market Perspective

A FOSS-based industry is not a not-for-profit activity; its associated economy is one in which many participants expect to derive monetary gain. A market undergirded by FOSS has several features that bring it close to being a true free market in software. First, free software licenses avoid much of the regulatory and transactional costs of a regime based on legislated intellectual property protections. These require legal monitoring, enforcement, and political activity to maintain and extend protection (Bessen and Meurer 2008; Boldrin and Levine 2008). These overhead costs are then reflected in the price consumers pay for software. Second, bargaining for copyright and patent licenses also involves transactional costs (Bessen and Meurer 2008; Boldrin and Levine 2008). The proprietary boundaries raised by these protections cause huge inefficiencies as independent actors are forced to reinvent software wheels. For example, the Free Software Foundation's list of high-priority projects (Free Software Foundation 2008) shows that many FOSS projects are "replacements" for proprietary products. Finally, proprietary software promotes free trade in software only until the point of initial purchase. From that point onwards, the customer "depends exclusively on the chosen-software's vendor to obtain changes or fixes needed in the software" (Oliva and Rezende 2008).

In a FOSS-based market, there is free trade in initial offering and support services alike. This freedom is available precisely because of FOSS licensing terms that allow competitors to modify and redistribute the original code. The customer is free to choose the appropriately modified or customized version of the code, perhaps even by picking the original vendor if it meets the customers cost and functionality constraints. The FOSS market extends to knowledge-based "services such as development, support, training, tuning" (Oliva and Rezende 2008). The vitality of this extended market depends again on the terms of FOSS licenses, which ensure a rich commons of code. The health of this market would be most directly assessed by the value of programmers' salaries rather than by the price paid for software licenses. These salaries would be a function of the value placed upon programmers' talents in an open marketplace by prospective employers who would bid for programmers depending on their competence in writing, repairing and enhancing various products in the FOSS commons.

We are already beginning to see the contours of what a free software-based market might look like. Certain sectors of the contemporary software industry, particularly server software and embedded software, are already dominated by free software, and other "mission-critical" business applications are expected to shift toward free software in the near future (Wheeler 2007). Free software projects, as well, show that programmers who work primarily on FOSS can indeed make

a living (Iansiti and Richards, 2006). This new market will still display differential power relations, but instead of being based on the power granted by code-hoarding, it will be the power inherent in any concentrations of programming knowledge, that might result a more meritocratic situation, one more amenable to change and not propped up artificially by a legal regime. Just as Richard Stallman supported himself by being an expert on Emacs customization (Williams 2002), individuals and corporations could become dominant through expertise in particular domains: perhaps Microsoft could become the leading expert in office productivity software, and Google could become the expert on software-as-a-service.

Information Justice and Software Justice

To this point, we have argued a software industry based on FOSS is a viable notion. But the argument for FOSS derives its true normative weight from social justice considerations. The evaluation of a particular economic regime should be guided not only by the economic qualities of the market in which it is based but also by consideration of the social and cultural states which are the ultimate goals of any economic arrangement. In the case of software, we must consider, then, how best to achieve the socially desirable states which the production and distribution of information are supposed to enable.

To aid our thinking, we consider three perspectives on the nature and importance of information:

1. "Information...is [a] core input into human welfare;" (Benkler 2006, p. 302)
2. "Information [is] a social good and a central element in the development and creation of a democratic society;" (Schiller 1996, p. 38) and
3. "[I]nformation '[is] a privately produced commodity for sale.'" (Schiller 1996, p. 38)

The first two statements speak to aspects of the social utility of information, while the third characterizes a particular view of information production and distribution; each statement is true, to some extent, of the current world. Attempts to reform the software industry could be characterized as efforts to protect the value of information as described here, while expanding the possibilities for production and distribution of information. Any such reform will have redistributive consequences: informational wealth will be transferred from those currently empowered by the exclusive rights granted them by the intellectual property regime that underwrites the software economy. The possibility of such a redistribution raises a fundamental ethical question—is such a redistribution morally justified?

It raises too, a more straightforwardly economic question-what would such a re-distribution look like? We consider the second question first.

In general, the economic justice question of how to maximize welfare can be given a theoretical answer based on computing the ideal distribution of resources. Analyses of economic justice are generally carried out in terms of rival resources, but in the case of software, rivalrousness is imposed by legal structures, so analyses of economic justice must acknowledge its non-rival and indeed, anti-rival nature (Benkler 2006, pp. 301-355). Given that software is inherently anti-rival the ideal distribution of software is simply the maximal distribution. Clearly, in a “free software world,” software would be maximally distributable. Such a distribution of code would be Pareto-efficient (no other feasible allocation is better for some individuals and worse for none) and envy-free (no individual would prefer having the bundle of another), though differential access to and costs of distribution will skew perfect availability. Armed with the knowledge of this desired end-state, the task for the social justice theorist is to give prescriptions for how it may be realized.

The most significant contemporary reform task is directed toward the intellectual property regimes that underwrite proprietary software. The most powerful arguments are those that place proprietary software owners under an ethical obligation to release their code under a free software license. The eloquent arguments of Peter Singer (2009) and Peter Unger (1996), directed toward those who could make a minimum contribution of monthly income to the alleviation of poverty in economically underdeveloped countries, can be recast and directed toward the “software-wealthy” who could license their software to alleviate the socially deleterious effects of widespread information deprivation:

1. The degradation of the quality of one’s life due to lack of access to the benefits of information is a bad thing.
2. If it is in your power to prevent something bad from happening, without sacrificing anything nearly as important, it is wrong not to do so.
3. By releasing your code under a FOSS license you can prevent the degradation of quality of life due to information deprivation, without sacrificing anything nearly as important.
4. Conclusion: if you do not release your code under a FOSS license, you are doing something wrong.

The crucial phrase in this argument is “without sacrificing anything nearly as important.” FOSS-licensing code does not remove access to the code itself from the owner. Thus, the only possible sacrifice is the owner’s ability to extract monopoly rent from this code. In other words, a software owner might argue that this ethi-

cal obligation imposes an economically intolerable burden. But such a move takes us back to the economic argument that we began by refuting.

Perhaps even the combination of our redistributive argument with our rejection of protestations of economic distress might not be persuasive to the software owner who is guaranteed higher profits by virtue of his monopoly on the product. This suggests a more active role for legislating such redistribution. When is such legislative compulsion morally justifiable? This depends not on contingent ownership relations but on whether such ownership was bestowed in just fashion: "where possessions have been acquired through unjust processes, purposive takings may be required to restore rightful possession" (Barry 2004). Such a rationale has been used in redistributive arguments implicitly underwriting many public policy decisions (e.g. Janani, McElroy, and Monroe 1975). For example, the compulsory licensing exceptions (on public health grounds) in the International Agreement on Trade Related Aspects of Intellectual Property Rights (TRIPS) are based on the understanding the monopoly rent that might accrue to patent holders can be withheld if there is a greater social imperative at hand (Ford 2000; Lang 2004). Thus pharmaceutical drugs makers are compelled to license their drugs for manufacture and sale without the intervening stage of patent license negotiation if there is a public health crisis that requires the broader availability of the patented drug. Here it is recognized that an intellectual property protection derives its validity from the social ends it ensures, and that its continuation is only morally justified if it does not bring about a social harm.

Reform of intellectual property legislation which applies to software is the most direct legislative strategy to support information justice in the software domain. These laws are best understood as imposing constraints on the flow of information to consumers and downstream producers with the intent of incentivizing the further production of informational goods. One contemporary direction of reform is to view availability and distribution of software as ends in themselves that should be supported by the law (Davis et al 1996; Samuelson et al 1994). The simplest formulation of these reforms which would achieve the ends we propose is the following triad: distributed software executables should not be copyrightable; EULA licenses should conform to the Free Software Definition; patents should not be granted on software and algorithms. Of these, the most reformatory energy is directed towards reform of the patenting system (especially in the US) as even proprietary software vendors are aware of the dangers to continued software innovation of excessive patent protections for software (and the ensuing expensive and time-consuming litigation) (Bessen and Meurer 2008). A recent ruling against business method patents by the US Court of Appeals for the Federal Circuit is an important step in the direction of removing patent protec-

tions for algorithms². Given the significant lobbying power of the US software industry, progress on the other two fronts of the proposed IP reform looks harder to achieve.

But legislation might not be the only strategy available. It may be that consumers could express their preferences for FOSS and governments might actively promote FOSS because there are public interests in supporting FOSS as a configuration for the software industry. Economist Dean Baker has suggested a governmental expenditure of \$2 billion a year to support the development of Linux and other free software (Baker 2009). Envisioning a situation in which computer ownership is ubiquitous, backed up by free operating systems and application software, Baker projects per capita savings of \$200 a year. Sales of \$20 million a year would suggest the savings in the US alone “easily exceed the cost of supporting software development.” Furthermore, the public sector is where “decisions on national economic objectives and strategy are made...[and which can] significantly affect the expansion of open source use within the government” (Deek and McHugh 2008, p. 309). Indeed, a variety of governmental agencies (including the National Security Agency) are engaged in developing and distributing open source software (Deek and McHugh 2008, p. 310).

There may be too, a need for protecting domestic markets from the monopolistic pressures of an international proprietary software player (this is most visible in the more-aggressive European antitrust actions against Microsoft that have gone further than comparable American efforts) (Meller 2008). As the FSF Europe notes, the FOSS imperative is not directed at particular companies but rather at the presence of monopolies: thus, while there might market leaders in a FOSS economy they will not be monopolists. Because an important problems facing the European IT industry is its reliance on international IT monopolies, “weakening these monopolies has become necessary for Europe to prosper” (Free Software Foundation Europe 2003).

Conclusion

Free and open source software is already demonstrating its economic significance, as governments, venture capitalists, and corporations move more fully to embrace it. Far from portending disaster, this trajectory suggests that a software economy based on FOSS is natural and, perhaps, inevitable. Such a reconfigured software industry will not preserve the economic disparities or power relations of the older one. But plenty of software will be written nonetheless: the dependence of the world’s economic and technical engines on software is guaranteed to drive

2. In re *Bilski*, 545 F.3d 943, 88 U.S.P.Q.2d 1385.

further innovation and production in software. As scholars of intellectual property never tire of pointing out, its protections are neither necessary nor sufficient for innovation and the public good, and indeed, might even be detrimental those ends. The important precedent of the public health provisions in the TRIPS agreement show conclusively that concern for social benefits can trump the narrowly economic demands of a particular business model. The redistribution of the economic, educational and technical benefits of software is a moral imperative in today's increasingly technical world. Free software's most salutary contribution to this debate is that it has shown, both by theory and praxis, how such ends may be achieved.

The proprietary software industry came to maturity in the US, and its most powerful players are still American. But it is more likely that FOSS companies will find it easier to attract manpower and funding in emerging software markets. Thus a combination of consumer preferences, governmental activism, and global competition, and the continuing commoditization of software might force software vendors to start competing in a FOSS world. If Microsoft's Open Source Labs are any indication, such a process is already under way. And as a new generation of computer science students comes of age in an increasingly FOSS dominated world, their understandings of the legal control of the software will be radically different. This change, of all the ones considered here, might well be most significant. A revolution in the intellectual atmosphere is more likely to be conducive to the deep reforms we need in the information sphere.

REFERENCES

- Aslett, Matthew.** (2008) Open source is not a business model. October 13, 2008. 451 CAOS Theory [online]. [Accessed April 3, 2009]. available from World Wide Web: <http://blogs.the451group.com/opensource/2008/10/13/open-source-is-not-a-business-model/>
- Baker, Dean.** (2005) Opening Doors and Smashing Windows: Alternative Measures for Funding Software Development . Center for Economic and Policy Research [online]. [Accessed March 15, 2009]. Available from World Wide Web: http://www.cepr.net/documents/publications/windows_2005_10.pdf
- Baker, Dean.** (2009) Yes, we can make the stimulus more stimulating. January 12, 2009. truthout.org [online]. [Accessed March 15, 2009]. Available from World Wide Web: <http://www.truthout.org/011209R>.
- Barry, Christian.** (2004) Redistribution. Stanford Encyclopedia of Philosophy [online]. [Accessed March 15, 2009]. Available from World Wide Web: <http://plato.stanford.edu/entries/redistribution/>.

Bessen, James and Meurer, Michael J. (2008) Patent failure: How judges, bureaucrats, and lawyers put innovators at risk. Princeton: Princeton University Press.

Bishop, Todd. (2008) Ballmer: Free software means no free soda. February 8, 2008. The Microsoft Blog [online]. [Accessed April 2, 2009]. Available from World Wide Web: <http://blog.seattlepi.com/microsoft/archives/131418.asp>

Boldrin, Michele and Levine, David K. (2008). Against intellectual monopoly. New York, NY: Cambridge University Press.

Butcher, Matthew. (2007) Software and the information oligarchy [online]. [Accessed March 1, 2009]. Proceedings of the 2007 Loyola University of Chicago Calling to Justice Conference. Available from World Wide Web: <http://www.luc.edu/ministry/pdfs/conference%20docs/CallingToJusticeConferenceFeb242007Butcher.pdf>

Butcher, M. (2009) At the foundations of information justice. Ethics and information technology 11(1), pp. 57–69.

Cave, Damien. (2001) Microsoft to schools: Give us your lunch money! Salon.com [online]. [Accessed April 5, 2009]. Available from World Wide Web: http://archive.salon.com/tech/feature/2001/07/10/microsoft_school/

Chopra, Samir and Dexter Scott (2007) Decoding liberation: The promise of free and open source software. Routledge, New York.

Dajani, Jarir, Egan, M. Michael, and McElroy, Marjorie B. (1975) The redistributive impact of the Atlanta mass transit system. Southern Economic Journal 42(1): 49–60.

Davis, Randall, Samuelson Pamela, Kapor, Mitchell, and Reichman, Jerome. (1996) A new view of intellectual property and software. Communications of the ACM 39(3), pp. 21–30.

Deek, Fadi and McHugh, James (2008) Open source technology and policy. Cambridge University Press.

Desmond, John P. (2008) Innovation alive and well. SoftwareMag.com [online]. [Accessed March 20, 2009]. Available from World Wide Web: <http://www.softwagemag.com/L.cfm?Doc=1175-10/2008>

Dixon, Rod. (2003) Open source software law. Norwood, MA: Artech House Publishers

England, J. and Harris, J. M. (1998) Alternatives to Gross National Product: A critical survey [online]. [Accessed March 20, 2009]. In Ackerman, Kiron, Goodwin, Harris, and Gallagher, eds. Human wellbeing and economic goals. Washing-

ton, D.C.: Island Press. Available from World Wide Web: <http://ase.tufts.edu/gdae/publications/archives/englandpaper.pdf>

Ford, Sara M. (2000) Compulsory licensing provisions under the TRIPs agreement: Balancing pills and patents, 15 Am. U. Int'l L. Rev. 941, 965.

Free Software Foundation. (1996) The Free Software Definition [online]. [Accessed March 26, 2009]. Available from World Wide Web: <http://www.gnu.org/philosophy/free-sw.html>.

Free Software Foundation. (2008) High priority free software projects [online]. [Accessed April 4, 2009]. Available from World Wide Web: <http://www.fsf.org/campaigns/priority.html>.

Free Software Foundation Europe. (2003) Free software in Europe: European perspectives and work of the FSF Europe [online]. [Accessed March 3, 2009]. Available from World Wide Web: <http://www.fsfeurope.org/documents/eur5-greve.en.html>.

Friends of the Earth. (2009) How ISEW terms are calculated [online]. [Accessed March 30, 2009]. Available from World Wide Web: <http://www.foe.co.uk/community/tools/isew/brief.html>.

Gasperson, Tina. (2007) Open source software lets Genuitec forgo venture capital. Linux.com [online]. February 13. [Accessed March 14, 2009] Available from World Wide Web: <http://www.linux.com/articles/60092>.

Ghosh, Rishab Aiyer. (2003) License fees and GDP per capita: The case for open source in developing countries. First Monday 8(12) [online]. [Accessed March 29, 2009]. Available from World Wide Web: http://firstmonday.org/issues/issue8_12/ghosh/.

Greene, Thomas C. (2001) Ballmer: "Linux is a cancer" [online] [Accessed March 25, 2009]. Available from World Wide Web: http://www.theregister.co.uk/2001/06/02/ballmer_linux_is_a_cancer/.

Iansiti, Marco and Richards, Gregory L. (2006) The business of free software: Enterprise incentives, investment, and motivation in the open source community [online]. [Accessed March 20, 2009]. Harvard Business School Working Paper Series, No. 07-028. Available from World Wide Web: <http://www.hbs.edu/research/pdf/07-028.pdf>.

Kelty, Christopher. (2002) How to do things with words [online]. [Accessed March 5, 2009]. Available from World Wide Web: <http://www.kelty.org/or/papers/Kelty.Hautodothings.2002.rtf>.

Kelty, Christopher. (2005) Free Science. In Feller et al, eds. Perspectives on free and open source software. MIT Press. Cambridge, MA.

Lang, Mark C. (2004) What a long, strange "TRIPS" it's been: Compulsory licensing from the adoption of TRIPS to the agreement on implementation of the Doha declaration. 3 J. Marshall Rev. Intell. Prop. Law 331.

Leonard, Andrew. (2006) Free software, big oil, and venezuelan politics. Salon.com [online]. January 12. [Accessed March 15, 2009]. Available from World Wide Web: http://www.salon.com/tech/htww/2006/01/12/venezuela/index_np.html.

Locke, John. (2002) The second treatise of civil government [online]. [Accessed April 3, 2009] Available from World Wide Web: <http://www.constitution.org/jl/2ndtreat.htm>.

Macpherson, Crawford Brown (1962) The political theory of possessive individualism: Hobbes to Locke, Oxford: Clarendon Press.

Madrick, J. (2008) The case for big government. Princeton: Princeton University Press.

Meller, Paul. (2008) Microsoft faces two new European antitrust cases. InfoWorld [online]. January 14. [Accessed April 9, 2009]. Available from World Wide Web: <http://www.infoworld.com/t/platforms/microsoft-faces-two-new-european-antitrust-cases-504>.

Montalbano, Elizabeth (2004) Intel joins venture capital firms as investor in Jboss. CRN.com [online]. May 10. [Accessed March 21, 2009]. Available from World Wide Web: <http://www.crn.com/it-channel/18842003>.

MySQL AB. (2003) MySQL AB Secures \$19.5 Million Investment Round Led by Benchmark Capital [online]. [Accessed March 20, 2009.] Available from World Wide Web: http://www.mysql.com/news-and-events/generate-article.php?id=2003_18.

New York Times. (2009) Software industry snapshot [online]. [Accessed April 2, 2009]. Available from World Wide Web: <http://markets.on.nytimes.com/research/markets/usmarkets/industry.asp?industry=57212>.

Oliva, A. and Rezende, P. On the constitutional preference for free software. Free Software Foundation-Latin America [online]. [Accessed February 24, 2009]. Available from World Wide Web: <http://www.fsfla.org/svnwiki/texto/pref-const-br-swl.en>

Radcliffe, Mark. (2008) Venture capital funding for open source shows significant increase in first quarter. Law & Life: Silicon Valley [online]. May 20. [Ac-

cessed March 24, 2009]. Available from World Wide Web: <http://lawandlifesiliconvalley.com/blog/?cat=69>.

Raymond, Eric S. (2002) The Magic Cauldron [online]. [Accessed March 28, 2009]. In The cathedral and the bazaar: Musings on Linux and open source by an accidental revolutionary. Sebastopol, CA: O'Reilly Media. Available from World Wide Web: <http://catb.org/~esr/writings/cathedral-bazaar/magic-cauldron/index.html>.

Redefining Progress. (2009) Genuine Progress Indicator [online]. [Accessed March 20, 2009] Available from World Wide Web: http://www.rprogress.org/sustainability_indicators/genuine_progress_indicator.htm.

Samuelson Pamela, Davis Randall, Mitchell Kapor, Reichman Jerome (1994) A manifesto concerning the legal protection of computer programs. Columbia Law Review, 2308-2431.

Sayer, Peter. (2008) Free software: It's about the money. ITWorld.com [online]. September 24. [Accessed April 3, 2009]. Available from World Wide Web: <http://www.itworld.com/open-source/55316/free-software-its-about-money>.

Schiller, Herber (1996) Information inequality. Routledge.

Sen Amartya (2009) Capitalism beyond the crisis. The New York Review of Books, 56(5).

Sen Amartya (1991) On Ethics and Economics. New York: Wiley-Blackwell.

Simmons, A. John. (1994) The Lockean theory of rights. Princeton University Press.

Singer, Peter. (2009) The life you can save: Acting now to end world poverty. New York: Random House.

Software Freedom Law Center. (2008) Software Freedom Law Center Announces Formation of Moglen Ravicher LLC [online]. [Accessed April 2, 2009]. Available from World Wide Web: <http://www.softwarefreedom.org/news/2008/mar/26/moglen-ravicher>.

St. Laurent, Andrew M. (2004) Understanding open source & free software licensing Sebastopol: O'Reilly.

Stallman, Richard (2002) Why software should not have owners. In Gay, ed. Free software, free society: Selected essays of Richard Stallman. GNU Press, Cambridge, MA.

Tellis, Philip (2005) Why FOSS in education makes sense. The Other Side of the Moon [online]. November 13. [Accessed April 2, 2009]. Available from World

Wide Web: <http://tech.bluesmoon.info/2005/11/why-foss-in-education-makes-sense.html>.

Uchitelli, Louis (2008) Hey, big number, make room for the rest of us. New York Times [online]. August 31. [Accessed March 25, 2009]. Available from World Wide Web: <http://www.nytimes.com/2008/08/31/weekinreview/31uchitelle.html>.

Unger, Peter (1996) Living high and letting die: Our illusion of innocence. New York: Oxford University Press.

United Nations Development Programme. (2009) The Human Development Index (HDI) [online]. [Accessed April 9, 2009]. Available from World Wide Web: <http://hdr.undp.org/en/statistics/indices/hdi/>

Williams, Sam. (2002) Free as in freedom: Richard Stallman's crusade for free software. Sebastopol, CA: O'Reilly Media.

World Trade Organization. (2005) Frequently asked questions about TRIPS in the WTO [online]. [Accessed March 30, 2009]. Available from World Wide Web: http://www.wto.org/english/tratop_e/trips_e/tripfq_e.htm#CompulsoryLicensing

Wheeler, David C. (2007) Why open source software / free software (OSS/FS, FLOSS, or FOSS)? Look at the numbers! [online]. [Accessed March 18, 2009]. Available from World Wide Web: http://www.dwheeler.com/oss_fs_why.html#market_share

Living in the Eye of the Artificial Other Artificial Intelligence, Moral Imagination, and Disciplining

Mark Coeckelbergh*

Department of Philosophy, University of Twente

Abstract

Developments in robotics and artificial intelligence suggest a near-future scenario in which technology explicitly assumes a social function. In this paper, I develop a conceptual perspective for understanding living together with artificial others and draw some implications for ethics. In particular, I argue that we can understand some robots and some AI environments as constituting quasi-others, which, to the extent that they are social others, may become 'generalised others' or 'impartial observers' and therefore subject us to socialisation and moral disciplining. I conclude that ethical issues raised by such intelligent systems are not be limited to the protection of privacy and are better re-conceptualized as problems we have as social beings trying to live good lives.

Keywords: robots, artificial intelligence, ethics, quasi-other, moral imagination, disciplining, privacy

Introduction

Developments in robotics and artificial intelligence (AI) suggest a near-future scenario in which technology explicitly assumes a social function. Robot designers expect that so-called personal robots or social robots will increasingly take part in our daily, personal lives. They put forward a vision of robots leaving the factories of industrial production and entering our households, our care institutions, and perhaps our personal relationships. Such a project raises ethical issues, such as 'Is it right to have sex with robots?' (Levy 2007) and 'Can robots contrib-

* *Mark Coeckelbergh* is Assistant Professor at the Philosophy Department of the University of Twente in The Netherlands. He received his Ph.D. from the University of Birmingham, UK. His publications include *Imagination and Principles* (2007), *The Metaphysics of Autonomy* (2004), and *Liberation and Passion* (2002), and articles on topics such as responsibility, risk, imagination, and justice. In his current research he responds to ethical, political, and anthropological issues concerning robotics, information technology, and human enhancement.

ute to the good life?' We also may want to invoke Orwell's image of 'Big Brother' in order to express our ethical concerns with AI systems that monitor us at home and respond to our (perceived) needs.

Usually the latter concern is phrased in terms of the 'privacy' principle: AI threatens our privacy. Privacy is a long-standing issue in ethics of information technology (see for example Van den Hoven 2008; Tavani 1999; Nissenbaum 1998). The application of AI technology in the personal sphere seems particularly threatening with regard to privacy. Do we want to live in intelligent electronic environments full of AI systems and AI entities that watch us day and night?

In this paper, I take one step back and ask about the nature of the 'social' role of intelligent technology. What would it mean to understand robots and AI systems as social entities that live with us? What is their precise social role? And what are the implications for ethics? I develop a conceptual perspective for understanding living together with artificial others and draw some implications for ethics.

Engaging with various perspectives from philosophy of technology, I first distinguish three ways in which technology can be understood as social. Using Smith, and Mead I then argue that AI technologies cannot only become quasi-others but also social others that launch our moral imagination as social beings and discipline us. Finally, I discuss the implications for ethics and show how this approach re-conceptualizes the problem.

Technology and the social

There are at least three ways in which technology can be understood as 'social':

1. Some artefacts are designed to have a social function. For instance, a telephone is meant to connect people. A robot could be designed to be a companion or to play the role of a parent, teacher, or child. On this instrumental view of technology, a robot can be social if it is meant to be social, that is, if the designer had this aim in mind, if it is intended by the designer. For ethics, then, the task is to make sure that designers have goals that are morally good. For instance, engineers can design robots that assist education, but they can also design military robots that collaborate with human soldiers on the battlefield. Both kinds of robots are 'social', in the sense that they interact and communicate with humans, but the latter application is usually seen as morally problematic. Either way, the ethical analysis concerns the goal(s) of the technology, not the means (the technology itself). For example, if an AI technology is meant to assist elderly people, make their life easier, and allow them to stay longer in their own homes, then it can be argued that this is good since the aim is good.

2. In a broad (and perhaps more radical) sense technology has always had a social function: contemporary philosophy of technology teaches us that artefacts mediate between people and do things (Verbeek 2005) Technology is not a 'mere instrument' that serves our pre-set aims and purposes, but does something more fundamental to the way we live together and changes our goals. For instance, a mobile phone changes how we organize our social lives – it changes that life itself. And from this perspective robots that are used in health care do not only serve the aims set by designers and health care professionals, but are also likely to change health care practices themselves. With the technology, the meaning of the practice changes. Computers became more than calculators and typewriters; they changed the way we work, live, and think.

My own version of this non-instrumental view of technology is inspired by Arendt's claim in *The Human Condition* that things are 'in-betweens'. The social life depends on a material world. By constituting a world, things relate people to one another. The table between us gather us, renders possible the social. Our social world is a world of people and things (Arendt 1958).

To live together in the world means essentially that a world of things is between those who have it in common, as a table is located between those who sit around it; the world, like every in-between, relates and separates men at the same time. (Arendt 1958, p. 52)

For an ethics based on this view, the task is to evaluate the social consequences of artefacts: what is their contribution to the social world? Do they relate people and how? For example, a baby robot in elderly care environments can relate people by making people collaborate to care for it, by making them talk about the robot, e.tc. (Kidd et al. 2006). A social world develops around the robot.

More generally, technology is not mere 'decoration', the background, or the stage on which then the social life develops. While it also provides part of the material conditions, it does more as well: it shapes the social life itself. It changes the script, the words, the scenes, and the actors. It does not play a marginal role; it co-constitutes the social by contributing to its conditions, its development, and its flourishing.

3. Viewed from these two perspectives, however, we cannot make sense of what makes the social role of artificially intelligent systems different from that of a mobile phone or a table. What is specific about the social dimension of this technology that warrants extra philosophical attention? In order to bring this to the foreground, let me explore an alternative conceptual route.

Apart from being designed for a particular purpose (social or other) and in addition to playing a more basic role as a condition and constituent of the social world,

some AI systems tend to become what Ihde calls a 'quasi-other' (Ihde 1990). Ihde uses the term in *Technology and the Lifeworld* when discussing what he calls 'alterity relations': the technology is neither in the background nor embodied (part of me), but becomes an other to which I relate (Ihde 1990, p. 97-108).

Let me explain this. With many AI systems we may have a background relation, for instance an intelligent climate control system. Some might be embodied, for example glasses that highlight information about our environment. But when we have a conversation with a personal, artificially intelligent household robot, we are likely to engage in an alterity relation with that robot: it appears an other. Perhaps we give a name to the robot, for example. This is very plausible. We already treat some pet animals as if they were others. We neither experience them as part of us, nor are they merely 'part of the furniture'. They are quasi-others.

In order to further develop this perspective, I propose to understand these alterity relations as social relations. Consider how we interact with robots that resemble us. Although we are usually perfectly aware that a particular robot is not human, our social response to that robot is likely to resemble our social response to human others. For instance, Hiroshi Ishiguro has used humanoid robots to study (human) interaction (Ishiguro 2006; MacDorman and Ishiguro 2006). This is only possible since we tend to treat humanoid robots as social others. We do not treat such robots as 'mere things' and we are usually unaware of the social nature of all things in an Arendtian sense (things that gather us); instead, we tend to include them in our social world as we would do with fellow humans (and indeed some animals).

Note that to have a social or quasi-social relation, AI robots or other AI system need not be humanoid. Presently, we often treat animals and computers in a 'social' way. For example, Reeves and Nass have shown that interactions with computers are similar to social relationships with humans (Reeves and Nass 1996).

Note also that from this perspective, it makes no difference whether or not an AI system has consciousness or is sentient. Our treatment of (some) robots as social others is not based on an ontology of mental properties, but on social function. What counts is that the entity in question appears to us (humans) as a quasi-other, regardless of the properties the entity actually has.

What does this perspective imply for ethics? As I said in my introduction, one worry we might have is that the technology becomes 'Big Brother', that is, a quasi-other that watches us and perhaps restrains our behaviour. But does this happen, and if so, to what extent is that a problem? To better understand the social dimension of AI technologies as quasi-others, we must turn from philosophy of technology towards moral and social philosophy tailored to humans. The concep-

tual resources available to us in this area are vast; let me select and employ one perspective: a social ethics that recognizes the role of moral imagination.

The individual, the social, and morality: Moral imagination and disciplining

There are at least two different and opposing views on the relation between the individual and the social (and the moral). According to the first view, there are 'first' individual, self-conscious individuals with a self, which raises the question how the social and the moral is possible. Hence we have theories such as contractarianism that try to understand and justify the existence of society and political social institutions. According to the second view, it is the social that is more 'original'; self-consciousness, individuality, self, and morality emerge from the social rather than the other way around. It is the individual, not the social that stands in need of explanation. Here I explore a route that has more affinity with the second view and try to apply it to the 'Big Brother' issue with robots and AI systems.

I already argued that some robots and AI systems can be understood as quasi-others. Let me now say more on how they can be interpreted as social others and what this implies for morality.

Some traditions in moral philosophy highlight that we always live in the eye of the other. We are moral beings as social beings. Let me explain this by drawing on moral sentiment theory and pragmatism. In *The Theory of Moral Sentiments* (1759) Adam Smith shows that when judging our own conduct our moral imagination is always directed to others. This takes on two dimensions. First, we judge our own conduct by placing ourselves in the situation of a particular other. We ask ourselves what that particular other would say of our conduct. Second, we also generalize the particular other. We ask what 'one' would say of our conduct. Smith uses the concept of 'the impartial spectator' to explain this:

We endeavour to examine our own conduct as we imagine any other fair and impartial spectator would examine it. If, upon placing ourselves in his situation, we thoroughly enter into all the passions and motives which influenced it, we approve of it, by sympathy with the approbation of this supposed equitable judge. If otherwise, we enter into his disapprobation and condemn it. (Smith 1759, pp. 109-110).

In a similar vein, Mead has argued in *Mind, Self, and Society* that humans are mainly social beings who develop themselves (morally) by taking the attitudes of our community as their values (Mead 1934). Mead explains this process by putting forward the concept of the 'generalized other':

He is putting himself in the place of the generalised other, which represents the organised responses of all the members of the group. It is that which guides conduct controlled by principles (Mead 1934, p. 162).

Thus, for Mead moral development is not so much a matter of gaining knowledge of abstract principles (and applying them) but of using one's moral imagination understood as social imagination. Principles embody the generalised other; they are not something that stands apart from the social. Both Smith and Mead share the view that morality is deeply social and imaginative: when we evaluate our conduct, we imagine how the other and 'one' would judge it.

If others are that important for our moral self-evaluations, this has implications for quasi-others as well. To the extent that we regard and treat AI robots and other AI systems as quasi-others, they can fulfil the following double role. First, they can act as particular social others. Humans are likely to evaluate their own conduct by asking questions such as 'What would the system say if I did that?' or 'How would the robot feel about that?' – even if the humans know that the robot is not a 'real' other. Second, when artefacts appear as social quasi-others they can come to be imagined as embodying the values of the community. They can be seen as representing the impartial spectator, the generalized other. In this role, the technology confronts the humans involved with what they imagine as the evaluation of their conduct by the community.

In both roles, the quasi-others will have a disciplining effect. In *Discipline and Punish* (1975) and other works Foucault has pointed to the many subtle ways in which power is exercised over individuals by institutions. This also depends on artefacts. Many technologies already have disciplining effects in virtue of what they do to us: they facilitate some behaviour while discouraging or preventing other behaviour. Consider a speed bump: it constrains our conduct. This was one of the ways to understand the 'social' role of technology I mentioned above. However, the disciplining by the particular or generalised other (in the sense meant by Smith and Mead) is more subtle and follows a less material route: individuals will (not) choose a particular course of action because they imagine that the robot would (not) approve of it or because they will experience the 'eye' of the technology as the eye of the community that would (dis)approve of their action. They adapt their behaviour to these expectations and imagined judgments.

For ethics, then, the question is not if imagining the view of the generalised other and if disciplining as such is right or wrong. As social beings, we do so anyway. We imagine what others would say of our conduct and in this way we discipline ourselves. With respect to the social-imaginative process described above, the ethical question must concern (1) the rightness and quality of the imagined moral judgment (whatever the real robot or real humans would say) and (2) the right-

ness and value of the behaviour and life that result from the disciplining. Thus, with regard to AI technologies as quasi-others, the main evaluative question is not: 'Does Big Brother watch me and is that right?' but 'Is his judgment – as I imagine it – right and does he make me do things that are right and good? Does this technology – in its role as quasi-other that makes me imagine how others would evaluate my behaviour and that therefore disciplines me – facilitate a good life?' Instead of limiting ethical issues to what harm can be done to the individual by others or by society, this approach first accepts that we are social beings and asks from that starting point how individual conduct and individual lives are to be shaped. Instead of imagining the individual mainly as a fortress that needs to be defended against invasion, humans are understood as heavily dependent on others and on their community, as deeply connected to those others and that community both in their interactions and in their moral imagination. Others are not 'outside'; they are already in our minds. To the extent that AI technologies can appear as an other, they become part of that social-moral constellation and have to be studied and evaluated in those terms.

Conclusions

I have argued that we can understand some robots and some AI environments as social quasi-others. This implies that as particular others and as launching sites for imagining the 'general other' or the 'impartial observer' they subject us to moral disciplining. As such, they can have a two-fold moral-social significance similar to human others: they too can be considered as particular quasi-others we have to live with and as gateways to imagining the 'generalized other'.

Whether or not this moral disciplining is morally acceptable is not the right question since it is unavoidable as part of how we think as social beings. In general, moral disciplining is not necessarily bad in itself. It plays an important role in education and moral development, for example. Social institutions cannot do without it. Consider what car traffic would look like without disciplining. And perhaps good care or good education always violates the privacy of individuals. This does not mean that all AI technologies are always good. Particular technologies will certainly be morally problematic or even unacceptable because they contribute to wrong actions or low-quality lives. But I have argued that, with respect to technologies as social quasi-others, the ethical question concerns not so much the technology itself but the quality of our self-judgments and the behaviour and lives that result from these self-judgments. The importance of moral-social imagination does not imply that we are bound to follow social conventions or have to accept any technology. The technology as quasi-other launches our social imagination and renders possible self-judgment, but what matters morally speaking is the quality of that judgment and of our lives with that technology.

Note, finally, that both the social imagination and the disciplining effects it has may differ between cultures and societies. In some cultures, the social dimension of life is more important than in others. This has to be taken into account when evaluating robots and other AI systems as quasi-others.

I conclude that my re-conceptualization of the 'Big Brother' issue as a problem of quasi-others and social disciplining, changes the problem formulation. Usually the issue is formulated as a 'privacy' problem or in terms of other principles that have to do with protecting individuals from the state or from others. But rather than a problem of individuals threatened by technology, it is here viewed as a problem we have as social beings that use our imagination to judge ourselves. The ethical question is then about evaluating that judgment and the role of particular technologies that contribute to that judgment as quasi-others: what do they do to our moral-social imagination, our conduct, and our lives? This, I conclude, is my tentative answer to the main question I asked in the beginning of this paper: what it would mean to understand certain robots and AI systems as social entities assuming social functions, and what this implies for ethics.

REFERENCES

- Arendt, H.**, (1958). *The Human Condition*. University of Chicago Press, Chicago.
- Foucault, M.**, (1975). *Discipline and Punish*. (trans. A. Sheridan) Random House, New York.
- Ihde, D.**, (1990). *Technology and the Lifeworld*. Indiana University Press, Bloomington/Minneapolis.
- Ishiguro, H.**, (2006). Android science: conscious and subconscious recognition. *Connection Science*, 18 (4), December 2006: 319-332.
- Kidd, C., Taggart, W., and Turkle, S.** (2006). A Sociable Robot to Encourage Social Interaction among the Elderly. *IEEE International Conference on Robotics and Automation (ICRA)*, Orlando, Florida, U.S.A., May 2006.
- Levy, D.**, (2007). *Love and sex with robots*. Harper Collins, New York.
- MacDorman, K.F. and Ishiguro, H.** (2006). The uncanny advantage of using androids in cognitive and social science research. *Interaction Studies*, 7(3): 297-337.
- Mead, G.H.**, (1934). *Mind, Self, and Society*. Ed. C.W. Morris. University of Chicago Press, Chicago.

Reeves, B. and Nass C.I. (1996). *The Media Equation: How People Treat Computers, Television, and New Media like Real People and Places*. Stanford: CSLI Publications; New York: Cambridge University Press.

Nissenbaum, H., (1998). Protecting privacy in an information age: The problem of privacy in public. *Law and Philosophy*, 17: 559-596.

Smith, A., (1759). *The Theory of Moral Sentiments*. (ed. D.D. Raphael and A.L. Macfie) Clarendon Press, Oxford, 1976.

Tavani, H., (1999). Privacy On-Line. *Computers and Society*, 29(4): 11-19.

Van den Hoven, J., (2008) Information technology, privacy, and the protection of personal data. In J. van den Hoven and J. Weckert (eds.). *Information Technology and Moral Philosophy*. Cambridge: Cambridge University Press, 301-321.

Verbeek, P., (2005). *What Things Do*. Penn State University Press, Penn.

The ethics of ambient computing for personal health monitoring

Göran Collste*

Centre for Applied Ethics, Linköping University

Abstract

Health care is undergoing a transformation caused by ambient computing for personal health monitoring. There are a number of foreseeable advantages with this development but also some ethical problems. This paper focuses on the consequences for health care values. The new medical landscape can pose a threat to privacy, imply technological paternalism, have implications for personal identity formation and impair a sound patient-doctor relationship.

Keywords: ambient computing, personal health monitoring, ethics, privacy, identity, social constructivism

Introduction

If anything can be said to characterise the present age it is constant change. A main impetus for change is technology. We find ourselves continuously in new worlds that offer new possibilities but also put new demands on us. But are we just puppets on the technologically driven strings? Are there any ways for us to ethically assess, control and perhaps even shape the emerging technologies? In this paper I will discuss these questions with the backdrop of a specific, emerging and revolutionary technology: ambient computing for personal health monitoring. According to the visionaries this new technology will transform health care and move medicine from the hospital to the home.

The paper is outlined in the following way: First, I introduce the general problem of assessing emergent technologies from an ethical point of view. Secondly,

* Göran Collste, Professor of Applied Ethics, Linköping University, Sweden.

Collste's research and teaching deals with problems in ethics and applied ethics, i.e. the principle of human dignity, work ethics, global justice and ethical issues related to information and communication technologies. He is the author of *Is Human Life Special? Religious and Philosophical Perspectives on the Principle of Human Dignity* (Peter Lang, 2002) and he has also published books and articles on ICT – ethics and global justice. Collste is coordinator of the EU-supported Erasmus Mundus Masters Programme in Applied Ethics and partner in the FP 7 – project Personalised Health Monitoring - Ethics.

I give an overview of ambient computing and personal health monitoring. There are a number of possible benefits of personal health monitoring and home medical care. They provide for rapid response in case of emergency, patients are not tied to hospitals, patients may stay longer in their homes and distant monitoring is a potential for underprivileged regions to have access to medical expertise.

However, the new medical landscape also raises a number of ethical questions. Will the emerging technologies pose a threat to patient privacy, to personal autonomy, to the relation between doctor and patient and perhaps even to personal identity? In the third part I will discuss whether important values are threatened by ambient computing for personal health monitoring.

The ethical questions also draw attention to the conditions for ethical assessment of technology. At an early stage of development the technology might be possible to assess and influence but at that time the consequences are uncertain. At a later stage the consequences are better known, including the non-intended and unpredictable consequences, but then the technology is set and more difficult to influence.

Social Constructivism and Assessment of Emergent Technologies

Different methods for ethical assessment of technology have been developed in recent years (Reuzel, 2001, Schot, 2001). However, normally the objects for assessment are already existing technologies. But is it possible to ethically assess emergent technologies? While the technology is still in a developmental stage this seems to be quite difficult. One does not yet know what the technology will be like, how it will be designed, its function or its effects.

As is emphasised by social constructivism, technology is emerging in interrelation with the social environment. New technology shapes but is also shaped by the social system. Horner argues: "We always have before us the reality of choice in the sense that we could do or have done otherwise" (Horner, 2005, 226)

There are different lessons learnt by the theory of social constructivism. It is a reaction against deterministic views of technology. According to technological determinism the development of technology is autonomous and set. There is a demand to do something, say writing, or get something done, say cooling food. The new technology - or better - technical artefacts are produced in order to meet the demand. At the end of the process of construction the final artefacts, for example the computer or the refrigerator are there and apprehended as closed, which gives the impression that there was only one way to construct it. As Andrew Feenberg express this: "Looking back from the later standpoint, the artefact

appears purely technical, even inevitable. This is the source of the deterministic illusion” (Feenberg, 1999, p.11).

In contrast to technological determinism, social constructivists insist that new technology is the result of social interests, forces and choices. The theory has both descriptive and constructive implications. It informs us that technologies are not neutral but instead serve the interests of some institutions and social groups. However, the insights may also have constructive and normative implications. If we are aware of the fact that a technology is not set but instead that there is a possibility to shape new technologies according to our needs and values, technological development becomes an ethical challenge. The process of technological construction is intimately connected to questions of what is a good life and what values we want to realise.

Let me now turn from the more general reflections on the social construction of technology to a particular example of emergent technologies. As a case in point I will discuss ambient computing for personal health monitoring. This is a bundle of technologies used for medical home care that likely will transform both health care and home environment. Health care will move out of the hospitals and into private homes and the homes will be “invaded” by sensors and cameras. The technologies necessary for this transformation are more or less developed. It requires cameras and computers that are already long established technologies but also micro-sensors and chips that will be implanted in human bodies and that are in the process of development. Ambient computing for personal health care is thus a good example of an emerging and transforming technology that is still possible to influence.

Ambient Computing for Personal Health Monitoring

In the home

This paper focuses on one kind of emerging technologies; ambient computing for personal health monitoring. According to the visions of the technicians behind, sensors and computer devices will in the future practically be omnipresent. Small application-specific, network-connected information appliances will be embedded in virtually everything around us (Kunze et al, 2001). This new stage of information and communication technologies (ICT) is given different names; ambient technology, ubiquitous technology and pervasive technology.

According to different forecasts we can foresee that our homes in the future will be littered with monitoring devices. There will be microphones, cameras, micro-processors and sensors attached to the toilet, to kitchen devices, lockers e.tc. with the aim of aiding and facilitating life in different ways. Ambient ICT will also be used for health care purposes. Sensors may be wearable and even implanted into

a person's body. These sensors can monitor and register body temperature, heart rate, blood pressure or any other bodily health related function. Furthermore, the monitoring possibility also includes a person's movements, fall detection, location tracking and gastrointestinal telemetry. Thus, the new electronic sensors/devices can monitor a person's activities and many of her bodily functions.

The devices are embedded in the environment and should ideally be unnoticed by the users. How is this possible? First, they are extremely small. Nano-technology is expected to provide material which will be miniature sized. Hence, the devices needed for ambient computing in general and not least for personal health monitoring will be discreetly applied in the environment. "It is a world of smart dust", writes et al (Wright et al, 2008, p 1). Secondly, some devices can also be directly attached to the person, for example by being connected to a person's watch, ring, and clothes and even implanted in his or her body.

Ambient computing is a resource for different purposes but not least for providing health care in the home. Medical treatments and drug deliveries can be distributed in the home through "intelligent" devices, i.e. automated functions programmed by health care personnel, and monitoring devices make it possible to treat patients at a distance.

The medical use of ambient computing in combination with communication technologies is called "m-Health". It is defined as "mobile computing, medical sensors, and communication technologies for health care" (Istepanian, RSH; Jovanov, E.; Zhang, YT, 2004). Telematics, i.e. the combination of information technology and telecommunication is one of the requirements for this development. It is furthermore strengthened through wire-less communication. Body-area network (BAN) is used for communication between sensors and the patient's body, personal area network (PAN) for communication in the personal environment of the patient and wide area network (WAN) for the connection to a central data pool and information services (Kunze, et al, 2001).

In sum, according to Nehmer, et al., in m-Health the following characteristics of ambient computing are used for health care purposes. Ambient computing is

- invisible, i.e. embedded in clothes, watches, glasses, e.tc.
- mobile, i.e. being carried around,
- context aware, i.e. equipped with sensors and wireless communication interfaces making it possible to scan the local environment for useful information and spontaneously exchange information with similar nodes in their neighbourhood,
- anticipatory, i.e. acting on their own behalf without explicit request from a user,

- communicating naturally with potential users by voice and gestures instead by key board, mouse and text on a screen,
- adaptive, i.e. capable of reacting to all kinds of abnormal exceptional situations in a flexible way without disruption of their service” (Nehmer, et al., 2006, p.46)

So, ambient computing for personal health monitoring implies that sensors and monitoring devices will be placed in the patient's homes. These sensors and monitors are connected to clinics and hospitals on the other end.

In the clinic – medical connectivity

So far, I have described how sensors and monitoring devices combined with wireless connections make it possible to send many different data about a patient to the health clinic. On the receiving end there are personnel at a clinic watching monitors, collecting information on computer screens and noticing signals. With the help of microphones, cameras and other communications devices, health care personnel are able to react to the incoming information and communicate with the patient.

On the basis of all the information received, health care personnel will perform their work. Physicians will make their diagnosis and suggest therapies, nurses will plan for patient care, occupational therapists will prepare necessary aid and social workers will analyse the patient's future social needs. And, in case of emergency, for example if the patient has fallen, an ambulance will in a short time be on its way.

Thus, it could seem as if ambient technologies for personal health care offer an ideal health care. The latest technology is used. There are human experts and expert systems that can diagnose any kind of incoming data and suggest treatments. Through telemonitoring health care professionals and patients can communicate with each other. The patient does not have to travel but can stay at home. He or she is observed and taken care of in a way equal to that of being at a hospital. Although the new systems for health care have a great potential there are some possible ethical problems connected to it that I will discuss in this paper.

What is the point?

Ambient computing is an instrument for prevalent personal health monitoring and distributed health support. It is envisioned as the future health care, and will, according to the “vision” of the European Commission “...take healthcare out of the hospital, bring it to the home and embed it into people's lives” (EurActiv, 16.6.2008).

It is not difficult to see a number of benefits with the emergent technologies. There are at least five reasons for introducing ambient computing for personal health monitoring. First, the system has a potential for monitoring patients suffering from chronic illnesses, such as diabetes, and elderly people. This is a category of patients that needs long term attention but they do not necessarily need continuous treatment. In the present health care system, when patients are too ill to live at home, but not ill enough to go to hospital, they normally move to institutions for chronically ill patients or old folk's homes. With the new possibilities for personal health monitoring it will no longer be necessary for them to go to hospitals or old folk's homes, or at least it will be postponed. Instead, the patients will be able to stay longer in their homes and they will still be monitored and get the care they need.

The benefits for chronically ill patients and the elderly who are able to live longer in their homes are both personal and economic. Both chronically ill patients and elderly typically want to stay in their homes as long as they can. They are familiar with the environment and they have neighbours and friends close by. Furthermore, the cost of institutional care is multiple compared to home care. Hence, ambient computing for personal health monitoring promises to benefit both the patients and the society.

Secondly, in cases of emergency and alarm personal health monitoring will facilitate and speed up necessary relief actions. For example, if an elderly falls in his or her home this will be registered by the monitoring devices. Thus, through cameras, sensors and other monitoring devices information about emergencies will immediately be transferred to the emergency units. The health care team can prepare for appropriate care and the patient will be more promptly taken care of.

Thirdly, the possibility to get expert opinions, for example medical advice, at a distance will have various beneficial consequences. There might be a need for immediate medical consultation when an accident happens far from a hospital as well as for a scientific expedition in distant regions or in space. Distant medical expertise for diagnostics and therapy might also be helpful for medical service in underprivileged regions and countries. The distribution of health care resources is usually uneven within a country and even more between developed and developing nations around the globe. The possibility of telemonitoring and distant care has the potential for limiting the gap in access to medical care.

Fourthly, personal health monitoring is not only a potential resource for the ill and elderly but also for healthy people. It might function as an early warning system for a variety of medical conditions. The monitoring system can indicate possible health-related problems even before they are noticed by the person him/herself.

Continuous health monitoring might also promote healthy life styles. It can indicate when the person behaves in a way that is detrimental to his or her health.

Fifth, in a time when there is a lack of resources for employing health care personnel, different kinds of technical support may compensate. Nehmer et al. write: "Autonomy enhancement services...make it possible to abandon previous manual care given by medical and social care personnel or relatives, and replace it by appropriate system support" (Nehmer, 2006, p.44). Hence, ambient computing for personal health monitoring may in the future replace health care workers and relatives.

There are thus a number of reasons for developing ambient computing for personal health monitoring. But what are the ethical implications?

Ethical Questions

Technology is often perceived as something given and set, as the result of invisible forces outside of anyone's control. The point of social constructivism is to "unveil" technology. It stresses that technology is the result of human decision making, social forces and interests, and hence, that technology is changeable. We can ask questions about technology: whose interests are served by technology, what is good technology and how would we like to change technology. These questions invite us to engage in ethical assessment of technology. We can assess the design, result and consequences of already existent technologies in order to improve them. We can also assess emergent technologies, although the assessment is more difficult to do. Regarding emergent technologies we do not yet know the consequences and the technological development can take different paths. On the other hand, the fact that emergent technologies are in a process of development makes the assessment even more pertinent. The ethical assessment might have a real – and presumable positive – effect.

We have noticed that personal health monitoring has a number of advantages. The patients can stay in their homes instead of having to visit a clinic or be admitted to hospitals, situations of emergency will be detected faster and medical diagnosis, advice and treatment can be provided at a distance. Besides the issue of more efficient use of societal resources, these consequences are benefitting the patients and contribute to a better health care. They can, from an ethical point of view, be motivated by both the principle of beneficence and the principle of autonomy. Personal health monitoring has in different ways a potential to benefit the patients and it may contribute to their self-determination. Thus, there seem to be many good reasons to introduce personal health monitoring. The emerging ambient technologies will very likely transform health care as well as the home environment.

However, in spite of the potential benefits of ambient computing for personal health monitoring, we can also envisage some threats and vulnerabilities connected to the emergent technology. Scientific commissions in the area have pointed out that there is also a dark side to a “world of ambient intelligence”. Increased surveillance of public space, in work and in the homes pose a threat to privacy. Furthermore, vital societal practices, infrastructure and management of personal data are dependent on the new technology. Therefore, security of the systems is of utmost importance. If - or rather when - the systems break down, it will cause damages to both the society and to individuals. For example, personal data run the risk of coming in the wrong hands and there are risks of identity thefts. According to the authors of *Safeguards in a World of Ambient Intelligence*, loss of control will also undermine public trust in the new systems (Wright, 2008). But what are the risks connected to future health care dependent on ambient computing? Are there any ethical problems related to the new medical landscape? In the next sections I will discuss possible ethical problems related to the advent of ambient computing for personal health monitoring.

Let me take a technical device as a case in point. Let us assume that a nano-scaled device is implanted in a patient’s body. The purpose is to monitor heartbeat and other bodily functions. We can envisage a number of ethical problems related to this new medical device. First, it may pose a threat to privacy. A right to privacy presupposes both a right to non-intrusion and a right to control information about oneself. How will privacy be affected by the fact that sensitive information is circulated in decentralised IT-systems, and how can privacy be protected with monitoring of patients in their homes? Secondly, ambient computing can be invisible and programmed to anticipate human action. For example, the implanted device might restrict the functioning of the patient in a way that was unanticipated and of which the patient was unaware. Does ambient computing imply a risk for technological paternalism? Thirdly, a motive behind both home-based medical care and telemonitoring is to replace health care personnel with technical devices. For example, the implanted device will automatically transfer information to the health clinic which will decrease the need for personal interaction between health care personnel and the patient. But how will the limitation of personal encounter change the relational aspects of health care? Is it a moral issue if computerised health care replaces human relations? Fourthly, how will wearable or implanted monitoring devices affect the identity and integrity of the patient? Will it imply a “medicalisation” of the personal identity so that the person’s self will be transformed into a patient’s self? Thus, it seems as if even existential problems of personal identity and integrity are raised by the emerging health care technologies. Now, let me discuss these possible ethical problems in more detail.

Personal health monitoring and privacy

The risk of privacy violations in the wake of ambient computing is stressed by the writers of the aforementioned book *Safeguards in a World of Ambient Intelligence* (2008). They warn about the threats of different facets of privacy invasions such as disclosure of personal data, surveillance and risks from personalised profiling. They also cite a report from the American National Academy of Science stating that due to a future scenario where monitoring devices might be embedded everywhere, privacy "...may be at greater risk than at any previous time in history" (ibid, p. 19).

How will ambient computing influence the home environment? What are the consequences for privacy if different kinds of monitoring devices are installed in the homes? "My home is my castle" is a well-known saying. Our homes are the place where we relax, are left alone and feel secure. We meet with whoever we like, say what we want and behave without anyone watching us. Will this still be possible with personal health monitoring? One can fear that the new technology will have negative implications for privacy. But in what way and what does privacy really entail? Why is privacy something that we should care about? These questions have been widely discussed in the last few decades due to increased surveillance and storing of information in the wake of the information society. Many ethicists have tried to find a more precise meaning of privacy and developed a normative underpinning of a right to privacy. In the following I will first – against the backdrop of ambient computing for personal health monitoring – try to demarcate the concept of privacy and, secondly, argue for the view that privacy is an important value; that it make sense to see privacy as a moral right. Thirdly, I will try to answer the questions if and how ambient computing and personal health monitoring will affect privacy and finally discuss whether there are possibilities to protect privacy in the new world of ubiquitous technology.

Privacy is an important concept in both law and ethics. Here I will discuss it from an ethical point of view. The exact meaning of privacy is a disputed topic. Philosopher Deborah Johnson argues that "...privacy is a complex and, in many respects, elusive concept" (Johnson, 2001, p. 120). However, in ethical discussions privacy generally has two meanings: (1) to be left alone and (2) to control information about oneself. Ambient computing for personal health monitoring seems to have implications for both these aspects of privacy. It will – as explained - imply that persons/patients have monitoring devices right in their homes and even in their own bodies and it will – as a consequence – mean that a lot of sensitive information about patients is transferred from the home to the clinic.

Is then ambient computing for personal health monitoring a potential threat to privacy? First, is it a threat to a patient's control of his or her private sphere?

Obviously, ambient computing implies that a person's private sphere is affected. Different kinds of monitoring devices are installed in the patient's home; hence the private space runs the risk of disappearing. Personal health monitoring makes use of sensors that perceive aspects of the environment. They may for example sense emotions like stress and excitement. Furthermore, ambient computing is omnipresent and invisible. This means that the user is not always aware of the monitoring that is going on. In this way, ambient computing seems to be a real threat to privacy in the sense of control of one's private sphere.

Even informational privacy is affected by personal health monitoring. There is always a risk of leakage when sensitive information is transferred, as in the case of personal health monitoring when health-related information is transferred to health clinics. Philip Brey writes:

The privacy risks of AMI/Ambient Intelligence are so great because of the often highly sensitive types of personal information that are recorded and encoded, the scale on which this information is recorded, and the ease with which it could potentially be disseminated and made available to other parties (Brey, 2005, p.164).

But does that mean that personal health monitoring is a threat to the patient's right to privacy? If a right to privacy means a right to decide who and under what conditions other persons can enter his or her private sphere, the right to privacy is violated if the monitoring is done without the consent of the patient. However, the presupposition behind personal health monitoring is that the patient has given his or her informed consent to the placement of computational and monitoring devices in the home and it is assumed that it is in the best interest of the patient. Furthermore, it is also assumed that the monitoring devices could be switched off when it suits the patient. Hence, if these requirements of informed consent are fulfilled, the patient's right to privacy is not violated while he or she is in control of his/her private sphere.

However, the patient is in control of his or her private sphere only if the patient has a real choice to say no to personal health monitoring. Let us imagine two scenarios: in the first, the patient can choose between personal health monitoring, with the benefits pointed out above, and traditional health care based in the clinic. In the second, the patient has a choice between on the one hand accepting personal health monitoring and on the other hand, if not accepting personal health monitoring, receiving very limited health care. In the first scenario the patient has a choice and is still in control. The patient may be aware of the benefits of personal health monitoring but he or she nevertheless choose traditional health care in order to avoid possible privacy intrusions. But in the second scenario, the patient's freedom is limited. In this case, if personal health monitoring is the only

option to get decent health care, the cost of saying no is so great that as a matter of fact, the patient has lost control and the new system implies a violation of patients' right to privacy.

Does then ambient computing and personal health monitoring pose a threat to the control of information about oneself? There will be a constant flow of medical information from the home to the clinic. The information is of a kind that is sensitive from a privacy point of view. Of course, the transfer of patient information in health care is nothing new. However, the amount of information and the potential numbers of receivers are aggravated through the use of information and communication technologies. It is safe to assume that the more sensitive the information that is transferred, the greater the risk for privacy violations. And there are those who have an interest in getting access to medical information. First we have the so called "trusted insiders", i.e. personnel who have legitimate access to medical information but who might use it in an inappropriate way. Then there are employers and insurance companies who might want access to patients' medical information (deCew, 1999).

Of course, even though system developers and designers have the best of intentions the possibility of abuse should also be considered. There is always a possibility that the wrong people will get access to sensitive information. Although this kind of risks is inevitable in any human practice, the awareness of possible misuse will raise the level of risk analysis. Even if the threat of information leakages, hacking and other kind of abuses should not be overvalued, personal health care information is of less interest for hackers than for example sensitive military information or information about company secrets, the threat to informational privacy should not be underestimated.

One way to avoid that personal health monitoring leads to privacy violations is to develop privacy-enhancing technology. This is an instance of so called "value-laden design" (Nissenbaum, 2000). When it is recognised that values are embedded in the technology already at the design stage, it is also possible to design technology so that for example the protection of privacy is taken into consideration. The aim of privacy-enhancing technologies is to minimise exposure of private data, through privacy protection (anonymiser and encryption tools) and technology for privacy management ensuring confidentiality (Ahonen, 2008, 158, 181). Another way to limit potential privacy violations and to safeguard that privacy protection is taken into account is to engage stakeholders in the design and implementation of systems for personal health monitoring.

This section has focused on the possible threats to privacy caused by ambient computing for personal health monitoring. Monitoring is installed in private homes which pose a threat to the private sphere and the systems transfer great

amount of sensitive information which pose a threat to informational privacy. Further, the right to privacy might be endangered if the patient has no other option than to accept the new technological systems. Privacy-enhancing technologies and a raised awareness among everyone involved in the development and use of the systems are ways to limit the threats.

Even if patients are aware of the installation of personal health monitoring, it might be going on in everyday life unnoticed. This may even increase the threat to privacy while the users may forget about the monitoring and live their private life as usual. However, the invisibility of ambient computing is not only a threat to privacy but also to the possibility to control technology. I will now turn to that problem.

A risk of technological paternalism?

We saw earlier that ambient computing in the homes will be embedded and also invisible. It is unnoticed but facilitates daily life in different ways. How can we choose and stay in control of a technology that is embedded and invisible?

In a seminal article from 1991, Martin Weiser forecast the advent of ambient or, as he calls it, “ubiquitous” computing. In his at that time futuristic scenario he writes: “We are therefore trying to conceive of a new way of thinking about computers, one that takes into account the human world and allows the computers themselves to vanish into the background” (Weiser, 1991, p.3) How will this be possible? According to Weiser, two aspects are crucial; location and scale. The computer devices must be placed unnoticed and so small that they practically disappear and come to be part of common awareness. Hence, location and small scale will make the new wave of computing “invisibility”. Weiser envisage a future where computers are invisible and when people use them unconsciously to fulfil their everyday tasks. (Weiser, 1991).

Weiser’s forecast seems to be realised today and ambient computing for personal health monitoring is one kind of application. Medical sampling, testing and monitoring, as well as diagnostics and therapy are delegated to the computerised systems. Does this imply a limitation of patient’s freedom? Are they targets of a new authoritarian regime, victims of paternalism?

The paternalistic threat of ambient computing is discussed by S. Spiekerman and F. Pallas. Paternalism implies that an agent, for example a doctor or some other authority, performs actions that control someone else, a patient, but in his or her best interest. When this action is performed against the will of the patient we have a case of strong paternalism, when it is performed in line with the will of the patient we talk of weak paternalism.

Now, Spiekerman and Pallas transfer the concept of paternalism to technological practice and coin the concept “technological paternalism”. They define technology paternalism in the following way:

Given a technology T controlled by a patron A that performs an action X which is affecting a subject B directly, X is paternalistic if and only if:

- X is perceived by subject B as limiting, punishing or in any other way cutting down on freedom
- X cannot be overruled or in any other way disregarded without sacrificing functionality
- X is claimed by patron A to be mainly in B’s own interest
- X is performed autonomously (Spiekerman and Pallas, 2006, p. 10-11).

The definition assumes first that the subject perceives that his or her freedom is being limited. However, this is due to the functionality of the technology at hand, hence without this limitation of freedom the subject would be without the supposedly beneficial technology. What is unique about ambient technology paternalism is that it is invisible and pervasive. The paternalistic technology is acting in a way that the users seldom have reason to question. Let me as a possible example of technology paternalism take one application of ambient technology mentioned above. The aim of a smart drug – system, in the form of a sensor of the pillbox connected to the mobile sensors monitor the correct drug intakes, was to enhance the control of drug addiction. Is this an example of technology paternalism? Yes, it may seem so. When the system is in practice it might very well be perceived as a limitation of the users’ freedom. Furthermore, it cannot be overruled without sacrificing functionality, it is in the users’ best interest and it is performed autonomously.

Is, then, this case of technological paternalism acceptable from a moral point of view? This question is analogous to the question whether paternalism in health care is justifiable. Arguments in favour of health care paternalism say that while it is the duty of doctors and nurses to act in the a way that is beneficial to the patient, they have this duty even when the patient is unable to agree, or when the patient due to lacking capacity is unable to take a stand. Arguments against paternalism emphasise the importance of autonomous decision making. According to this view health care paternalism is a violation of the principle of respect for persons. It seems that the same arguments are relevant for and against technological paternalism. As Spiekerman and Pallas rightly concede, there is a potential conflict between anti-paternalism and ambient computing. They write, “There is a clear disaccord between the concept of disappearing technologies and the attempt to remain in control” (ibid, p.12).

One can envisage a future development of ambient computing where the pervasiveness and invisibility of the technology may make the problem of technological paternalism more acute. The users are becoming less and less aware of the applications, for example monitoring or smart devices for health information that are installed in their best interests. These technologies function automatically and while the patients are unaware of their existence they lack the freedom to accept or not accept them.

Ambient computing for personal health monitoring has a potential for paternalism. But it also has a potential to affect the way we perceive ourselves. This question will be discussed in the next section.

Implications for identity

Ambient computing for personal health monitoring implies that sensors and other monitoring devices are wearable, for example in rings or clothes, or implanted in a person's body. Hence, the devices will be situated very close to, or even be part of, a person's body. How will this closeness influence a person's life and well-being?

In this section, I will discuss some possible implications of personal health monitoring for a person's sense of identity and for her life-world. Let me start with an example of another device. Most of us living in the affluent world have a wrist-watch. It is convenient to be able to check what time it is and the watch is a necessary means for showing up in time for appointments. The watch is intimately integrated in our work and everyday life. In order to sense the contrast, some people put their watch in a drawer when they are on holiday. Then there is less need to keep track on time and to be without a watch gives them a sense of freedom.

The example illustrates that the watch is something more than just a practical device. The watch has implications for our sense of time and for our sense of freedom and dependency. It makes us more conscious of time and it structures and even controls our day. Thus the watch has not only a practical function but also a symbolic.

The wrist-watch is an example of a known device that we carry close to our bodies. I imagine that this example also tells us something about the more far-reaching implications of ambient computing. How, then, will ambient computing and personal health monitoring influence our sense of ourselves and of reality? Will it have implications for our identity?

"Who am I?" This is the basic question of identity. In answering the question, we may refer to our gender: I am a man, our nationality: I am a Swede; our family life: I am a father; our work: I am a university professor; our belief: I am a Christian e.tc. Our different identities are of different weight for us. For some people

their work is most important, for others their family position e.tc. Our identities also have different weight over time. In one period of life being a father is in the foreground, while in a later period the professional identity becomes more important. Identities are also situated. When I attend a church service, my religious identity becomes dominant, and when I visit a hospital due to some illness, my identity as a patient prevails.

What, then, will it mean for our identities to have sensors and monitoring devices attached to our clothes or in our bodies with the aim of measuring health related bodily functions? Even if the devices are disguised and hardly noticeable, the bearer is well aware of their presence. Their presence is a constant reminder of the fact that the person suffers from some kind of chronic disease. Consequently, the person wearing the devices is constantly in a role as a patient. We can conjecture that this will affect his or her sense of identity. Who am I? I am foremost a patient, might be the natural answer.

To be a patient (from Latin *patientia*, meaning “to suffer hardship”) is to be monitored and dependent on caring and treatment. It is to be in a state of worry and hardship; will I recover or will I get worse? A possible effect of ambient computing for personal health monitoring is that the personal identity is being “medicalised”. A person’s self will be transformed into a patient’s self. If this happens, the emergent technologies affect our identity as well as our integrity. An integrated self is a self where multiple identities are balanced. In contrast, in a disintegrated self, a specific role or identity is dominant, not only temporarily but constantly. According to my analysis of the implications of ambient computing for our identities, this is what might happen as a consequence of the new technology. A person is becoming a patient. The identity of being a patient takes over other identities.

The new technologies may not only be used for monitoring sick people but also for monitoring healthy. This has been suggested as a means for promoting healthy life styles (Istepanian, RSH; Jovanov, E.; Zhang, YT, 2004). Then the monitoring devices are not there to detect possible malfunctions in chronically ill patients, but to report the implications for health of a person’s life style. If a person smoke or drink too much this will be noticed, as well as if a person fails to take part in some health promoting activity. A similar kind of identity transformation as the one from person to patient might then also occur as a consequence of continuous personal health monitoring of healthy people. Even here one might ask whether this would imply a “medicalisation” of a person, who as a matter of fact is healthy. Perhaps a constant reminder of our health conditions does not improve our health?

We noticed in the opening of this article that new technologies will have both anticipated and unanticipated consequences. In this section I have pointed at some possible unanticipated consequences of personal health monitoring. It might lead to a transformation of a monitored person's self. It might imply a change in self-perception when the identity as a patient prevails over other identities. If this happens, the emergent technology has far-reaching psychological and existential implications.

The vision of ambient computing for health monitoring is to move health care into the patient's home. One argument for this transformation is to save money by substituting health care personnel with technical devices. How will this transformation affect the social aspects of health care? This question will be discussed in the next section.

How will personal health monitoring influence the patient-doctor relationship?

Ethical aspects of the clinical encounter or consultation has for many years been an issue for discussions in medical ethics (Ramsey 1970, Pellegrino & Thomasma, 1981, Beauchamp & Childress 2001, Svenaeus, 1999). The patient is in a vulnerable situation when his or her health or life is threatened and the clinical encounter is a means to recovery and/or of caring, with the doctor as a mediator (Pellegrino & Thomasma, 1981). Medicine is an "interpretive meeting" between the patient and the care giver with the aim of understanding and healing the patient who seeks help (Svenaeus, 1999, p.28). This seemingly simple meeting can be seen as a complicated encounter, with scientific, emotional and normative content.

How, then, do emergent technologies used for personal health monitoring influence the patient-doctor relationship? I will discuss the question from two angles. First, an argument for personal health monitoring is that technology will substitute health care personnel. For example, nurses and assistants will be redundant with the presence of monitoring and "smart" devices. This, in its turn – it is argued – will save money for health care. In a time of prioritisation of resources such a gain is an important impetus for developing the new technology.

There is some evidence that remote communication via telecare is considered as beneficial by patients and care providers (Sävenstedt, 2004). On the other hand, those investigations originate from a situation when the geographic distances render regular contacts between patient and health care personnel difficult. Telecare is in this case not replacing personnel but instead adding to the possibility of getting care. Still, one might ask, considering the importance of the doctor-patient relationship, how will drainage of personnel affect the quality of care? If

we can assume that the caring relation is crucial for medicine, the answer seems to be obvious: less caring personnel will have detrimental effects on the quality of health care. Hence, those introducing information and communication technologies for personal health monitoring should consider the words of Marsden S Blois regarding an earlier stage of medical informatics. When reflecting over the apparent distrust and disinterest of doctors towards information technology in health care, he comments:

The most important question appears not to be 'Where can we use computers?' but 'Where must we use human beings?' Until this matter is thoroughly explored, tension between physicians and computer advocates will persist (Blois, 1980).

Secondly, a technology-driven implementation of personal health monitoring runs the risk of promoting a one-sided engineering perspective on the clinical encounter. In such a scenario, the relationship between patient and doctor will be of less importance. But, why should one expect such a development and why is it an ethical issue?

In his historical survey of the "reign of technology" of medicine, historian Stanley Reiser investigates the consequences of an increasingly technological driven health care. Before the advent of modern medicine in the nineteenth century the doctor had to rely on two sources of information: one source was to listen to the patient's narratives and another through a direct connection to the patients' bodies carrying out physical examinations. Needless to say, this practice presupposed a dialogue with the patient. The introduction of medical technology puts the patient in the background. And, as Reiser argues, something important was lost. He writes:

...the machines are denied complete access to a whole range of non-measurable facts about human being that a physician can only obtain through his own senses – questioning, observing and making judgements (Reiser, 1978, p.229)

Reiser's conclusion after an historical survey of how new technologies were introduced in medicine is that in spite of all the advantages of new medical technology, an increasing dependence on technology has in some respects been detrimental to health care quality and the values of health care.

Reiser points at two problems with a reign of technology over medicine: first it might create a distance between doctor and patient detrimental to the principle of patient-doctor relationship. This will not only harm the personal and psychological aspects of the clinical encounter, but also limit the access to the personal information the doctor needs in order to make an adequate diagnosis.

Secondly, it might also have detrimental effects on the doctor's own diagnostic capacity. Diagnostic expertise, like all kind of expertise, is developed and maintained through practice. Without a continuous contact with real patients, e.g. unique persons with unique symptoms, the - often - tacit knowledge of how to diagnose a patient's disease might get lost.

Obviously, a beneficial relationship between doctor and patients is not necessarily lost in a high tech medical environment. That is neither Reiser's view. There is no predetermined path to distance and alienation. Instead, if developed and used in a way that takes the basic values of medicine into account new technology can both be helpful and contribute to a good health care. However, such an alternative way of implementing emergent technologies need a conscious effort to assess it from a moral point of view.

A Paradox: From Visibility to Disappearance

The vision of an ethically informed technology development presupposes visibility. In line with the theory of social constructivism I have argued that technology is the result of a chain of decisions taken by product designers and producers. In order to assess technology, questions like: What choices are made? at what moment? by whom? should be raised.

However, the visibility of a new technology and the public and moral discussion of its consequences tend to slowly vanish. This will happen when the technology is integrated in our everyday life and it becomes commonplace. Then, we no longer reflect on the pros and cons of the technology, we just take it for granted. As Mark Weiser argues, the disappearance of technology "...is a fundamental consequence not of technology but of human psychology" (Weiser, 1991, p.3). When we learn something well enough we cease to be aware of it.

Philosopher Deborah Johnson sees technology as "instrumentation of human action" (Johnson, 2001, p. 27). Ethics is about human action and, hence, when technology is set it is the human action instrumented by technology and not the technology as such that is ethically assessed. As a consequence, Johnson argues that a focus on the ethics of a particular technology, for example computer ethics, will also slowly disappear. Johnson writes: "...once the new instrumentation is incorporated into ethical thinking, it becomes the presumed background condition" (ibid, p. 30).

So, we can conclude that it is of utmost importance to focus on the process of design and development of new technology from an ethical perspective. The reason is, of course, that technology is influencing our world and our lives. However, when a technology is set and become integrated as instrumentation of human ac-

tion, it will disappear and the ethical discussion will move to another new technology.

Conclusions

This article deals with ethical problems and ethical assessment of ambient computing for personal health monitoring. We have noticed that one can foresee that the emerging technologies will benefit health care in different ways. It will facilitate for elderly and chronic ill patients to stay in their homes instead of being forced to go to hospitals or caring centres. It will speed up relief in case of emergency, it will facilitate the possibility to get expert medical opinion at a distance and it has a potential to lower the costs of health care.

However, the transformation of health care due to ambient computing for personal health monitoring can also be expected to have some challenging ethical consequences. It might affect privacy in two ways. Ambient computing invades the patient's private sphere and even – ultimately - his or her body which means that almost everything, except perhaps his or her thoughts and hopes, might be monitored. Secondly, even informational privacy might be threatened due to the transfer of sensitive information about the patient.

The introduction of ambient technology for personal health monitoring is of course done with the best of intentions. However, the combination of invisibility and pervasiveness might create difficulties for the patient to control his or her environment. Hence, the emerging technology has potential paternalistic implications; the patient loses control and her autonomy is in this way confined.

We have, thirdly, noticed that personal health monitoring might influence a person's sense of identity. He or she might comprehend him/herself more and more as a patient. In the footsteps of the emerging technology one might find an insidious medicalisation of identity.

Finally, as a consequence of ambient technology for personal health monitoring the distance between doctor and patient might increase. There will be fewer reasons for personal encounters and more of distant monitoring. Health care runs the risk of being ruled by technology.

These four possible consequences are examples of unintended but ethically relevant implications of emerging technologies in health care. Can they be avoided? Are there any options for alternative strategies? Can health care benefit from the emerging technologies without threatening important health care values? The answer to these questions depends on many factors. The problematic effects of ambient computing for personal health monitoring are not predetermined. Through a continuous, constructive, imaginative, interactive and ethically informed technolo-

gy assessment involving different stakeholders they might be avoided. Then emerging technologies could be designed so that their potentialities are realised but their negative impacts will be avoided. If this happens, in the long run ambient computing for personal health monitoring will become helpful instrumentation of health care and embedded in our everyday life.

REFERENCES

- Axisa, F., Dittmar, A. & Delhomme, G.,** (2003) Smart clothes for the monitoring in real time and conditions of physiological, emotional and sensorial reactions of human, Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE.
- Beauchamp, T.L. & Childress, J.F.,** (2001) Principles of Biomedical Ethics, 5. ed., Oxford University Press, New York.
- Beaudin, J.S., Intille, S.S. & Morris, M.E.,** (2006) To Track or Not to Track: User Reactions to Concepts in Longitudinal Health Monitoring, Journal of Medical Internet Research, vol. 8, no. 4.
- Blois, M. S.,** (1980) Clinical Judgment and Computers, North England Journal of Medicine, 303, 1980, p. 192-197
- Buber, M.,** (1923) Ich und Du. Insel-Vlg, Leipzig.
- Collingridge, D.,** (1980) The Social Control of Technology, Frances Pinter, London.
- Collste, G., Shahsavar, N. & Gill, H.,** (1999) A decision support system for diabetes care: Ethical aspects, Methods of information in medicine, vol. 38, no. 4-5, pp. 313-316.
- Daniels, N.,** (1996) Justice and justification: reflective equilibrium in theory and practice, Cambridge University Press, Cambridge.
- DeCew, J.W.,** (1999) Alternatives for protecting privacy while respecting patient care and public health needs, Ethics and Information Technology, vol. 1, no. 4, pp. 249-255.
- Dieng, D.,** (2001) The Hemacard Project: Applying the Constructive Technology Assessment Method to Computerised Health Cards, in Technology and ethics: a European quest for responsible engineering, Peeters, Leuven.
- Horner D.S.,** (2005) Anticipating ethical challenges: is there a coming era of nanotechnology? in Ethics of New Information Technology, ed. Brey, P, Grodzinsky, F, Introna, L, CEPTEs, Univ-. of Enschede, pp. 217-228.

Inness, J.C., (1992) *Privacy, Intimacy and Isolation*, Oxford U. P. (N. Y.), New York.

Johnson, D.G., (2001) *Computer ethics*. 3. ed. Upper Saddle River, N.J.: Prentice Hall.

Korhonen, I., Parkka, J. & Van Gils, M., (2003) Health monitoring in the home of the future, *Engineering in Medicine and Biology Magazine, IEEE*, vol. 22, no. 3, pp. 66-73.

Kunze, C., Grossmann, U., Stork, W. & Muller-Glaser, K., (2001) Application of ubiquitous computing in personal health monitoring systems, *Bio-medizinische Technik*, vol. 47, pp. 360-362.

Milenkovic, A., Otto, C. & Jovanov, E., (2006) Wireless sensor networks for personal health monitoring: Issues and an implementation, *Computer Communications*, vol. 29, no. 13-14, pp. 2521-2533.

Miller, R.A. & Masarie, F.E., Jr., (1990) The demise of the "Greek Oracle" model for medical diagnostic systems, *Methods of information in medicine*, vol. 29, no. 1, pp. 1-2.

Nehmer, J., Becker, M., Karshmer, A. & Lamm, R., (2006) Living assistance systems: an ambient intelligence approach, *Proceedings of the 28th international conference on Software engineering*, ACM New York, NY. USA,.

Nissenbaum, H., (2000) *Values in Computer System Design: Bias and Autonomy*, *Ethics in the Age of Information Technology*, ed. Collste, G, Linköping, pp. 59-69.

Pellegrino, E.D. & Thomasma, D.C., (1981) *A philosophical basis of medical practice: toward a philosophy and ethic of the healing professions*, Oxford U.P., New York.

Ramsey, P., (1970) *The patient as person: explorations in medical ethics*, Yale U-P., New Haven.

Reiser, S.J., (1978) *Medicine and the reign of technology*, Cambridge U.P., Cambridge.

Sollie, P., (2007) Ethics, technology development and uncertainty: an outline for any future ethics of technology, *Journal of Information, Communication, and Ethic in Society*, vol. 5, no. 4, pp. 293.

Spiekermann, S. & Pallas, F., (2006) Technology paternalism—wider implications of ubiquitous computing, *Poiesis & Praxis: International Journal of Technology Assessment and Ethics of Science*, vol. 4, no. 1, pp. 6-18.

Svenaesus, F., (1999) The hermeneutics of medicine and the phenomenology of health: steps towards a philosophy of medical practice, Tema, Univ., Linköping.

Sävenstedt, S., (2004) Telecare of Frail Elderly: Reflections and Experiences Among Health Personnel and Family Members, Umeå University.

Tavani, H.T., (2007) Philosophical theories of privacy: Implications for an adequate online privacy policy, *Metaphilosophy*, vol. 38, no. 1, pp. 1-22.

Wright, D. Gutwirth, S. Friedewald, M. Vildjiounaite, E. Punie, Y., (Eds.) (2008) Safeguards in the World of Ambient Intelligence, The International Library of Ethics, Law and Technology, Vol. 1, Springer.

Internet sources

www.healthyaims.org (accessed February, 20, 2009)

<http://www.intellidrug.org/> (accessed February, 20, 2009)

<http://www.fp6-minami.org> (accessed February, 20, 2009)

<http://www.euractiv.com/en> (accessed February, 20, 2009)

Trusting Invisible Strangers in Open Source Communities: About the Assumption, Inference and Substitution of Trust

Paul B. de Laat*

Faculty of Philosophy-University of Groningen

Abstract

Open source communities squarely rely on the contributions of invisible strangers in cyberspace. How do these communities handle the problem of trusting them to have good intentions and adequate competences? This question is explored for the peer production of software (FreeBSD and Mozilla in particular) and encyclopedic entries (Wikipedia in particular). It is argued that in the informal phase trust is mainly assumed: the open approach empowers potential contributors and makes them step forward. Additional assurances can be inferred from an underlying ethic that is either already there or has to be created on the spot. In the formal phase rules and regulations are introduced. Here projects face a design choice, in particular as concerns the continuum between a high-discretion design (that grants a large amount of trust to contributors) and a low-discretion design (that substitutes some amount of trust and accordingly grants only a small amount of trust to contributors). Surprisingly, open source designs for software and encyclopedias are found to be converging, towards a mid-level of discretion – the anonymous user is no longer held in absolute trust.

Keywords: design, FreeBSD, hacker ethic, Mozilla, open source software, strangers, trust, Wikipedia

Introduction

Since the advent of the Internet open source communities have been growing in participants and spreading from software to various other kinds of content. Not only software can be produced in open source fashion, also encyclopedias, mov-

* *Paul B. de Laat* is Assistant Professor of Philosophy of Science, Technology & Society, at the Faculty of Philosophy, University of Groningen (Groningen, The Netherlands). He obtained a Masters Degree in theoretical physics (University of Utrecht) and a PhD in organizational sociology (University of Amsterdam). His current research focuses upon governance of open source software, commons for informational resources and associated IPRs, and the concept of trust in cyberspace.

ies, biology and more. I will rely on the following standard definition of open source: peers producing knowledge together as volunteers, without being directed by markets or managerial hierarchy, and posting their ever evolving products in a readily accessible format in a virtual 'commons'. We observe social movements here that develop knowledge both 'by the people' and 'for the people'.

The most amazing aspect of such a mode of production has always been the openness to the outside world. The gates are open to everybody, no questions asked. With software, one is invited to submit comments, code patches or new features; with encyclopedias, one is invited to change existing entries or suggest new ones. This open approach must even be considered the very basis of its success.

It has to be borne in mind that anonymous strangers on the Internet are the perfection of strangers in real life: not any of the usual characteristics (like sex or race) can be ascribed to them. So *tabula rasa* is the perfect description. The following question, then, imposes itself: How do such communities manage to cope with the problems and anxieties inherent in such an open approach? How do they manage to reap the fruits while keeping possible abuses in check?

As can readily be seen, the issue of trust is central to this matter. Open source communities seem to put complete trust in total strangers, in the sense of relying on their goodwill and professional competencies in contributing to the common cause. Notice that the problematic issue is not so much that proposals deposited in the commons for public access are entrusted, as it were, to the outside world. These are simply gifts to the community, be it with some strings attached (as specified in the open source licensing conditions). Rather, the more vexing issue to be investigated here is whether contributions in return can be trusted. Can unknown contributors to the collective project be considered trustworthy? From the point of view of the community, risks emanate mainly from the uploads as performed by strangers; hardly from anyone downloading the public content. This is so while the official contents of the collective project have to remain on course, protected from unwanted intrusion and disturbances (like viruses and worms in source code, or vandalism towards encyclopedic entries).

It will be argued that this problem of trust is handled differentially, depending on the phase in which an open source projects happens to be. I suggest that two phases can usefully be distinguished: an informal and a formal phase. This is an analytical distinction; in reality, the two phases often co-develop together. In the initial phase, project leadership relies on informal processes to keep a project on course. By implication, strangers are fully trusted to contribute to the common cause. Trustworthiness, it will be argued, cannot be inferred directly in any plausible way. Participants are therefore considered trustworthy simply by assumption. I will explore whether such assumption may have rational underpinnings.

The answer will not be found so much in the mechanism of seeking esteem (as proposed by Pettit), but rather in the mechanism of substantial hope (as proposed by McGeer). Moreover it will be argued, that the community may obtain some assurances from a common culture, specific to the community involved. For open source this is the hacker ethic, for Wikipedia this is a common code of conduct for 'proper Wikipedians'. While adherence to such a culture cannot reliably be signalled by strangers in cyberspace, only a weaker form of inference is involved based on statistical reasoning.

In the next, formal phase, rules and regulations are applied to manage open source projects. This is a common development as soon as projects grow, both as regards the number of participants and the size of created contents. Projects then come to embody structure in order to manage the complexities and retain some amount of order and efficiency. A whole range of governance tools are in use (such as division of roles and decision-making structure). The link with trust becomes more intricate now. The usual view is that rules may substitute for trust - and so reduce the trust needed. It will be argued instead that governance by rules and regulations may transform the problem of trust in a variety of ways.

On the one hand, rules may be designed starting from the premiss that participants can be fully trusted. A maximum of discretion for participants will be designed in, so to speak. On the other hand, the leading presumption may be the opposite: participants cannot be trusted to deliver reliable content of their own accord. Therefore as little discretion as possible - without stifling voluntary contribution altogether - is granted by the structural design. A low-discretion design signalling low trust is the outcome. In between these extremes, a continuum ranges from high to low discretionary design. All along, trust is indeed substituted by rules, ranging in extent from a small to a large substitution. It will be argued, finally, that the more trust actually gets substituted, the more assuming trustworthiness makes way for inferring trustworthiness (by means of entry requirements).

This argument will be developed below by a close analysis of developments in both open source software communities (FreeBSD and Mozilla in particular) and encyclopedic communities (Wikipedia in particular). The selection of cases is meant to cover some typical open source communities currently in existence.

Open Source Software: Initial Phase

The origins of the open source software (OSS) movement go back to the 1980s. Hackers - as they liked to call themselves - were used to freely exchange pieces of source code they had written. Then large companies started to enforce some of their alleged intellectual property rights on software. In particular, AT&T did so

concerning UNIX. In response, hackers rallied together in an effort to keep source code free (i.e., freely available). As a result, famous packages like FreeBSD (with a BSD-license) and the GNU Emacs editor (with a GPL) were developed. In the early 1990s, the Internet – itself largely the fruit of such open source practices – boosted participation in open source projects. With one click of the mouse people from anywhere around the globe could join. As a result, the numbers of OSS projects and participants rose sharply. Estimates of the total number of OSS projects currently underway amount to over 100,000 (on platforms like Freshmeat and Sourceforge).

Central to my investigation is precisely this Internet-boostered era. At least initially, regulating rules were far and few between. Usually someone initiated a project by putting a source code proposal on the web and inviting comments, patches and new features. This initiator – usually male – then operated as project leader, trying to manage the whole undertaking. The number of people responding could assume astonishing proportions. Larger projects easily attracted the attention of thousands of people out there (as evidenced by their downloads); among those, hundreds might actually send input back to the project, whether comments or code.

An astonishing feature of this open source process is the near total trust that was put in strangers – outside a core of close friends that often existed. After having made a few useful patches, contributors were easily welcomed as developer, often with permission to upload code into the official tree of the project. So it was a big ‘bazaar’ indeed, almost without ranks and distinctions, all babbling together and hacking away on the code tree. But what about the quality of these return gifts of code? Might some of these possibly be misguided, poorly formulated, misleading, outright irrelevant, or even poisoning? Might as a result the code tree become corrupted? All of these objections notwithstanding, leaders – at least in this initial phase – practiced near total trust towards strangers and did not attempt to delineate trust more carefully – as happened later.

Assumption of Trust: Pettit and McGeer

What were the bases of this trust? In real life, when we meet people, we are used to being able to infer some amount of trustworthiness from their characteristics. People’s family background, ethnicity or sex may be interpreted as providing trust in the prospective transaction (characteristic-based trust; cf. Zucker 1986). Put otherwise, these ascribed characteristics serve as flags that signal trustworthiness towards observers. But on the Internet, no such inferences about contributors are possible. All characteristics that might give clues about trustworthiness are hidden from view. It is only IP-addresses presenting themselves with hopefully use-

ful contributions. By definition Internet-participants cannot reliably signal anything at all.

So we necessarily arrive at the conclusion that open source leaders simply assume that fellow hackers are trustworthy enough. But then, we may continue to ask, what bases can be discerned that lend some credibility to this assumption? Can any good reasons be supplied for this assumption? More than a decade ago, Philip Pettit (1995) proposed the mechanism of 'secondary trust': while people are sensitive to the esteem of others, they will favourably reply to acts of trust in order to actually reap this esteem. The chance to be admired cannot be forfeited. Or alternatively, people cherish some amount of self-esteem and will therefore behave in a trustworthy fashion in order to avoid feelings of shame. In our context of OSS this reads as follows. Fellow hackers that stumble upon an open source proposal and decide to return reliable code, do so while they are seduced by the prospect of being admired by its project leader. Alternatively, they will refrain from poisonous contributions while they want to avoid embarrassing themselves. This imputation of 'normative pressure' (in the Luhmannian sense) to a source code proposal, as written in C or Python or any other programming language, might have some plausibility.

Recently, this mechanism of normative pressure has been reformulated by Victoria McGeer (2008). She tries to move away from the calculative and cynical conception of as-if trust as formulated by Pettit. Instead, she focuses on moves inspired by the kind of trust that does not rely on coldly weighing the evidence available but is prepared to go beyond ('substantial trust'). It is based on a vision on and hope in the capabilities of the other. By the trusting move the other is hopefully energized to realize his capabilities to the full. Such trust is empowering the other, not – à la Pettit – seducing or manipulating the other. As the prototype of this hopeful trust she presents the example of parents who sometimes have to let their children go and engage in risky adventures. Such trust is a hopeful bet on a future in which their children will finally be able to take care of themselves.

To me, such theorizing about the assumption of trust seems the more plausible avenue in the case of OSS communities. By openly exposing their gifts of code hackers can be seen to appeal to the hacking capabilities of unknown others out there. These are urged to show in return what they are worth as writers of code. Hackers are challenged to show their true potential. It is not so much esteem in the eyes of the project leader or avoidance of shame in one's own eyes, but the exercise of one's hacker capabilities that is spurring participants into action. A continuing cycle of sound code contributions may ensue.

Hacker Ethic, Old and New

Nevertheless, open source hackers did not operate in a total void as suggested above. When the Internet opened up avenues for massive participation, they had the cooperative experiences of more than a decade behind them – be it on a much smaller scale. These were imbued with what has come to be labelled the ‘hacker ethic’. This conception was first coined by Steven Levy (1984), describing computer wizards from the 1960s to the 1990s. True hacking as a way of life revolves around spectacular and novel ways of using the available capabilities of computing. All along, the emphasis is on constructive cooperation and sharing. Bureaucracy, security, passwords, and copyrights are detested as ever so many bureaucratic impediments to fruitful exchange. In 2001, Pekka Himanen suggested even grander dimensions for the ethic of the 1990s hackers, closely tied to OSS development. In his vision, such hacking is a creative passion that is embedded in a new work ethic for the information age, which focuses on sharing of information and keeping the Internet open for all, in a spirit of caring for all.

In the 1990s hackers of such persuasion took the Internet route for developing OSS together. By that move, a much broader audience came to be addressed. Could they be supposed to be bound by the same ethical standards as ‘true’ hackers?! Was there any ground for optimism on that score? My answer is in the affirmative. I would argue that it was not just a blind gamble. That broader audience was, on the one hand, composed of ‘true’ hackers living elsewhere on the globe, and on the other hand, of members of the computer underground, the ‘new’ hackers so to speak. Writings of the latter have been analyzed by Steven Mizrach (1997). After distinguishing several categories of this underground (such as system intruders, phone phreaks and virus writers), he gave an overview of the do’s and don’ts in their ethical self-conception. He finds a considerable continuity between old and new hacker ethic. The following principles in the new hacker ethic are particularly relevant for my purposes: share and exchange information with other people; do not take information and software from other people only (no hoarding, no freeloading); do not damage anything upon entering other computers or data systems; do not crash others’ systems by destroying hardware or data, by unleashing viruses, Trojans, or logical bombs.

So also new hackers predominantly cooperate and share, while avoiding damaging and hurting - in spite of the bad press reports about the few that deviate from this moral baseline (‘crackers’). This suggests that opening up source code proposals to the world at large was not so irrational after all. Amidst the potentially thousands of downloaders only a small fraction could be expected to be knowledgeable enough to be able to reciprocate. And that fraction would seem to be bound by some kind of hacker ethic, whether of the old or of the new variety.

Trusting their contributions seemed to be warranted. It was only much later - to be explored below - that these assumptions could no longer be upheld.

Therefore my conclusion is that some kind of hacker ethic was shared by the open source audience at large - at least by those who were returning comments or code. As a result, the blind gamble of assuming trust on a global scale was given at least some foundation: the hacker ethic was instrumental in making the OSS experiment successful. Obviously, the question imposes itself whether and to what extent its participants shared my analysis? That is, did they make an estimate of the constructive attitude of global audiences, was that estimate the same as mine, and did they consciously interpret it as a support for their open source experiment on the scale of the Internet? If the answer is in the negative, it implies that OSS participants embarked head-on upon the gamble of assuming trust worldwide, without any assurances. If the answer is positive, it implies that some amount of inferred trust was also supporting the decision to go ahead and share. Anyway, to my knowledge, the answer is unknown.

Notice that in my analysis culture plays at most a very modest role in creating trust. In the usual analysis, members of a culture are supposed to flag their allegiance to it to potential trustors. Visible membership is a sign from which observers may infer an amount of trust (characteristics-based or institutional-based trust). In my analysis of open source communities, culture recedes to the background. Being a true hacker cannot be reliably signalled, due to the nature of virtual communications. So a potential trustor can never be sure that a specific potential trustee indeed is a member of the hacker tribe. (S)he only may obtain assurances of a statistical nature that the hacker ethic obtains in general. The inference that trust is warranted in a specific case becomes more fragile. The support of culture in creating trust in virtual life can only be much weaker than in real life - a strong form of inference is replaced by a much weaker form of inference.

Wikipedia: Initial Phase

The movement for producing encyclopedic entries in open source fashion is of more recent origin. It all started in 2000 with the American Nupedia, written and reviewed by experts. While that undertaking was slow to take off, Wikipedia, under the leadership of Jimmy Wales, was launched as a kind of experiment. Everybody, unregistered and anonymous, became entitled to read and 'edit' entries in the online encyclopedia ('editing' meaning changing, deleting, or adding content). Three 'pillars' have to be observed in the process. 'Neutral point of view' means that articles should represent all significant viewpoints to an issue fairly, proportionately and without bias ([http://en.wikipedia.org/wiki/Wikipedia: NPOV](http://en.wikipedia.org/wiki/Wikipedia:NPOV); henceforth

for all English Wikipedia references the prefix <http://en.wikipedia.org/wiki> will be omitted); 'no original research' means there is no room for original research of one's own that has not yet been published elsewhere (Wikipedia: NOR); and 'verifiability' means that all content that is likely to be challenged should be traceable back to a reliable source (Wikipedia: V).

And soon enough this was a great success, at least numerically. Local Wikipedia's were created, in languages other than English. These number over 250 by now. The biggest English-language Wikipedia contains over 2.8 million articles, while a small one like for example the experimental one in Kuanyama contains as yet only 5 articles.

This movement is to a large extent modelled upon OSS experiments. The basic software tool, a wiki, allows distributed participants to work on the same body of text simultaneously (just like versioning systems do for source code). Wikipedia also admits everybody as contributor (denominated as 'editor'), and makes entries available for everybody (with a GNU Public Documentation License tailored to texts, the equivalent of the GPL for software). So here again, just like in OSS, we find almost unlimited trust in strangers from all over the globe. The numbers involved are impressive. For the English Wikipedia alone, apart from anyone being allowed to edit entries, 9.2 million people have registered, and, as a result, may additionally start a new article of their own. Roughly, 1/3 of edits comes from anonymi, 2/3 originate with registered users. And similarly for other language versions, with diminishing numbers of users.

Again we pose the question: what mechanism of trust is involved? What grounds can be advanced for trusting outsiders not to damage entries, introduce minor or major mistakes, or edit the details of their own biographies? Inferring trust from available signs is impossible – anonymous contributors can only be identified by their IP addresses, not very revealing in themselves. So again, we are dealing with an absolute assumption of trust – the point of departure for dealing with outside contributors being that these can be fully trusted to contribute to the worthy cause of an encyclopedia by all and for all.

It is my contention that this assumption of trust creates a kind of normative pressure - in the Luhmannian sense – to contribute loyally. Openly editable entries invite not only to read, but also appeal to the specific capabilities and knowledge of people out there. These are urged to employ and reveal their abilities and return comments of their own, however tiny these may be. In short, the open policy motivates others to show themselves to be real Wikipedians in the making. As such, Wikipedia can be interpreted as an expression of substantial trust in unknown others (à la McGeer), based on a vision of the ultimate attainment of the encyclopedic ideal of knowledge accessible for all and developed by all.

Wikiquette

At the same time, astonishingly enough, the Wikipedia adventure was not supported by an underlying ethic from the beginning (cf. the hacker ethic above). At the time that Wikipedia started, the approach was completely novel. No relevant ethic of any kind had crystallized; no one could forecast how people out there actually would react. An encyclopedic community was terra incognita. In that sense, Wikipedia was more of an adventure than OSS, with its decades of experience before the Internet boosted collaboration.

But soon enough, problems did surface. Several varieties of 'disruptive' behaviour emerged (Wikipedia: BP): on the one hand, vandalism towards entries (like changing small details, inserting nonsense, adding obscenities or crude humour, blanking pages, and changing details of one's own life; cf. Wikipedia: VAN); on the other hand, gross incivilities, persistent harassment, and threats or attacks against editors personally (on discussion fora, on talk pages, by e-mail, e.tc.). A specific term was coined: 'edit warring', referring to contributors fighting over the contents of an entry from their own point of view by repeatedly deleting each other's changes (in the Wikipedian jargon these are called 'revert edits', returning an article to an earlier version). Notice that the reverts as such need not be unjustified, but it is the lack of any explicit comment or justification that may make the act, to many a participant, rude and insulting. All kinds of rules were devised to deal with the phenomenon after the fact; these will be analyzed below. What matters here is, that simultaneously a kind of cultural offensive was relinquished, to develop a kind of 'wikipedian ethic' and fill the ethical vacuum that existed at the outset of the experiment. The true Wikipedian had to be constructed *de novo*.

Several texts testify to these efforts. The focus is on constructive argument in order to be able to reach consensus over articles. Proposals for textual changes or deletions should always be accompanied by arguments (in the 'talk pages', with one's name and date attached). In the process, one should be civil and avoid incivility (Wikipedia: CIV). Civility means: a considerate, polite and respectful attitude towards others (remember the Golden Rule) in discussing differences of opinion. Incivility - to be avoided - means: being rude, uttering insults or profanities, personally attacking or harassing other editors, and the like. The atmosphere, moreover, should be open and warm. Turn the other cheek if necessary, give praise, and forgive! (Wikipedia: EQ). All along, assume good faith with other editors (Wikipedia: AGF). Assume that they want to help the Wikipedia project, not hurt it. Help them gently to correct their mistakes if any (like introducing 'original research' or attacking someone personally). Do not accuse anyone lightly of bad faith; and above all, do not forget to show your own good faith.

This 'Wikiquette' – as it is called - is repeated in guidelines for newcomers. As regards content newcomers are urged to not contribute an article about themselves or their company, to add or delete content with caution and with arguments only, and to avoid chatting or flaming (Wikipedia: ACM). As for treating newcomers, established wikipedians are advised not to bite them and so scare them away with hostility (Wikipedia: BITE). Be respectful and constructive in correcting them (= civility); make them feel welcome (= warm atmosphere); assume good faith on their part (= good faith). Give them a chance! Ignorantia juris (i.e., of Wikipedia law) and inexperience may be excused. One can have faith in this approach, it is asserted, while 'many new users who lack an intuitive grasp of Wikipedia customs are gradually brought around once the logic behind them becomes more clear' (Wikipedia: AGF). This approach is carried as far as making available pre-fabricated templates cordially welcoming newbies into a project (Wikipedia: WT).

So here we clearly find the articulation and fostering of an attitude of trust towards fellow-wikipedians, in the sense of assuming good faith in co-editors one is continuously dependent upon in writing the entries together. It is a civilization campaign in order to keep the open approach to editing viable and alive. The true Wikipedian had to be co-constructed with the design of Wikipedia (to be explored below). While in OSS culture preceded structure, with Wikipedia these have been evolving simultaneously. As a result, some confidence in fellow-wikipedians has become warranted. The trust involved is no longer purely an assumption, but has partly become based on inference; in particular, on weak inference based on statistical reasoning.

Open Source Software: Rules and Regulations

In OSS, the 'bazaar' soon came to be regulated. Especially in projects that grew in size, the 'simple structure' of project leaders-cum-followers did not suffice anymore. For reasons of efficiency and manageability rules, regulations, and prescriptions had to be introduced. The most important governance tools to be introduced in this formal phase are the following (as described in De Laat 2007):

- **Modularization.** In many larger projects, the code tree is divided into several subtrees. In this fashion, dozens of modules may be carried out in parallel.
- **Division of roles.** In almost any project nowadays, roles are distinguished that define what the occupant is allowed - and expected - to do inside the project. A common role division is between observer, developer and project owner.
- **Decision-making.** In every project, decisions have to be made about a range of matters, for example, the methods to be used, acceptance of code in the

main tree, preparing new releases, e.tc. Powers of decision-making in such matters are formalized and in some way distributed over participants.

- **Formalization.** Technical tools and standardized procedures have been introduced to streamline virtual cooperation. This applies to mutual discussion, reporting of bugs, working on the code tree simultaneously, and testing.

As a result, OSS projects typically assume a design of a kind. Various arrangements are in use, some of which will be explored below. What matters here is the relationship with trust. How can the introduction of governance in a project – i.e., creating more structure than the ‘simple structure’ alone – be related to our issue of trusting virtual hackers in cyberspace? Any project design is a mixture of discretion and prescription. Certain activities are allowed to participants (terms of discretion), within the limits of certain formal rules (terms of prescription). The amount of discretion and prescription, by the way, are not necessarily (negatively) correlated. These designs can all be interpreted as reducing the dependence of a project upon the whims of outside contributors. The risks incurred by an almost total absence of structure are curbed. Precisely while their ever voluntary activities become more precisely circumscribed and channelled, participants retain less leeway to damage, harm, or obstruct (whether on purpose or not). The connection with trust can easily be established by now: the amount of trust granted to the world outside is reduced. As a result, less trustworthiness on the part of virtual strangers is needed; the structural arrangement partly substitutes for the need for trustworthiness (cf. Sitkin and Roth 1993). Whatever amount of trust of necessity remains to be invested in potential participants has to be inferred or simply assumed (cf. discussion above).

Obviously, differences in size, technology, phase of maturity and the like, will all impinge on the shape of a proper design. As a rule, assumed designs will vary among projects. More importantly, for any individual OSS project there will be a variety of possible designs that may be adopted; not one but many options are open. As in organizations, it cannot be assumed that there is one best way of organizing. As a corollary, this implies that project leaders usually face choices in this respect: what design to adopt and on what grounds? Note that variety in design necessarily implies variety in trust granted – and so of variety in the amount of trust remaining from the former situation of total trust. The amount of trust that is substituted is in direct proportion to the amount of rules and regulations in a specific design. This observation is not just an academic remark: choices and variations on this score may heavily influence mutual relations within a project. So any design choices to be made have to carefully considered and given attention.

High-Discretion and Low-Discretion Designs

Obviously, this assertion needs elaboration. Let me distinguish between two kinds of design variables. On the one hand, parameters may be controlling/coercive: they are the means for project leaders to obtain control on the workings of the project and to be able to enforce compliance. This is the usual function associated with bureaucratic rules. On the other hand, parameters, if appropriately designed, may actually facilitate contributing to a project and reduce technical inefficiencies. Such variables will be called enabling. These denominations are modelled after Adler and Borys (1996) who alerted us to the fact that bureaucratic rules – in organizations – may sometimes be designed to enable employees to master their tasks, rather than to enforce their compliance.

Looking back at the four parameters distinguished above, some of these would clearly seem to be enabling in the sense just coined. Both modularization and formalization streamline an otherwise chaotic and unordered process. It is structuration as a minimal condition for fruitful collaboration to occur. Formalization tools in particular may function as tools that enhance programming capabilities (cf. also Andrews et alii 2005). As a result, although actually less trust is granted to contributors, this is welcomed as a useful ordering of processes – not detested as encroaching control. Mutual relations and atmosphere can only be strengthened. It is precisely for that reason, I would suggest, that the whole OSS community seems to have standardized on the use of such formalization tools (Robbins 2005).

The introduction of role divisions and decision-making powers, on the other hand, is a more delicate affair. These parameters may easily be considered controlling, as they touch directly on the amount of discretion that outside collaborators may enjoy. The choices lie on a continuum. On one end of the scale, the role division employed may be minimal and decision-making decentralized. Discretion granted remains high – as high as the particular project seems to allow. In such a high-discretion design, trust in virtual strangers remains high. As a result, outside contributors may be expected to remain committed to the project and continue contributing code or comments. On the opposite end of the scale, a low-discretion design can be introduced: an elaborate division of roles is carved out, with minimal discretion at the lower end, and decisional powers highly centralized. Such a design effectively awards little trust to contributors. While hackers may be supposed to be attached to autonomy in their voluntary activities, this might well be interpreted as a manoeuvre of control, as a way of expressing distrust in their very capacities and/or intentions. As a result, volunteers might be chased away from continued participation and enlist elsewhere. Of course, the above is an adaptation, for my own purposes, of the famous distinction between

high-discretion and low-discretion work role patterns in organizations as developed by Alan Fox decades ago (1974).

Take for example Tigris, a well-known platform that hosts a great many OSS projects. On the site it is explained that a role consists of a set of permissions granted; a permission allows specific activities (like reading, or editing) to be performed upon specific resources on the project's site (like project documents or source code files) (<http://www.tigris.org/scdocs/DomAdminRoles.html.en>). A common 3-fold division of roles is the following (cf. <http://www.tigris.org/scdocs/ProjectRoles>). An observer has read-only access to most of the project's documentation and source code files, and may return comments and/or code proposals and patches. A developer obtains more permissions: (s)he also obtains write-access to the official source code tree and project text files. A project owner is someone at the top who manages the project as a whole (and part of the job is precisely to grant membership roles as just discussed). Notice the pivotal role of developers: they are the ones who are empowered to incorporate changes in the code tree. For getting role permissions, candidates have to qualify – though standards seem quite relaxed. After surveying a project as anonymous guest (who are allowed to see most of the project's files), one may ask to get observer-status; this will as a rule be granted to anybody. Then, after delivering contributions to the project of sufficient quality, one may obtain the status of developer. So only for obtaining the status of developer one has to prove oneself to be professional enough.

For many smaller projects, consisting of just a few modules, this design will do. Each module is run with an 'owner' at the top. The design is still quite similar to the 'simple structure' from the initial, informal phase – and with a rather high amount of discretion still. But for larger projects, the structure will be evolving: of necessity, the design will move towards curtailing discretion. An often-discussed example, at the lower end of discretion, is Linux. In that vast project with a range of modules, 'trusted lieutenants' (above the layer of maintainers of the modules) are the ones who take all proposed changes into consideration, with the final say still exercised by Linus Torvalds. So this is a very centralized design. In order to show that intermediate designs do exist, and highlight what choices can be made regarding role division and decision-making, two other larger OSS projects will now be analyzed in some detail: FreeBSD and Mozilla. The two have made slightly different choices as far as design is concerned.

FreeBSD and Mozilla

Let us first consider FreeBSD, an operating system developed by volunteers for decades now (the following account is based on Jørgensen 2001, and made up-to-date by consulting http://www.freebsd.org/doc/en_US.ISO8859-1/books/dev-model/). Roles (called 'general hats' inside FreeBSD) distinguished as far as module development is concerned are: contributors, committers and maintainers (quite comparable to the Tigris distinctions). After accessing and exploring files contributors (several thousands now) are those that start contributing comments and/or code. Committers (392 active ones at the moment) have obtained write-access to the code tree, and actually may commit code – either of their own, or from fellow-developers (without write-access). Maintainers (to be recruited from the committers) are at the top of a module and coordinate the incorporation of new code. Notice that for this large project a whole range of other administrative 'hats' are also in use.

The procedure to be followed by a committer for getting his code accepted is instructive. After writing a contribution, he is urged to discuss the changes first with fellow-committers. Preferably, code has to be reviewed several times. As the next step, he has to test the proposed changes by integrating them in the module involved and trying to compile the module as a whole (the build should not 'break'). This is a pre-test in his own downloaded copy of the current development version. If the test has succeeded, it is up to the committer to actually integrate the changes into the official code tree. From then on, the changes are open for debugging and commenting by anyone. So in effect, committers decide (semi-)autonomously about the incorporation of pieces of code. The fact that the results of one's work immediately become visible and available for all, is reported to be a gratifying experience.

Not just anybody may rise upwards in this hierarchy. True, anyone may look around and/or contribute. But in order to become a committer, one has to qualify (http://www.freebsd.org/doc/en_US.ISO8859-1/books/dev-model/model-processes.html). After having made a number of high-quality contributions, a developer may ask for committer status. It is upon recommendation by a committer that the so-called Core Team (above the layer of maintainers, consisting of 9 members) will vote about granting that status.

In Mozilla, a web browser, a mail client (and more), the same division of roles is employed (Holck and Jørgensen 2005; updated by consulting the Mozilla site, cf. below). For each of the by now some 80 modules, contributors (i.e., users and developers), developers with write-access (committers) and module owners (helped by 'peers') can be discerned. The procedure for code contribution is iden-

tical to FreeBSD: discuss informally, test in your own working copy of the tree, incorporate in the main tree, and let others debug. There is one significant difference however. The decision about incorporating code changes in the official tree is no longer within the committers' authority. For several years now they need to ask for permission in 2 steps (<http://www-archive.mozilla.org/hacking/code-review-faq.html>). First a contributor asks approval from his module owner. Thereupon, so-called 'super-review' is called for: someone preferably with expertise in another domain has to judge whether the patch fits into the broader Mozilla code base. Officially appointed super-reviewers (some 30 in all) have to give their approval.

In accordance with this design, the procedure to become a committer is quite strict (<http://www.mozilla.org/hacking/committer/>). Basically, the contributors first need to demonstrate that they know what they are doing, by having contributed some good patches. After that, they may formally ask to obtain committer status. For this, they need two people that want to vouch for their competences. These vouchers, usually one's module owner or a peer, are responsible for them during 3 months. Moreover, one of the super-reviewers must stand behind their nomination.

From these descriptions it transpires that both projects are in the middle range of discretion granted towards outside contributors. Write-access to the code tree is never granted immediately or indiscriminately. However, with Mozilla, in comparison with FreeBSD, decision-making is more centralized. Supervision has tightened. It is the layers of module owners and super-reviewers that are the judges of incorporation of code, no longer the code developers themselves as in FreeBSD.

Some influential Mozilla leaders (B. Eich and D. Hyatt) even forcefully argue that the trend of more control has to be strengthened:

The faux-egalitarian model of CVS access and pan-tree hacking that evolved from the earliest days of Mozilla is coming to an end. (...) [One of the key elements in the new roadmap is to] continue the move away from an ownership model involving a large cloud of hackers with unlimited CVS access, to a model, more common in the open source world, of vigorously defended modules with strong leadership and clear delegation (...).

(<http://www-archive.mozilla.org/roadmap/roadmap-02-Apr-2003.html>)

Trust Substituted

It may be concluded that as OSS projects grow in size and complexity, the trend towards introducing structure is ineluctable. In the process, every project will

have to find its own design solution – within a range of possible solutions. Whichever design is chosen, some amount of trust needed before is substituted by structure. Less trustworthiness has to be supposed to be present with outside anonymous contributors, whether by assumption or by inference, while rules and regulations take care of this.

As a corollary it should be noted, that a recent trend can be discerned for many an OSS project of perceiving outsiders as less trustworthy at the outset. Fears seem to be mounting about sloppy code, or code with bugs, or even Trojan horses being introduced into the main tree of a project. Moreover, concerns over intellectual property rights are mounting: whether on purpose or not, contributors could incorporate code carrying a license which is incompatible with the existing licensing arrangement, or could import patented matter – with all attendant dangers of patent litigation later on. So in a way, the hacker ethic – as described above – is seen to be eroding.

In response, potential committers of code are likely to be more subjected to screening – that is, more screening than the usual procedure for ascertaining technical skills which amounts to showing some good work (cf. the procedures as described above for FreeBSD and Mozilla). In Debian, for example, contributors who want to become ‘real’ developers with write-access to the tree not only have to demonstrate their technical capabilities, but also have to reveal and prove their real life identity, and show their factual understanding of and ideological attachment to the cause of OSS (‘Debian new maintainer process’; described on <http://debian.org/devel/join/newmaint>; cf. also Coleman and Hill 2005). This tightening of entry qualifications is a clear move from the assumption to the inference of trust. Whatever amount of trust is not yet covered by structure and still needs to be granted to outside contributors is henceforth secured by the more reliable process of inference. The pure assumption that one is dealing with a ‘true’ hacker seems no longer justified.

Wikipedia: Rules and Regulations

It is time to turn to developments in Wikipedia now. Also in this project, rules and regulations soon came to be applied. I would contend that the same type of analysis developed above for OSS applies. Governance tools were introduced: modularization (several subprojects were introduced under the umbrella of Wikimedia such as Wikipedia, Wiktionary and Wikibooks; within them, each and every entry is a module in itself), division of roles & decision-making (to be discussed below), and formalization (rules for editing pages, for discussing changes, for reporting and handling vandalism, e.tc.). And so a design was introduced, with terms of discretion and terms of prescription for participants. As

a result, trust towards the outside world was substituted in part. Depending on the particular configuration of design parameters, the result may vary: from a high-discretion design on the one hand, to a low-discretion design on the other. Which kind of design did actually crystallize within Wikipedia? I intend to show, first, that the design as gradually introduced is characterized by very high discretion – even higher than for OSS –, and accordingly by high trust towards outside contributors. Furthermore it will be shown that at the time of writing pressures are mounting to reduce this discretion and be more careful about granting trust; all of this as a response to increasing vandalism and harassment.

The best way of exploring the design is to start with the division of roles (as employed within the English Wikipedia in particular) (Wikipedia: UAL). Just like in the Tigris division of roles, it is all about obtaining permissions for performing specific activities upon specific resources on the site. Users come in three varieties. The anonymous user (no account) may read and edit all entries. As soon as (s)he has created an account (username, password) and logs in, a user may in addition create new pages, as well as email other users who publicly mentioned their address. A user automatically obtains the autoconfirmed status with special privileges (such as uploading files, and editing semi-protected pages, to be explained below), as soon as (s)he has made 10 edits and has been registered for 4 days in a row.

Above these user levels other roles have been defined that mostly have to do with protecting Wikipedia against disruptive behaviour (as described above). In order to keep disruptions in check, administrators (aka sysops) obtain the right to protect pages, delete pages and block users (see below); over 900 of them are active right now. Very trusted users may become bureaucrats, who are entitled to appoint users as administrators or bureaucrats (almost 30 active at the moment). Both role occupants are not simply appointed from above, but should formally apply and be accepted by the broader community after – usually lengthy – public discussion. Although these are the main roles, notice that a whole array of others can be found that I will not go into here (such as the ‘checkuser’ - who may run a thorough check upon users suspected of misdemeanour, and the ‘rollbacker permission’ - allowing a fast way of reverting vandalist changes).

The administrators, as ‘police officers’ (my terminology) of Wikipedia, have several measures at their disposal to deal with incidents of disruptive behaviour. They may introduce page protection: a page involved in a dispute can no longer be edited or moved, usually for one or two weeks (Wikipedia: PP). With ‘full protection’ not any user is admitted, with ‘semi-protection’ autoconfirmed users still are. Whichever protection applies, administrators may still touch and edit the page involved. Protection is useful as a period of cease-fire, in order to let

contestants resolve their conflicts. Similarly, administrators may delete pages as a last resort, in order to deal with vandalism, lack of reliable resources or copyright violations (Wikipedia: DEL).

Another measure that administrators can apply in case of serious disruptions by particular users, is to block them, that is, bar them from any further activities on the site (except for reading) (Wikipedia: BP). Such a block can last a day or longer, depending on the circumstances. It is intended to protect the Wikipedia project, not to punish the user involved. A particular example of behaviour considered incontrovertible evidence of edit warring is making 3 reverts anywhere within 24 hours. Therefore, anyone indulging in it may be blocked from further editing (the 'three-revert rule'; Wikipedia: 3RR). In the same vein, someone who stealthily creates an alternative account in order to be able to push his personal point of view twice or stir up controversy (a 'sock puppet') also risks to be blocked (Wikipedia: SOCK). The same measure may apply to a user who lets someone else register in order to push his own point of view ('meatpuppet'), or alternatively, push the opposite point of view ('straw man sock puppet').

Notice that the police officers involved also get assistance from software bots that are entitled to revert changes that most obviously emanate from vandalism or edit warring. As a result of both men and machines policing Wikipedia, such changes are reportedly corrected within a very short time span (median correction time from 2 to 3 minutes; figures referring to October 2005; Viegas et alii 2007). This fast repair mechanism is a form of so-called 'soft security' – protecting the system in unobtrusive ways, almost invisibly, and after the fact.

How may the design as crystallized within the English Wikipedia be characterized? It has always been a high-discretion one. Users may edit at will, change whatever they like, even unregistered and anonymous. Trust in their potential contributions is high. Meanwhile, as we have seen above, limits have been set upon contributing: no edit warring, no vandalism, no harassment. In those cases, disciplinary sanctions may be applied. Rules such as the 3-revert-rule or the sockpuppet rule are clearly bureaucratic rules that coerce users. But apart from such rules for irregular behaviour the regular user is still at the helm. This is so, presumably, while users are considered the main assets upon which the destiny of the encyclopedia depends. Without their massive voluntary collaboration, Wikipedia would have suffered the same fate as Nupedia or Knol. Peer production of (encyclopedic) knowledge is taken to its logical limit: 'democratic' production by all.

Remarkably, the trust granted is even higher than was ever the case in OSS. For software, contributors who want to acquire write-access to the code tree have always been held to some proof of their coding capabilities. Becoming a commit-

ter was always a privilege. Moreover, maintainers often do own modules in the sense of actively interfering with code patches from committers. In Wikipedia, in contrast, users have always had immediate write-access to articles and other pages, no capabilities test or edit review are required. The motto is: let everyone be a Wikipedia committer!

Review and Super-Review

However, signs are increasing that the era of total trust in strangers is coming to an end. Discretion is circumscribed in tiny steps. A first, hardly noticeable step had to do with the incident about the journalist John Seigenthaler: an anonymous user created a biographical entry about him containing false content (May 2005) (Wikipedia: Seigenthaler hoax). It went undetected for several months. In response, Wales barred unregistered users from creating new pages. From December 2005 onwards, an anonymous visitor may no longer create a new page but only read and edit existing entries.

A more serious encroachment upon full discretion for all is the call for review: all changes should be checked on vandalism before incorporating them in the 'stable' version of a page. While the system is still under discussion for the English Wikipedia, it has already been unrolled gradually in the German, Russian, Hungarian and Polish versions (since 2008). The software involved can create many varieties of reviewing systems. As regards the kind of entries to be reviewed, some argue for reviewing only the most sensitive ones (like biographies of living persons), others for reviewing all of them. Furthermore, who is to be censured: only anonymous users, or contributors at large? As for the reviewers themselves, should these be a select group of trusted users, or all (registered) users? And finally there is the question of what a visitor actually gets to see on the screen: the 'stable' version, or the 'experimental' version containing one or more as yet unreviewed edits?

The approach chosen for the German Wikipedia is the following (German Wikipedia: Gesichtete Versionen; Wikipedia: Flagged revisions/Sighted versions; <http://de.labs.wikimedia.org/wiki/Hauptseite>). All entries fall under the system, and edits from all contributors are to be reviewed (but see exception below). Reviews are carried out by Sichtern (literally: sifters): registered users who performed at least 300 edits and without block for 60 days automatically obtain these rights - about 6000 Sichtern are in function now. Highly experienced reviewers - at least 3000 edits, one year without a block and with public email address - are exempt from the process: their edits or articles do not need to be reviewed and automatically turn up in the stable version ('autoreview rights') (German Wikipedia: Gesichtete Versionen). On the screen, unregistered users get to see the so-called sighted (flagged) version (as default) - although they can also click on the newest

unsighted version if they wish. For registered users the most recent version is the default. The trick of this setting of the default is that anonymous vandalism no longer obtains immediate gratification: their changes do not show up in the official version.

While Wikipedians from Germany (and other Eastern European countries) are mostly satisfied with the flagged revisions scheme, it is hard to swallow for those of English tongue. Heated discussion is still raging on talk pages. Various alternatives are being proposed such as 'delayed revisions' - delaying edits made by users below the autoconfirmed status, after 2 hours they turn up in the official version (Wikipedia: Delayed revisions); or 'deferred revisions' - only suspect edits as identified by an abuse filter come under the flagged revisions scheme (Wikipedia: Deferred revisions). Also it is vehemently defended that the most recent version is shown to all users.

An opponent of the scheme observes: "The idea that we trust some users more than others on content is terrible" (17 January 2009; on the talk page about flagged revisions). And indeed, trust granted is effectively differentiated by the scheme. Lines of division are drawn: between those who may carry out reviewing, and those who may not (Sichtern vs. other users); between those who immediately get to see the most recent version, and those who are referred to the - possibly less recent - stable version (registered users vs. anonymous users); and, above all, between those who may edit without review and those who may not (Sichtern with autoreview rights vs. ordinary Sichtern and users). Especially the autoreview rights create a new elite by themselves. Trustworthiness as regards editing is no longer assumed from the outset, but has to be proven by one's editing track record within Wikipedia.

Effectively, Wikipedia design is moving towards a lower level of discretion. A move that is laudable to most German Wikipedians, but detestable to most English ones. This difference in appreciation can be linked to a distinction made earlier: between enabling and constraining rules. Obviously, the English feel offended by the curtailing of their discretion and the attendant differentiation of privileges. The scheme is interpreted as outright constraining – to them, bureaucracy (in the pejorative sense of the term) is setting in. On the other hand, the German interpretation of flagged revisions is that these enable the proper working of the encyclopedia, more vandalism-free than before. Discretion is gladly sacrificed, in order to gain in efficiency. Why, in the end, English speaking and German speaking Wikipedians differ so much in their diagnosis is an intriguing question – ripe for investigation.

In the near future, Wikipedia – more surely the German than the English language version – intends to introduce also a more severe kind of review, one that

checks upon the quality of articles. This will be done by Prüfer, also referred to as über-reviewers or surveyors. With this last move Wikipedia would move closer to conception and design of Citizendium. That online encyclopedia of recent origin (2007) does explicitly honour expertise, and distinguishes roles such as 'authors', 'editors' and 'constables' (cf. <http://en.citizendium.org/>). Their authors are comparable to registered users in Wikipedia, their constables are the equivalent of administrators in Wikipedia. And the Citizendium editors? These are supposed to guide the crafting of articles and approve the various versions; therefore, they come close to the proposed Prüfer in Wikipedia. And in comparison with OSS: Wikipedian Prüfer and Citizendium editors are looking quite similar to module owners in the average OSS project. In that sense, the design for Wikipedia is not only moving toward Citizendium, but also towards the average OSS model.

Conclusions

Open source communities squarely rely on the contributions of anonymous strangers in cyberspace. Therefore the problem of whether, and to what extent, such volunteers can be trusted to contribute in good faith and in a competent fashion is a central concern. It has been argued that such communities do indeed have a whole array of mechanisms at their disposal to handle this matter of trust. Based on the experiences of FreeBSD, Mozilla and Wikipedia in particular, the following picture emerges – table 1 lists the mechanisms involved.

In the initial, informal phase the assumption of trust seems to be the prominent mechanism. Making contents available for modification is a sign of 'substantial' trust (in the sense of McGeer 2008). In that vein, opening up source code is an appeal to fellow hackers to show their capabilities, and putting encyclopedic entries on the Internet as a wiki is an appeal to fellow Wikipedians to show their editing capabilities. The open approach empowers potential contributors and makes them step forward. This theorizing about the 'normative pressure' emanating from opening up content would seem to confirm the notion that, especially in cyberspace, assuming trust is a very important mechanism for creating trust in the first place (cf. De Laat 2005).

At the same time, some assurances may possibly be generated by underlying cultural processes. When hackers started to use the Internet for developing OSS, they had been cooperating amongst each other for decades in 'real life'. The 'hacker ethic' had crystallized, revolving around spectacular hacks in a spirit of cooperation and sharing. On the Internet they now bumped upon a younger generation of new hackers. If Mizrach (1997) is right, their 'new ethic' had considerable overlap with the old one. As a result, the open source audience as a whole was bound to some kind of hacker ethic. Their trustworthiness would seem to be

guaranteed. Whether and to what extent OSS initiators were actually aware of this cultural support for their cause when they started, is unclear.

The Wikipedia experiment was not so fortunate: no 'encyclopedic ethic' of any kind existed. So when vandalism and harassment started to emerge, making clear that not all anonymi could be assumed to be trustworthy, the community hurriedly started a campaign to educate actual and potential Wikipedians. A 'Wikiquote' was developed, prescribing respect, constructive argument, a warm attitude, and good faith towards fellow Wikipedians. As far as contributors feel bound by this ethic, they can surely be trusted to further the encyclopedic cause.

Notice the differences. In OSS the ethic involved preceded the Internet age, and its support for the OSS experiment possibly remained latent for the actors involved. In Wikipedia, the relevant ethic had to be fabricated from the start and was undoubtedly a conscious effort to develop good manners where they were lacking. Moreover, this cultural support for inferring trustworthiness of participants is of a peculiar nature. Inference usually relies on solid signs from the particular trustee. Ascribed characteristics, cultural allegiances, all can be flagged. Open source communities, however, depend on anonymous strangers who cannot reliably signal anything of the kind. They are just IP-addresses, pretending to be willing to contribute in a loyal fashion. Instead, such communities can only gauge the existence of culture among their contributors in a general statistical sense. If a culture obtains, they may infer, in a weak sense, that the anonymi that they will be dealing with probably can be trusted. So normally, open source communities are barred from inferring the trustworthiness of a specific contributor from available signs, and are condemned to making probability estimates. Probabilities replace certainties, due to the veil of the Internet.

In the next, formal phase, open source communities introduce rules and regulations, in order to manage the complexities involved. Governance tools distinguished above were modularization, formalization, division of roles and decision-making. As a rule, projects will have a range of choices in adopting a design. An important choice has to do with the discretion granted to collaborators (in analogy with Fox 1974). In a high-discretion design, role division is minimal and decision-making decentralized, yielding a high amount of discretion to virtual participants. A low-discretion design, with elaborate division of roles and centralized decision-making, on the other hand, leaves little discretion to collaborators. The link with our issue of trust is that any design, to some extent, substitutes for the need for outsiders being trustworthy. The rules and regulations, to some extent, eliminate the problem of trust by explicitly regulating dependencies. However, that extent is variable: the more one moves towards the side of a low-

discretion design, the more trust is actually substituted. Eliminating discretion saves one the trouble of having to trust those exercising that discretion.

The introduction of design has been explored in OSS projects such as FreeBSD and Mozilla. A typical division of roles consists of contributors, committers and maintainers (or module owners). It was shown how FreeBSD tends towards higher discretion (with committers autonomously deciding on inclusion of code in the main tree), while Mozilla tends towards lower discretion (with committers having to obtain approval in this regard from their module owners and 'super-reviewers').

At the same time, an associated tendency has been observed: the less discretion remains in the design involved, the more role occupants have to prove their worth before being accepted. Committers, for example, are getting more heavily screened about their real life identity, their professional as well as their ideological competences (Debian is a particularly striking example). In terms of our trust problem: the more trust has been 'designed out', the more the trust remaining needs to be secured by inference. That is, the assumption of trust makes way for the inference of trust. And this is inference in the classical, strong sense of the term - no longer based on statistical considerations.

Similarly, the encyclopedic Wikipedia project soon became subjected to design. Various roles and regulations were introduced. Fact is, however, that contributors retain a high measure of discretion. Anonymous users may still edit entries of their choosing; upon registration, they may in addition create new entries. Rising vandalism was not fought by reducing the discretion of ordinary users, but by introducing 'administrators' (and 'bureaucrats') who obtained powers to block particular users and freeze articles involved in edit wars ('page protection'). As a result, users remain at the helm in Wikipedia.

Nevertheless, because of rampant vandalism, pressures are mounting within Wikipedia for another measure: review of all changes (edits) as regards vandalism, before including them in the so-called stable (or flagged) version. The review is carried out by reviewers (with considerable editing experience) who patrol recent changes. Very highly experienced reviewers become exempt from the process (they may edit undisturbed). So users are only treated as really trusted, once they obtain these 'auto-review' rights. Every 'lower' kind of user remains subjected to vandalism review.

As such, this is a step that reduces users' discretion, no doubt about this. Interestingly enough, though, the system is not met with universal (dis-)approval. It has already been introduced in the German Wikipedia (and some others). The majority of Wikipedians involved applauds the system as a contribution to the

fight against vandalism. Their fellow English-speaking Wikipedians, however, vehemently resist its introduction as an encroachment upon their editing rights. To them, this looks like bureaucracy is setting in. This finding suggests that inter-cultural perceptions of open source design may differ considerably. As such, an interesting research field seems ready for exploration.

In addition, plans are afoot of reducing users' discretion even more. So-called 'super-reviewers' will in the future review articles upon quality. Such a tightening of design is a logical extension of the vandalism review scheme. In that case, with (registered) users, reviewers and super-reviewers, the division of roles within Wikipedia will come to resemble those usually distinguished in OSS. That is, open source design in software on the one hand and encyclopedic content on the other seem to be converging.

Although converging, historically the two kinds of communities have taken opposite trajectories to arrive there. In OSS structure came about mainly as a response to growing technical complexities of code integration. Only recently, project leadership have become aware of the problem that anonymous contributors may be less of a true hacker – and more of a vandal – than originally supposed. In response, screening is tightened. In Wikipedia structure mainly came about in order to fight vandalism and harassment. At this time of writing, quality concerns are growing as well: can Wikipedia entries ever become reliable enough? The means contemplated to guarantee quality are reviewer roles that steer upon content. As soon as Wikipedia would decide to screen potential contributors (as is usual in Citizendium), the designs in both domains of content would be almost indistinguishable. Wikipedia meets Mozilla. As open source ventures, software and encyclopedias would be managed in similar ways. The user, meanwhile, no longer reigns supreme in either of them. Absolute trust has become a relic of the past. Will they stay on board, or cut their losses and leave?

REFERENCES

Websites mentioned in either text or references were last visited on 15 April 2009.

Adler, P.S. and Borys, B. (1996). Two types of bureaucracy: Enabling and coercive. *Administrative Science Quarterly*, 41(1), 61-89.

Andrews, C.K., Lair, C.D., and Landry, B. (2005). The labor process in software startups: Production on a virtual assembly line? In Barrett, R. (Ed.), *Management, Labour Process and Software Development*, Routledge, London/New York, 45-76.

Coleman, E.G. and Hill, B. (2005). The social production of ethics in Debian & free software communities: Anthropological lessons for vocational ethics. In Koch, S. (Ed.) (2005), *Free/open source software development*, Idea Group, Hershey, 273-295.

De Laat, P.B. (2005). Trusting virtual trust. *Ethics and Information Technology*, 7, 167-180.

De Laat, P.B. (2007). Governance of open source software: State of the art. *Journal of Management and Governance*, 11(2), 165-177.

Fox, A. (1974). *Beyond Contract: Work, Power and Trust Relations*. Faber and Faber, London.

Himanen, P. (2001). *The Hacker Ethic and the Spirit of the Information Age*. Vintage, London.

Holck, J. and Jørgensen, N. (2005). Do Not Check in on Red: Control Meets Anarchy in Two Open Source Projects. In Koch, S. (Ed.) (2005), *Free/open source software development*, Idea Group, Hershey, 1-26.

Jørgensen, N. (2001). Putting it all in the trunk: incremental software development in the FreeBSD open source project. *Information Systems Journal*, 11, 321-336.

Levy, S. (1984). *Hackers: Heroes of the computer revolution*. Penguin, London/New York.

McGeer, V. (2008). Trust, hope and empowerment. *Australasian Journal of Philosophy*, 86(2), 237-254.

Mizrach, S. (1997). Is there a Hacker Ethic for 90s Hackers? Retrieved from <http://www.fiu.edu/~mizrachs/hackethic.html>.

Pettit, Ph. (1995). The Cunning of Trust. *Philosophy and Public Affairs*, 24(3), 202-225.

Robbins, J. (2005). Adopting open source software engineering (OSSE) practices by adopting OSSE Tools. In Feller, J., Fitzgerald, B., Hissam, S.A., and Lakhani, K.R. (Eds.), *Perspectives on free and open source software*, The MIT Press, Cambridge, MA, 245-264.

Sitkin, S.B. and Roth, N.L. (1993). Explaining the Limited Effectiveness of Legalistic "Remedies" for Trust/Distrust. *Organization Science*, 4, 367-392.

Viegas, F.B., Wattenberg, M., Kriss, J., and van Ham, F. (2007). Talk Before You Type: Coordination in Wikipedia. *Proceedings of the 40th Annual Hawaii International Conference on System Sciences*.

Zucker, L.G. (1986). Production of trust: Institutional sources of economic structure, 1840-1920. *Research in Organizational Behaviour*, 8, 53-111.

Table 1. Open source communities and mechanisms used for handling the problem of trusting strangers		
	Open source software	Encyclopedic entries
Assumption of trust (substantial trust)	Appeal on hacking capabilities	Appeal on “encyclopedic” capabilities
Inference of trust (weak form)	Hacker ethic	Wikiquote
Inference of trust (strong form)	Entry requirements	(Entry requirements)
Substitution of trust (from a small to a large amount)	Design (from high-discretion to low-discretion)	Design (from high-discretion to low-discretion)
Examples discussed in more detail	FreeBSD Mozilla ↓	Wikipedia ↓
↓ means: pressure on design to move towards lower discretion		

Ethical aspects of e-Government: Social Actors, Politics and the Digital Divide

Georgia Foteinou & George Pavlidis
Department of Business Administration
University of Patras - Greece

Abstract

In this paper we examine the ethical issues arising by the evolution and deployment of ICTs in the public sector and pose some fundamental questions concerning the value of e-Government systems for the society and the government. To this end, we consider e-government as a sociotechnical system whose structure is studied using the systematic approach. We perform a cross reference between a system's functional, structural and hierarchical views and ethical issues at different levels. As an example we discuss such issues arising in the most popular and successful e-government system in Greece, TAXISnet. We find that this e-service surprisingly exacerbates long standing gender discrimination.

Introduction

The modern society exhibits an increasing dependence from digital means, while ICTs have a significant role at the contemporary social transformations [Laudon & Laudon, 2003; Osorio, 2002]. The way we work, interact and socialize change day by day by discovering new ways and means for communication and collaboration. These technological advances in social interaction lead to what we call

* *Georgia Foteinou* is a researcher at the University of Patras - Greece and a postgraduate student at the University of Oxford – United Kingdom. She is currently working as an expert in development projects funding by EU in the countries of Eastern Europe. Her research interests include ICT strategy in public sector and intergovernmental digital services between European countries.

Professor George Pavlidis has been working as ICT-specialist, director, senior consultant, top-level manager, and a teacher. His record includes publications in international journals and conference proceedings, Springer-Verlag lecture notes in Computer science and books. His main area of expertise and interests includes ICT-management and development issues, software engineering, prototyping of distributed environments, distance learning, large scale management, expert, decision support, strategic and operational information systems for banks, insurance companies, private enterprises and software houses, applied business intelligence, data warehousing, customer relationship management, document and workflow control, technology capacity planning and performance tuning, e.tc.

“virtual society” or “digital society” [Mullen & Horner, 2004; Floridi, 2001]. Inevitably, new roles and standards emerge in this digital society [Huyer & Sikoska, 2003], while its members have no time or space limitations and exploit the potential of new technologies to achieve better quality of life, to save resources and to gain some kind of profit [Floridi, 2001]. However, the potential benefits of ICTs are not guaranteed and new strategies are required for taking advantage of the new opportunities brought by new technologies [Osorio, 2002].

The public sector is obliged to follow the social command by modernizing its services and procedures and offering new services which fulfill existing and new needs. The social map has changed radically in the last few decades in Europe and new forms of social, economical and political interaction have developed. The European Integration, the labour mobility, the migration and other factors have contributed to the creation of a new multi-cultural and multi-national society, where new technologies have a major role. This society has completely new requirements and needs, while older rules and solutions do not always cover the reality which is formed through the use of information technology. There are many examples of how ICTs change our social and political life: Immigrants who were threatened by social exclusion have now the ability to communicate and socialize with other people from the same ethnic minority in every neighbourhood, city or country just by creating a blog. People with disabilities can have a direct approach to other people or services. They can communicate and express opinion publicly and they can overcome practical difficulties in their lives. Elderly people, formerly suffering by loneliness have now the ability to meet friends electronically, to speak with their family, to shop, to ask for services. Knowledge is available to every one and there are limited barriers in communicating ideas. A new situation is formed and it is evident that ICTs create new opportunities and new challenges. The public Administration itself has undergone many significant changes in recent years which raise ethical issues, while the evolution and deployment of ICTs creates even more ethical issues which seem more complex in this particular area [Bilhim & Neves, 2005].

Contribution

This paper examines the ethical issues rising by the evolution and deployment of ICTs in public sector and poses some basic questions concerning the value of e-Government systems for the society and the government. There are contradictory views about the potential effects and outcomes (i.e. the dot-gov phenomenon) of e-government and it is obvious that a systematic examination is required in order to achieve a broader overview. The ultimate goal is to provide a framework for answering some basic questions, such as: Do the investments in ICTs in public sector create the required public value or is just a trend? Do we consider

technology in public sector as value neutral or value loaded? Is the technology itself enough for justifying all these huge investments? Is it morally right to assess e-government with e-business criteria? Do we need other methods of assessment? For this purpose the system approach is adopted and a model for the examination of ethical issues related to e-government is presented.

Moreover, this paper poses the basic argument that existed laws and regulations do not always cover the sophisticated needs of e-citizens and we have to reconsider some basic assumptions concerning e-Government. For this purpose a case study of the most successful Greek e-Government system is presented. The case of Greek taxis-net raises significant ethical questions about what we call digital divide and requires a clear delimitation of what is legally right, what is ethically right and what is the role of politics in e-Government.

E-government as sociotechnical system

In general terms, a sociotechnical system is a combination of human actors and technological elements in an integral system, where their mutual interdependence creates the need for a balance between the two [Symons & Walsham, 1988, Sitter et al., 1997]. Hence, a sociotechnical system describes an interlinked mixture of human actors, technology and their environment which interact in order to achieve a goal or to produce a specific result [Fountain, 2000]. Thus, the ICTs in public sector are considered as a component of a more complex sociotechnical system [Gil-Garcia & Pardo, 2006; Mumford, 2000]. Although the initial approach to e-government was mainly technical, at present, the sociotechnical perspective implies that a holistic, interdisciplinary approach of e-government is needed, because the interdependence between technology and human actors creates intended and unintended consequences [Cherns, 1976]. This may partly explain the high rates of failure of e-government projects which according to Heeks reach the 85% in developing countries [Heeks, 2003], while some authors argue that this failure is due to a lack of understanding of the context of use. Each technological change may produce unplanned consequences in an organization. So, we have to think the organization, its sub-systems (where one of them is the information system) and its hyper-system as an integral whole. In our case this hyper-system is the political system.

The system approach of e-government allows us to map the structure and the functions of the system, the system's dynamics as well as to recognize the main actors and the relationships among them and how the e-government changes the system's stability by creating new interrelations [Fountain, 2000]. The parts of the system have structural, as well as functional and hierarchical interrelations and interact with their environment [Ropohl, 1999]. Thus, the e-government

is an open, multilevel system. This means that information flows into the system, out of the system [Ollhof & Walcheski, 2002] and within government levels. Moreover, the system approach helps us to think the broadest framework in which e-Government is nested.

The e-Government is developed and deployed in a specific context, serves specific purposes prescribed from its hyper-system, creates new interrelations by connecting different parts of other systems and finally changes the distribution of power [Orlikowski, 1992; Markus, 1983]. The following figure shows a hypothetical e-government system which connects three public organizations and it is nested to another hyper-system which is the public administration and this, in turn, is embedded to another wider system which is the political system and the society.

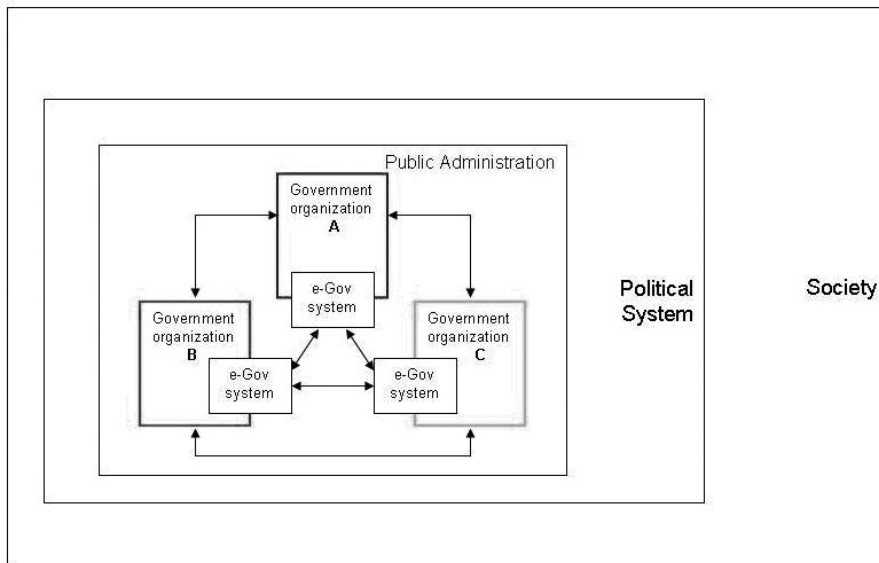


Figure 1. A hierarchical view of an e-Government system.

According to general system theory each system has its own goals and also it attains the goals of the system in which its use is embedded [Ropohl, 1999]. This means that the system which is presented in figure 1, has its goals (for example the automation of a process), it attains the goals of the organizations in which its use is embedded (for example the facilitation of a service delivered by the organizations A, B, C). Moreover, it attains the goals of the political system (for example greater efficiency and effectiveness in public administration), and in a higher level of abstraction it attains the purposes of a broader system which is the society (for example greater transparency, better response to citizens' needs,

social welfare). We suppose here that the society choose the political system and the ultimate goal of the public administration is to serve society.

Additionally, the system approach helps us to specify the broadest context of use and to recognize the “power” actors. In our case it leads us to identify that there are three different organizations, with three different sets of goals, sometimes contradictory to each other, three different organizational structures and a hyper-system which has direct control over them.

A system approach

The system approach of e-government allows us to specify its functions, its attributes, its constituent parts, and the relations among them. The e-government is a distributed Information System that encompasses diverse elements and functions. Each element is an entity that interacts with, or is related to at least one other entity in order to produce a specific result. An entity is something that has a distinct, separate existence, like a person, an organization, a hardware unit, a department, a project or even a task [Chen, 1976]. Each entity is associated with n other entities with a relationship in order to produce a specific result. A relationship is defined as an association between two or more entities [Batini et al, 1992, Chen, 1976], while each entity has a role in a relationship. An entity that has an active role is called actor. The entities are classified into entities sets according to their attributes. By this way it is possible to classify different users' groups as well as the power actors of the system.

There are three different views of a system:

The structural, is the basic, where a system is characterized not only by its elements, but also by the relationships among them. The structural view of the system offers an overview of the systems' actors and relationships between them, what constitutes the system, and how these relationships and the roles of e-government actors' change in time.

The functional, where a system is an entity, sometimes called “black box”, which transforms inputs into outputs, depending on specific internal states. The kind of transformation is called a “function”. The functional view of the system can help us to have an overview of how the system works, in what way, and what are the main functions. It describes the system's functionality.

The hierarchical, where the elements are regarded as subsystems and the original system is considered as a subsystem of a more extensive hyper-system. It is used to describe the different systems in which e-government is embedded and also the hierarchical levels of the public administration. Public administration is still based on the bureaucratic organizational model.

Therefore, a hierarchical model which can also prescribe the organizational structure of different public organizations is the most appropriate. Moreover, we can see who is responsible for the system and the context in which this system is embedded.

The general view of a system, which encompasses the three aforementioned views, is described by a basic mathematical form which, is:

$$S = \{E, R\}$$

Where S is the system, E is the total of entities and R is the total of relations among these entities [Ropohl, 1999].

Ethical aspects of e-government

Some authors argue that there is a lack of fully developed models of ethical behaviour in e-environments [Mullen & Horner, 2004]. If that is true then a significant threat for the democratic processes and social justice, due to a failure to develop appropriate ethics for a digital society, surfaces [Mullen, & Horner, 2004]. This scenario is even worst in the case of e-government, because the evolution of technology in government structures affects not only the public administration itself, but also the citizens' lives, the social fabric and the level of trust in government. In addition, we have to consider the influence of the political decision making and the government practices over society and economy. The government laws and rules define or, in some cases imply, what is ethically correct and very often reflect and affect the deepest beliefs of a certain society. For example in Greek municipal registries the family records are based on father's name and details. In contrast, a married woman belongs to her husband's registry. Moreover, a child could have her/his father's or mother's surname as family name, but he/she cannot have her/his mother's name as middle name. Although the gender equality is legally established in Greece, this kind of practices show that the patriarchic model of family is still present not only in social life but also in government procedures.

It is definitely true that most of these practices originate from past decades, with different values and rules and just represent antiquated habits of public administration which no one politician decided to change until today. In this case a main question which technological evolution rises is: should these norms be retained or should they be altered in a digital structure in Greek registries? Because the preservation of these norms in the digital age will eternalize the social inequalities and will probably create new ones. It is obvious that the system designer has no right to change these procedures unless a politician undertakes the responsibility. This example demonstrates that e-government is closely linked to specif-

ic political actors' decisions. We cannot consider ethical issues in e-government without analyzing the social and political situation and certainly we cannot use e-business criteria for assessing e-government systems [Stahl, 2005]. E-Government transactions have a binding character for the citizens while e-business transactions are based upon a person's choice. Moreover, the evolution of ICTs in public structures raises significant ethical questions in various levels and thus a systematic way of thinking is required.

A multilevel perspective

A theoretical model for the ethical issues related to the use of information systems in public sector is presented in the next paragraphs. The ethical issues rising by ICTs in public sector are analyzed on the individual, organizational, political and social level. Some of the moral dimensions of e-Government systems, such as system quality and accountability and control, are common with other IT systems [Laudon & Laudon (2004)], but e-Government have additional moral dimensions, because of the public and democratic nature of governments [Chircu, 2008]. Subjects such as e-inclusion and e-participation are of great importance in e-Government. Besides, it is necessary to have in mind, that the size and scope of information in e-government is huge compared to those of private sector. The government is the greatest producer, collector and user of information. In the following Table a classification of the basic ethical issues related to e-Government for each level of analysis is presented. This table also shows what system view best describes each issue.

Table 1 – Interrelations between ethical dimensions and system's views

System's View → Ethical issues ↓	Functional	Structural	Hierarchical
Individual level	Personal Data Protection versus Surveillance	Privacy policies & Rules Definition	Access Rights to Services and Personal Data
Organizational level	Horizontal versus Vertical integration	Accountability, Liability & Control	Bureaucratic models versus Networked Governance
Political level	Decentralization versus Centralization	Trust & Transparency	Power Distribution versus Leadership Coordination

Social level	Innovation & Flexibility versus Reliability & Trust	Digital divide versus Development Cost	e-Democracy versus e-Totalitarianism
---------------------	---	--	--------------------------------------

Individual level

The concept of Democracy puts an emphasis on individuals' freedoms, while the respect of a person's preferences and special characteristics is a priority in democratic societies [Hayek, 1994]. Thus, technological policy in public sector should not be exempt from the norms of democratic governance [Laird, 2003]. The protection of a person's freedoms and personal life is of great importance in e-Government, but unfortunately, very often, personal data are used by government bodies without for diverse purposes. The ICTs nowadays makes the surveillance and control by governments easier than ever [Homburg, 2008]. There are many reasons why governments use personal data collected by various sources, the question is: do they have this right? If the answer is yes, than we have to define, under what circumstances and what rules is this practice acceptable. Any rules and regulations must be known and acceptable from all. The citizens concern on privacy and confidentiality of the personal data collected by government is closely linked with the achievement of a higher level of trust to e-Government systems by public body.

Another main concern at this level is technical and is related to the degree of the system's security and confidentiality [Reddick, 2005]. The safeguard of the transmitted information and personal data demands technical solutions and a clear privacy policy [Darrell West; Layne & Lee, 2006]. Hence, from a Functional point of view the main issue at individual level is the personal data protection and the elimination of the possibility for surveillance and control by government actors. This means that it is necessary to clearly define: who has access rights to these data, from what services (applications) and organizations and for what purposes.

Organizational level

E-Government is a great opportunity for public organizations' improvement and reform [Silcock, 2001], while most of the changes which e-Government brings in public administration are realized in organizational level. Thus, many ethical issues must be addressed in this level whereas the traditional regulations and organizational hierarchy may not be sufficient in network working places [Bilhim & Neves, 2005]. The core issue at this level is the transformation of street-level

bureaucracy to a non-bureaucratic model of governance with different values and normative behaviour [Reddick, 2004].

The face of the public administration has changed radically in the last few decades. The previous bureaucratic (Weberian) model has given way to a more complex organizational reality in public sector, where public agencies and divisions do not just offer services directly to citizens, but also are part of a grid of inter-organizational, intergovernmental and multisectoral relationships. Although the traditional hierarchical model of government is still the base of the public structure, isn't enough to produce value-added services, because the value which is produced through the collaboration of different public agents is greater than the sum of what each lone agent can produce [Goldsmith, 2004]. The ICTs have a main role in this collaboration and offer the tools for effective communication, interaction and knowledge and data sharing.

Another core issue, closely connected with the political actors is the question of the borders of functional integration among interorganizational systems. Some authors like Scholl and Klischewski, (2007) stress the importance of the decentralized character of government and argue that an extensive functional integration across e-government levels (vertically) would violate basic democratic principles and result in a vague organizational structure with limited vertical accountability [Homburg, 2008]. Therefore we have to determine what level and kind of integration we have to reach, in which services and in what way. The automation and integration of processes has a risk of losing vertical accountability. Hence we have to redefine some basic rules of the processes (not only the way) and certainly we have to redefine the goals. Moreover, we have to specify: who is responsible/liable/accountable for the processes right now and after integration has completed, who has control over the process (which system/organization/person/department)? Who will have control after integration? What are the purposes prescribed from the hypersystems (social, political, institutional etc)? Are there any contradictory goals?

The following figure represent a simplified view of what we call vertical and horizontal integration among two different government organizations.

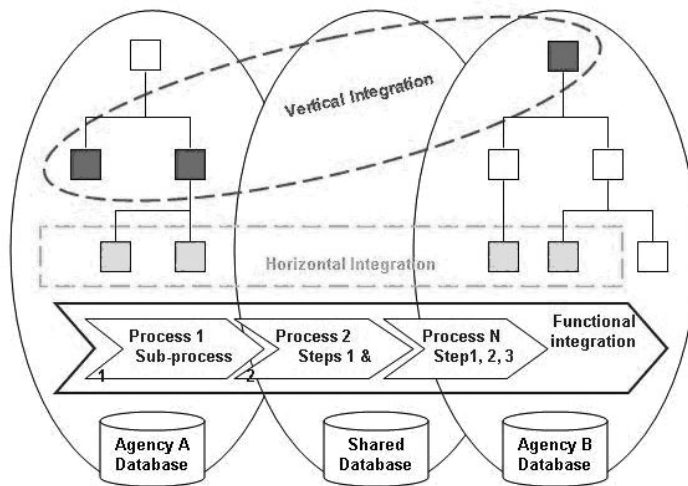


Figure 2. Vertical and Horizontal Integration between government agencies

Political level

As, we already mention, the option of a more centralized government control, because of the use of ICTs and the integration of various services, is a core organizational and political issue that requires balance between decentralized innovation and flexibility, and centralized coordination and leadership [OECD, 2008]. Moreover, the achievement of greater transparency and accountability of public administration without loosing the decentralized character of government is another core matter as well. The role of political actors in the stabilization of these matters is crucial because the politicians are considered as power actors because they have a determining influence on decisions concerning e-Government [Hofkirchner & Fuchs, 2003].

Another important ethical aspect of e-Government concerns the concept of digital democracy. One reason why e-Government is considered important for the society is its potential to enhance democracy. Technology is of course democratically neutral, but the way we use it determines if it is a tool for or against democratic processes [Silcock, 2001]. It can be used for encouraging democratic participation and social debate or, contrarily, as a tool for surveillance and control over the internet. There are examples of both situations in a number of countries [Homburg, 2008]. When the political actors' incentives are not clear, then a lack of trust to public IT systems is emerging.

At the social level

Social actors are very often circumspect about using e-Government services, while e-commerce systems have gained their trust in a satisfying level. They are skeptic in using e-Government systems and very often don't even know anything about it. At the same time, the formal researchers are focused on the supply side of e-Government; therefore, the causes of this attitude are not known. However it is important the prospective benefits of e-Government to be available to every citizen. Unfortunately, it is apparent that these benefits are not evenly distributed between social groups.

The core ethical issue at the social level is the digital divide and its implications. The equal access of all citizens to digital services has both social and political dimensions. A number of demographic factors, such as gender, family income, age, and region affect the e-Government usage by individuals. Recent researches in this area reveal that the equal access to the digital services is still infeasible and the digital divide remains a main obstacle for the e-Government take-up [Eynon & Margetts, 2007]. In the most countries the percentages of female users is substantially lower than men's percentages and this fact reveals that the digital divide has an additional aspect: the woman's position in our society. This is particularly true for the less developed countries and the countries with problems to many other sectors. Hence, we can talk about gender digital divide in many countries in many parts of the world [Huyer & Sikoska, 2003].

An additional ethical aspect of e-Government concerns the cost of these systems for the society. Huge government budget it is devoted to this purpose and an analysis of impacts and outcomes is required in order to answer a basic question: why this budget, is it worth?

A case-study: Taxis-net and gender discriminations in Greece

TAXISnet (taxation information system) is the most successful Greek e-Government system which offers services directly to the citizens through a web site. A variety of services concerning taxation issues are fully available electronically to the public while the system exploits existed information infrastructures [Stamoulis et al., 2001]. It has extremely high rates of usage in enterprises (which reach the 90% of Greek enterprises) and the highest rates in citizens, comparing to other Greek e-services. Actually, this service is the only well-known e-Government service in Greek population and probably its great performance is the reason why e-Government services in Greece have satisfactory usage rates, compared to other EU countries (although they are still under the average EU percentages).

TAXISnet had an overall budget of 60m euro and it was funded by the Greek Government and the European Union. Until today it remains an efficient and well running information system which saves millions of euros every year for the Greek government [Stamoulis et al., 2001].

However, a recent evaluation regarding the social aspects and the citizens' satisfaction of TAXISnet revealed that the system has many weaknesses. This is due to a lack of support for people with disabilities, for immigrants, for foreigners and for other social groups [Terzis & Economides, 2006].

This evidence coupled with some important remarks concerning the Greek tax law, which was fully implemented in the case of TAXISnet without any changes for embracing the current social situation and values, unveils that this successful e-service promotes social discriminations in many ways, especially gender discriminations.

According to Greek tax law no married woman has the right (or obligation) to declare her personal income or property to tax office (unless she owns a business). The husband is alone responsible to declare his personal income and the family income (which includes the woman's income), as well as his property and that of his wife. Greek women pay of course their taxes, but through their husbands' tax declaration. The most impressive is that only the man's details are mentioned at the section of "taxpayers details", while the woman's details are only mentioned as "taxpayer's spouse" to another section of the tax declaration form. The man alone handles all the tax affairs of the family, while the woman pays her taxes but without having any right to handle her personal tax issues. The husband is still considered by the Greek government to be alone responsible to declare the family income, even if he has no personal income or property himself.

As a result Greek married women have no actual access to TAXISnet. This is the paradox of TAXISnet: every married woman has "access rights" to the service but these access rights are reversed by the Greek tax law. At the same time, the statistics shows limited usage of e-Government services of women in Greece. These figures are presumable because the "system" itself excludes women from it.

The following table shows the percentage of individuals (Male and Female) aged 16 to 74 who use e-Government services in recent years.

**Table 2. E-government usage by individuals (demand side)
total and by gender (Total, Male, Female)**

Percentage of individuals

geo/ time	2002A00			2003A00			2004A00			2005A00			2006A00			2007A00			2008A00		
	T	M	F	T	M	F	T	M	F	T	M	F	T	M	F	T	M	F	T	M	F
AT	11	13	9	20	23	16	21	24	19	29	33	25	33	39	27	27	32	23	39	45	33
BE										18	20	16	30	33	28	23	26	20	16	18	14
BG							5	6	5				8	9	8	6	6	7	8	7	10
CY							11	14	8	11	12	11	13	13	12	20	21	19	16	17	14
CZ							7	8	5	5	5	4	17	18	16	16	17	15	14	15	14
DE	17	20	13	26	30	23	33	37	30				32	36	29	43	47	39	33	37	30
DK	37	43	31	40	45	35	44	49	39				43	50	36	58	62	55	44	46	42
EA										25	29	22	27	30	24	33	36	30	31	33	28
EE							20	19	22	31	36	27	29	30	28	30	29	32	34	31	36
ES													25	28	22	26	29	24	29	32	26
EU15										26	29	23				34	37	31	32	34	29
EU25										23	26	20	26	29	23	32	34	29	29	32	27
EU27										23	26	20	24	27	21	30	33	28	28	30	26
FI	34	35	32	40	42	38	45	46	44	47	47	47	47	50	44	50	51	50	53	53	53
FR													26	28	24	41	42	40	43	43	44
GR							8	10	6	7	8	6	9	10	7	12	14	9	10	13	8
HU							16	18	14	18	17	18	17	18	16	25	25	25	25	23	26
IE							14	14	14	18	18	17	26	27	24	32	34	31	27	29	24
IS				56	60	51	58	63	54	55	61	50	61	65	56	59	63	54			
IT										14	17	12	16	19	13	17	19	14	15	17	13
JP				18																	
LT				7			10	9	10	12	12	12	13	12	13	18	17	19	20	18	21
LU	16	21	12	28	33	23	45	55	35	46	58	35	46	57	35	52	62	41	48	55	41
LV							13	12	14	13	12	15	25	25	25	18	16	20	16	15	16
MK													15	19	11						
MT										19	20	18	17	19	16	25	28	21	20	19	20
NL										46	53	38	52	61	42	55	61	49	54	61	46
NO				43	50	36	37	44	31	52	58	46	57	61	54	60	65	55	62	67	57
PL							13	13	12	13	13	12				15	15	15	16	16	16
PT							13	14	11	14	16	12	17	19	14	19	22	17	18	21	15
RO													3	3	3	5	6	5	9	10	9

RS																4	6	3			
SE	42	47	37	44	49	39	39	40	37	52	56	47				53	55	50	52	54	50
SI							13	13	13	19	19	19	30	33	28	30	29	31	31	31	32
SK							25	28	23	27	28	26	32	35	29	24	23	24	30	29	31
TR							6	8	4	6	8	4									
UK				21	24	18	22	24	20	24	27	22				38	42	34	32	35	29
US				23																	
geo/ time	2002A00			2003A00			2004A00			2005A00			2006A00			2007A00			2008A00		
	T	M	F	T	M	F	T	M	F	T	M	F	T	M	F	T	M	F	T	M	F
AT	11	13	9	20	23	16	21	24	19	29	33	25	33	39	27	27	32	23	39	45	33
BE										18	20	16	30	33	28	23	26	20	16	18	14
BG							5	6	5				8	9	8	6	6	7	8	7	10
CY							11	14	8	11	12	11	13	13	12	20	21	19	16	17	14
CZ							7	8	5	5	5	4	17	18	16	16	17	15	14	15	14
DE	17	20	13	26	30	23	33	37	30				32	36	29	43	47	39	33	37	30
DK	37	43	31	40	45	35	44	49	39				43	50	36	58	62	55	44	46	42
EA										25	29	22	27	30	24	33	36	30	31	33	28
EE							20	19	22	31	36	27	29	30	28	30	29	32	34	31	36
ES													25	28	22	26	29	24	29	32	26
EU15										26	29	23				34	37	31	32	34	29
EU25										23	26	20	26	29	23	32	34	29	29	32	27
EU27										23	26	20	24	27	21	30	33	28	28	30	26
FI	34	35	32	40	42	38	45	46	44	47	47	47	47	50	44	50	51	50	53	53	53
FR													26	28	24	41	42	40	43	43	44
GR							8	10	6	7	8	6	9	10	7	12	14	9	10	13	8
HU							16	18	14	18	17	18	17	18	16	25	25	25	25	23	26
IE							14	14	14	18	18	17	26	27	24	32	34	31	27	29	24
IS				56	60	51	58	63	54	55	61	50	61	65	56	59	63	54			
IT										14	17	12	16	19	13	17	19	14	15	17	13
JP				18																	
LT				7			10	9	10	12	12	12	13	12	13	18	17	19	20	18	21
LU	16	21	12	28	33	23	45	55	35	46	58	35	46	57	35	52	62	41	48	55	41
IV							13	12	14	13	12	15	25	25	25	18	16	20	16	15	16
MK													15	19	11						
MT										19	20	18	17	19	16	25	28	21	20	19	20
NL										46	53	38	52	61	42	55	61	49	54	61	46
NO				43	50	36	37	44	31	52	58	46	57	61	54	60	65	55	62	67	57
PL							13	13	12	13	13	12				15	15	15	16	16	16
PT							13	14	11	14	16	12	17	19	14	19	22	17	18	21	15

RO														3	3	3	5	6	5	9	10	9
RS																	4	6	3			
SE	42	47	37	44	49	39	39	40	37	52	56	47					53	55	50	52	54	50
SI								13	13	13	19	19	19	30	33	28	30	29	31	31	31	32
SK								25	28	23	27	28	26	32	35	29	24	23	24	30	29	31
TR								6	8	4	6	8	4									
UK				21	24	18	22	24	20	24	27	22					38	42	34	32	35	29
US				23																		

Source: Eurostat

The rates in Greece are extremely low to both categories but the percentage of women who use the internet for interacting with government is significantly lower than this of men (38% less women users). Moreover, the percentage of the Greek women users of the internet is the lowest in EU27 and is hardly reaches the 23%, while the European average is 57% [Observatory, 2008]. It is clear that a number of factors affect the internet usage of women in Greece, certainly one of them is the denial of their right to handle personally their tax affairs. It is obvious that the patriarchic model of family is still dominating the Greek Society and it is also true that this model is still promoted by the Greek Government in some ways.

The way the TAXISnet works is legally right, according to the Greek tax law, but at the same time, it raises ethical and social issues, because it practically excludes every married woman from the service, and promotes gender discriminations against them. If someone asks a married woman in Greece if she uses a digital government service, probably will answer no. Then it is possible to start talking about digital divide, because women use the e-government services less than men do. But what about the political decision about it?

The Greek Ombudsman and the General Secretariat for Gender Equality have repeatedly asked the Greek Ministry of Economy and Finance to change this practice, but nothing changed for years. It is obvious that the matter is not legal only; its principal nature is ethical. Why the Greek Ministry does not change this regulation? Why a digital service eternalizes an ethical problem? Who is responsible after all?

Someone will answer that is a politician's fault and someone other will argue that is a systemic shortcoming, not an intended one. There are many contradictory opinions, but the main concern has to be; what the government does with citizens' taxes? Are they spent in an appropriate way? Do the e-government services really fulfill the society's needs? Is the huge cost of e-government always justified? Does it really serve the social welfare? Is it contradictory to the social values?

The consequences, positive and negative, of a government policy may be huge. Can we consider Greek TAXISnet as a successful digital service? It has of course substantial economic benefits for the citizens, but it is contradictory to principal ethical values. The citizens do not “feel equal” when they access the service, because what someone can or cannot do is depended on his or her gender.

The main conclusion is that the aforementioned e-service exacerbates the existed discriminations and we can even suppose that it creates new ones by preventing young, intelligent women from using e-Government. The potential female users of e-services are wealthy, well educated women who, until recently, had never felt what really means gender discrimination. Moreover, this case study indicates that many factors can produce biased statistical results, because of a lack of understanding of the context of use. In our example we are not sure that the Greek women do not really want to use TAXISnet. There is not this choice for them.

REFERENCES

- Bilhim J. & Neves B.**, (2005). “New ethical challenges in a changing public administration”, Paper presented at conference Ethics and Integrity of Governance: The first transatlantic dialogue, 2-5 June 2005, Leuven.
- Cherns, A.** (1976). Principles of sociotechnical design. Human Relations vol. 29(8) pp. 783-792
- Darrell M. West**, (2007). “Global Perspectives of E-Government” chapter in “Governance and Information Technology: From Electronic Government to Information Government” edited by Mayer – Schonberger V. & Lazer D., MIT Press, Cambridge.
- Eynon R. & Margetts H.**, (2007). “Organizational Solutions for Overcoming Barriers to eGovernment”, European Journal of ePractice, No.1
- Fountain J.**, (2000). Building the Virtual State: Information Technology and Institutional Change, Chapter 6 – Enacting Technology (Prepublication draft), Brookings Institution Press, John F. Kennedy School of Government, Harvard University, Cambridge.
- Gil-Garcia, R., Pardo, T.**, (2006). Multi method approaches to understanding the complexity of e-government, International Journal of Computers, Systems and Signals vol. 7 (2), pp. 3-16
- Goldsmith, S.**, (2004). “Governing by Network: The New Shape of the Public Sector”, Brookings Institutions Press, Washington, USA

Heeks R., (2003). "Most e-Government-for-Development Projects Fail: How Can Risks be Reduced?", eGovernment Working Paper Series, IDPM, University of Manchester, UK

Hayek, F. A., (1979). *The Political Order of a Free People*. Law, Legislation and Liberty, volume 3, Routledge & Kegan Paul: London

Hofkirchner, W. & Fuchs, C., (2003). "The Architecture of the Information Society", In: Wilby, J., Allen J.K. (eds.), *Proceedings of the 47th Annual Conference, ISSS (The International Society of the Systems Sciences)*, ISBN 0-9740735-1-2

Homburg, V., (2008). *Understanding e-government: Information systems in public administration*. Routledge, New York

Huyer S. & Sikoska T., (2003). "Overcoming the Gender Digital Divide: Understanding ICTs and their Potential for the Empowerment of Women", Instraw research paper series No. 1

Laudon K. & Laudon J., (2004). *Management Information Systems – Managing the Digital Firm*, International Edition, Prentice Hall, New Jersey

Layne, K. & Lee, J., (2001). Developing fully functional eGovernment: A four stage model, *Government Information Quarterly* 18(2) pp. 122-136

Laird F., (1993). "Participatory Analysis, Democracy, and Technological Decision Making" *Science, Technology & Human Values*, Vol. 18(3) pp.341-361

Markus, L., (1983). "Power, Politics and MIS Implementation," *Communications of the ACM*, 26, pp. 430-444

Mumford, E., (2000). "A socio-technical approach to systems design", *Requirements Engineering* 5(2), pp. 125 – 133

Mullen H. & Horner D., (2004). "Ethical Problems for e-Government: An Evaluative Framework", *Electronic Journal of e-Government*, vol. 2 (3), pp. 187-196

Orlikowski W., (1992). "The duality of Technology: Rethinking the Concept of Technology in Organizations", *Organization Science* 3(3), pp. 398-427

Ollhoff, J. & Walcheski, M., (2002). *Stepping in Wholes: Introduction to Complex Systems*. Sparrow Media Group, Inc., Eden Prairie, USA (2002)

Osorio, C., (2002). "Public Ends by Digital Means: some thoughts on e-government and the creation of public value", NSF digital government workshop, White Paper. Available at: www.umass.edu/digitalcenter/events/2002Workshop/Papers/osorio_white.pdf

Reddick C., (2005). "Citizen interaction with e-Government: From the streets to servers?", *Government Information Quarterly* vol. 22, pp. 38-57

Ropohl, G., (1999). "Philosophy of socio-technical systems", *Techne: Research in Philosophy and Technology* 4(3), pp 59-71

Scholl, H. & Klischewski, R., (2007). E-government integration and interoperability: Framing the research agenda. *Intl Journal of Public Administration* 30 pp. 889-920

Silcock R., (2001). "What is e-Government", *Parliamentary affairs*, vol. 54, pp.58-101

Sitter, L.U. & H.J.F., Bankbaar, B., (1997). "From complex organizations with simple jobs to simple organizations with complex jobs", *Human Relations* 50(8) pp 497-534

Stahl B., (2005). "The Ethical Problem of Framing e-Government in Terms of e-Commerce" *The Electronic Journal of e-Government* Vol. 3(2), pp 77-86, available online at www.ejeg.com

Stamoulis D., Gouscos D., Georgiadis P. & Martakos D., (2001). "Revisiting public information management for effective e-government service", *Information Management & Computer Security*, vol. 9(4), pp. 146-153

Symons V. & Walsham G., (1988). "Evaluation of Information Systems: A Social Perspective", *Information Technology for Organizational Systems*, pp. 204-210

Terzis N. & Economides A., (2006). "Internet – based tax filling", *Proceedings of the 1st Intern. Conference in Accounting and Finance, ICAF*

United Nations, Economic & Social Affairs Dept, E-GOVERNMENT SURVEY 2008: From e-Government to Connected Governance

The Great Catopticon

Jean Gabriel Ganascia*

LIP6 – University Pierre and Marie Curie

Abstract

Since a couple of years ago, some well known computer scientists, such as Steve Mann, argued that the development of information technologies and their capturing of everyday life with webcams and RFID (Radio-Frequency Identification) would not lead to surveillance, but to what they called *sousveillance*, i.e. to a state where everybody would be watching everybody. For them, the diffusion, on the web, of all available information, would prohibit the existence of a central superpower that has the exclusive privilege to watch without being watched. This paper is aiming to demonstrate that the underlying structure of *sousveillance* is a Catopticon, i.e. a mirror tower based architecture. The notion of Catopticon has been defined in opposition to the Bentham's Panopticon, which rationally organizes surveillance. The paper draws some possible ethical consequences of the extension of the Catopticon to the entire planet. It tries to show that a universal Catopticon is necessarily unique in nature and that it can coexist with multiple Panopticons. As a consequence, a generalization of *sousveillance* principles does not prohibit the existence of surveillance societies.

Keywords: Panopticon, Catopticon, *sousveillance*, surveillance, transparency, privacy, EyeTap, JenniCam

Introduction

With search engines, webcams, RFID (Radio-Frequency Identification), mobile technologies, cloud computing and chips embedded cards, all we do, i.e. all our movements, talks, expenses e.tc., may be continuously recorded, stored and exchanged. Many of us fear such archiving of our private life. The general argu-

* Jean-Gabriel Ganascia was first educated in mathematics and physics. Then, he studied Philosophy — “licence de Philosophie” université Paris I (Sorbonne) — and computer science — DEA Paris VI university. He got his “Doctorat d'ingénieur” in 1983 and his “Thèse d'état” in 1987. He was successively named assistant professor at Orsay University (Paris XI) (1982) and Professor at Paris VI University (1988). He acted as a program leader in the CNRS executive from November 1988 until April 1992 before moving to direct the Cognitive Science Coordinated Research Program since January 1993 until 2000. He led for 20 years now the ACASA team in the LIP6 that is the computer science laboratory of Paris VI University.

ment is easily understandable: a Big Brother could use these continuous records of all our personal activities against us and consequently, it would restrict our freedom and our secrecy. However, some scientists don't share this view. For instance, Steve Mann¹ argues that the diffusion of all private information will contribute to establish a new social equilibrium called *equivoillance*, where everyone is watching everyone. These scientists distinguish the classical surveillance, where a few are watching the majority without being viewed, from the regime of *sousveillance*, in which everyone has an equal access to all information about the others.

This view seems to be reinforced by the current development of social networks, like Facebook, where members give freely private information available to everybody. Nowadays, many of our contemporaneous, especially youth and teenagers, don't fear surveillance. They like authenticity. They don't want to hide themselves. They agree providing access to their intimacy and to everything about themselves; in return they demand a total transparency. The success of the Jennifer Ringley's website attests this tendency. This young girl had installed webcams in her student bedroom and then, during seven years, from April 1996 until the end of 2003, she has continuously diffused views of her intimacy on the web. She quickly became very popular. There were more than five millions visits per day on her website, which has been seen as a social phenomenon. Jennifer Ringley² herself has been considered not only as a young exhibitionist but also as a conceptual artist who could anticipate the future state of a society.

The goal here is not to analyze the social web or the Jennifer Ringley success from a social, a psychological or a mathematical point of view, but to understand the philosophical meaning of such evolutions. The aspiration towards transparency corresponds to some sort of utopia, i.e. a kind of ideal view of the society where nothing would be hidden. This utopia is not new; for instance, in the 19th century, Nikolai Chernyshevsky had formulated it in a famous novel entitled "What is to be done?" that subsequently greatly influenced Lenin and many other Russian revolutionaries. Later, in the beginning of the 20th century, many people, such as André Breton in his novel "Nadja", Paul Valéry in "Monsieur Test" or Walter Benjamin in "Experience and Poverty", have also expressed a desire for total transparency. At that time, the extensive use of glass architectures, for instance the "Crystal Palace", designed by Joseph Paxton and built for London Great Exhibition of 1851, seemed to render this total transparency reachable. However, even in glass

-
1. *Steve Mann* got a PhD from the MIT; he is now professor at Toronto University, where he is working on human computer interaction.
 2. The interested reader may refer to the Wikipedia article about Jennifer Ringley - http://en.wikipedia.org/wiki/Jennifer_Ringley.

buildings, the transparency remained restricted to a local area. Nowadays, things appear different. Modern techniques, which are the information and communication technologies, render the achievement of a total transparency feasible at the earth scale. This might greatly affect the structure of the society, as the separation between the private and the public spheres, that organized the society for the last three centuries, since the institution of a legal state, has now become blurred.

By introducing the notion of Catopticon, derived from the Bentham's Panopticon, this paper attempts to investigate the logic of the generalized *sousveillance* that underlies, according to Steve Mann, all these contemporaneous phenomena. The paper is divided into five parts. The first one is dedicated to the introduction of the concepts of *sousveillance* and *equivveillance*. The second part describes the architecture of both the Bentham's Panopticon and the Catopticon. The third part shows how these architectures spread to the entire planet by the generalized use of informational technologies. The fourth part shows the properties and the limitations of these extensions. The fifth and last part concludes on the ethical and political perspectives of which the great Catopticon clarifies the issues.

Surveillance, *sousveillance* & *equivveillance*

Surveillance

According to Steve Mann, the surveillance characterizes situations where a watcher is positioned above the watched, "above" being understood from both a positional and a social point of view. It follows an asymmetry between the watcher, who may use any information he has about the people he observes, and the observed individuals, who have no information about who watches them. Consequently, the watchers dominate the watched, who are under their total control. The aim of the police in totalitarian state is undoubtedly to establish such a regime of surveillance. The 20th century offered many cases of generalized surveillance in totalitarian societies like the Italian fascism, the German Nazism or the soviet Stalinism. Many authors, such as Franz Kafka and George Orwell, exacerbated the logic of the surveillance society. For instance, "1984", the famous George Orwell's novel, depicts the paroxysm of the surveillance society where one individual is continuously watching everybody and everything. Note that the logic of surveillance is not restricted to the gaze, i.e. to the information access. Information dissemination is also one of the characteristics of the surveillance societies in which the diffusion follows the line of sight and where the broadcasting is totally controlled by a central organization that imposes its absolute censure.

Sousveillance

By opposition to the surveillance, the sousveillance characterizes situations where the watchers are placed under the watched. The term sousveillance is a neologism derived from the French word *surveillance*, which is composed of “sur”, above in French, and “veiller”, to watch. By analogy, *sousveillance* is built with the prefix “sous” that means under. Examples of sousveillance are numerous, for instance, citizens who film the police, or civilians who photograph government officials are typical cases of sousveillance. In the present days, the information technologies make everybody a potential broadcaster: it’s enough to have a mobile phone or a webcam, to record pictures and sounds and to put them on a personal weblog or anywhere on the web. Many cheap devices exist that can easily be managed for that purpose. For instance, Jennifer Ringley developed a special webcam, the so-called “JennyCAM” (Jimroglou 1999), giving her the opportunity to continuously broadcast videos on a personal weblog; Steve Mann designed a new device called the “EyeTap”³ worn in front of an eye and acting both as a camera, which captures the continuous visual flows, and as a screen that displays computer-generated imagery, the latter coming from other “EyeTaps” or from any visual recorder. As a consequence, it is now possible for everyone to broadcast the information he wants, when he wants, without fearing any censorship.

In the sousveillance regime, the observer, who is situated under the watched, can’t control him. According to Steve Mann, a new social and political order derives from this physical disposition. More precisely, he argues that, when the sousveillance is generalized, everybody is watched by everybody, which makes impossible for a few to govern or impose their views. Let’s suppose that an incident happens in the street: patrolmen are beating youths. Anyone may record this scene with its mobile phone and diffuse the pictures online earlier than any journalist could, and before the police authority has heard about it. This easy information broadcast is not restricted to surveillance; more generally, it may help people exchanging information about any concern in the everyday life. For instance, in Paris subway, travelers of the line 13, very often delayed, have decided to exchange information with their mobile phones using the Twitter messaging facilities. However, people do not always communicate freely through technologies for the best: everybody remembers when, the end of 2008, during the Mumbai attacks⁴, terrorists used Twitter, Flickr and other social web technologies,

3. Interested readers may read the “EyeTap” home page (<http://wearables.blu.org/>) or the wikipedia article dedicated to the “EyeTap” (<http://en.wikipedia.org/wiki/EyeTap>).

4. Among the many papers published on this topic, the interested reader may refer to the Telegraph (<http://www.telegraph.co.uk/news/worldnews/asia/india/3530640/Mumbai-attacks-Twitter-and-Flickr-used-to-break-news-Bombay-India.html>).

by amplifying the scope of their action through the international medias, and, in a more active manner, by exchanging and obtaining strategic information about the current situation. Whatever our appreciation may be, positive or negative, of the role of communication technologies, from their general use may result a new social organization, where the officials in charge of the administration, the police or other intermediary bodies can be short-circuited more and more easily.

Equiveillance

Note that the logic of *sousveillance* is not restricted to information access and to information diffusion. It also concerns all aspects of the society where the traditional roles are transformed. The authority, i.e. the legal power, is submitted to the domination of the watchers who continuously observe and comment each of their acts. Politics, diplomats, police, physicians and all those traditionally licensed to act confidentially, now need to modify their practice and to act in open air. As a consequence, the role of administration is changing rapidly. Its place becomes less and less central. It does not mean that it disappears, but it is evolving. The bureaucracy understood, by reference to Max Weber (Weber 1969), as the ideal-type of a rational and legal form of domination, is vanishing and is giving way to a new form of administration, which does not act as dominant, but as a partner or a facilitator. Nobody will really complain about the loss of the bureaucracy; however the resulting social organization may generate new offensive forms of domination, without any of the intermediaries to smooth the antagonisms.

According to Steve Mann (Kerr & Mann 2006), there is no such a risk: the generalized *sousveillance* spontaneously conducts to a regime of equilibrium that he calls *equiveillance*. In such a regime, we are all under the permanent observation of all. Everybody is acting under the control of everybody. Therefore, everyone is obliged by everyone. Steve Mann claims that it forces each person to adopt an ethical behavior. In a way, this permanent control of the controller (i.e. of those who have the authority) may result in a new form of transparency. Steve Mann asserts also that the continuous record and retrieval of everyone's personal experiences lead to a state of *equiveillance*, since everybody will be able to get access to all personal archives and consequently to judge everyone's acts with respect to the precise informational context of his or her behavior. It follows, that the development of wearable computer that supplies people with continuous images and sounds capture devices would facilitate the emergence of *sousveillance*.

Steve Mann argues that those techniques contribute to both the empowerment of the individuals and the enforcement of the democracy. According to him, the elites will no more abuse of their power because they will be continuously cleared. Without weakening Steve Mann's originality, note that this idea is not new. During

the French revolution, some people promoted the institution of the so-called “iron mouths”, iron mailboxes where everybody was free to send public opinion on any subject, which was then printed and spread using the new postal techniques. Inspired by Jean-Jacques Rousseau, the advocates of “iron mouths” argued that it supported what they called the fourth power, i.e. the power of the opinion, against the power of the representatives, suspected to act in their personal interest.

Nevertheless, *sousveillance* may have some opposite effects. For instance, the institution of “iron mouths” also made possible an easy slander and calumny of personal enemies. More generally, the *sousveillance* changes the “syntax” of the society, which may introduce confusions. As a consequence, power has more and more difficulties in imposing its will, which may lead to a state of anarchy (by reference to the etymology of the word “anarchy”, which comes from the Greek word *anarkhia* – an- “without” + -arkhia “power”), or an absence of leadership. Let us illustrate this with one example of the possible confusion that can result from the generalized *sousveillance*. It comes from one recent episode of the French political life. In November 2006, a few months before the March 2007 French presidential election, M Alain Duhamel, a politic commentator,, was invited to participate to an academic debate in the Paris school of political sciences. During his intervention, he criticized the campaign of one presidential candidate, M François Bayrou⁵. But, to weaken his critics against the politician, he said incidentally that he would vote for him. It happened that someone made a non-authorized video record of M Duhamel’s intervention⁶ and diffused it on the web some months later, in February 2007. Having being accused to publicly support M François Bayrou, Alain Duhamel was condemned to stop his activities of political commentator in the media during the presidential election campaign. This example shows how the so-called *sousveillance* leads to the confusion between an academic freedom of speech and an official declaration in the public sphere. This could have tragic consequences for us, as academics, if our debates and discussions would be permanently diffused to everybody in the world. More generally, the confusion between the different spheres of reality could really be a nightmare. The nature and the necessity of these differences need to be understood, but, before, we have to understand the structure of this new world. Our aim, in this paper, is to investigate it. As we shall see in the next section, we propose to use and to tweak the architecture of the Bentham’s *Panotipicon*, originally designed for surveillance, for this purpose.

5. *François Bayrou* was one of the candidates to the French presidential election in 2007. He got 18,57% of the votes in the first round, and was positioned just after Nicolas Sarkozy (31,18%) and Ségolène Royal (25,87%).

6. http://www.dailymotion.com/video/x151aw_duhamelvotebayrou

Panopticon and catopticon

The architecture of Panopticon

The architecture of Panopticon has been designed, at the end of the 18th century⁷, by Jeremy Bentham (Bentham 1838) to facilitate the surveillance of prisoners in jails. According to Michel Foucault (Foucault 1975), Panopticon defines a structural schema, which played a key role in modern societies since the establishment of a rational legal fundament of the social order. Its role is to teach, to redress and to amend. Even if there were, during the 20th century, some attempts to generalize Panopticon to the overall social order, for instance in Soviet Union, most modern societies are only in part organized on Panopticon schema. Some key institutions like prisons, schools, hospitals, and asylums are organized on this model. However, Panopticon influences the overall structure of the society, as it threatens all the deviants, i.e. criminals that are put in jail, sick persons in hospitals, crazy people closed in asylums, e.tc.

Briefly,, the Panopticon is built as a ring around a central tower (cf. Figure 1), where observers can see all prisoners' actions. The cells are transparent: they receive and transmit the sunlight in such a way that the inspector may observe every movement of a prisoner without being viewed. In addition, prisoners are totally isolated from each others. To summarize, the three main Panopticon principles are:

1. the total transparency of cells,
2. the fundamental dissymmetry, which makes the observer watch all the prisoners, without being watched, and
3. the isolation of the prisoners who can't communicate each others.

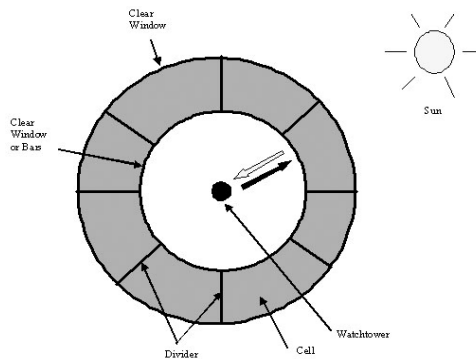


Figure 1: the schematic principle of Panopticon

As a consequence of such a structure, the Panopticon society is basically hierarchic one. The status of those who are located in the watchtower differs strong-

7. The first papers on the Panopticon have been published from 1780.

ly from the status of cell occupants. The first ones watch without any risk; the second ones are continuously submitted to the watch of the controllers, without having any information about who is controlling them. Originally, Panopticon has been designed by Bentham just as architecture for prisons, the goal being to rationalize the surveillance in a utilitarian perspective. Its initial role was to re-educate, to cure and to teach the law; it has then been extended to other social institutions that have been progressively seen as “curative”. For instance, schools, hospitals, army, factories e.tc. have been organized in conformity with this model. Even if Panopticon was not designed as a model for the overall society, it influenced its constitution and most of the social bodies were structured on the model of Panopticon. More precisely, as Foucault says, Panopticon corresponded to a new state of the society where an impersonal law replaces the person of the Sovereign. Panopticon is a symbol of the logic of the surveillance society where few privileged people, i.e. the administration or the occupant of the watchtower, take advantage of their position. It explains the hierarchy of the society, which is no more family or heredity based. More generally, the differences in social status depend on the places that men occupy in different Panopticons that constitute the society.

Catopticon

By analogy and by contrast to the three surveillance principles on which Panopticon is based, there are three fundamental principles of sousveillance that are:

1. the total transparency of society,
2. the fundamental symmetry, which gives everybody the ability to watch – and consequently to control – everybody, and
3. the total communication, which makes everyone able to exchange to everyone.

In practice, it means that there is no hierarchy, since there is no central tower, and that everyone may communicate to everyone in a total transparency. This structure differs radically from the Panopticon schema. The design of an edifice that renders possible such a total communication among its inhabitants is here derived from Panopticon schema.

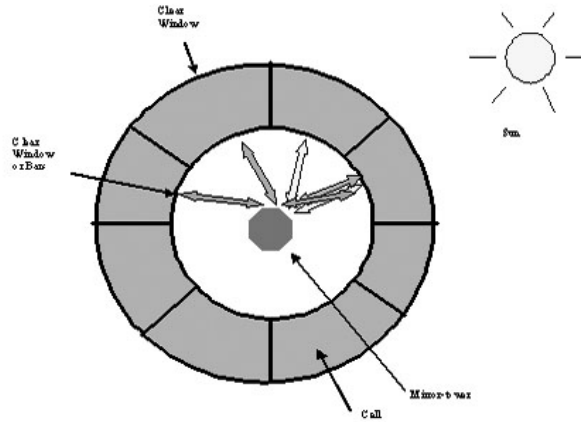


Figure 2: the schematic principle of Catopticon

This architecture of *sousveillance*, which obeys to the three above mentioned principles, replaces the Panopticon watchtower by an empty mirror-tower that renders everyone able to communicate with everyone (cf. Figure 2). It is based on the notion of “reflectionism”; a word invented by Steve Mann to describe the procedures using technology as mirrors against bureaucratic organizations. Since this architecture is a derivation of Panopticon based on mirrors, we decided to call it Catopticon (from the *catoptrics*, the study of light reflection and mirrors).

The two main properties of Catopticon are, first, that it does not generate inequalities, since it does not introduce any difference of status between the occupants, and secondly, that the space is totally transparent.

Extensions of catopticon

Panopticon was designed to be a building. Consequently, it is geographically located and limited to at most, a few hundreds of meters length. The Catopticon architecture that is derived from the Panopticon architecture supposes, implicitly, that it is also restricted to a building. As a consequence, its size is also limited. With the information and communication technologies, the structure of both Panopticon and the Catopticon may be expanded in huge proportions. Our aim, in this section, is to examine such extensions.

Geographical Extension

Modern information technologies, e.g. webcams or wearable computing, render now possible the extension of Catopticon to the global human society throughout the entire planet. More precisely, new devices, for instance the Jennifer Ringley’s

JenniCam (Bailey and Kerr 2007) or the Steve Mann's EyeTap, are designed to record continuously personal information and to retrieve it anytime, anywhere, throughout the globe. As a consequence, it is possible for anyone to get information about anyone, which corresponds exactly to the principle of *sousveillance* on which is based Catopticon.

Let us remark that, simultaneously to this extension of Catopticon to the global society throughout the entire planet, Panopticon itself may be considerably extended. Nevertheless, we can also prove that, by nature, Panopticon cannot be extended to the entire society, since it has been designed to correct and to reform people and it is – at least in principle – transitory.

Extension to the Virtual World

In parallel to its physical extension to the entire planet, Catopticon has also been enlarged to the infosphere: not only human, but other informational organisms – the so-called *inforgs* (Floridi 2008) – belong to Catopticon. More generally, a modern Catopticon is part of a virtual world built on the infosphere. Equipped with artificial intelligence techniques, *inforgs* may freely communicate among each others; some of them are humans while others may be artificial intelligent agents, virtual robots or chatterbots. All human beings, artificial intelligent agents and other *inforgs*, belong to Catopticon. Nevertheless, even if Catopticon can be extended to the infosphere, it is not sure that it contains all the infosphere. This point needs a further discussion, which goes beyond the scope of this paper.

Note that Panopticon may also be extended to the overall infosphere with the use of new information technologies. But, the meaning of both extensions, the extension of Panopticon and the extension of Catopticon, differ. Since the inhabitants of the Panopticon periphery, i.e. the cells, cannot communicate to each others, no matter the presence of artificial agents there. In contrast, artificial agents that would have been admitted in the watchtower could act as efficient controllers. These agents could continuously check that the activities of the inhabitants of the periphery are conforming to the rules. Consequently, it would considerably decrease the amount of work of the official warder. In the future, one could even imagine that for the sake of equality nobody would allow anyone, except artificial agents, to act as controllers.

The logic of universal catopticon

The extension of Catopticon to the entire planet, to all the humanity and then to the overall infosphere defines what we call universal Catopticon. Our postulate is that the concept of universal Catopticon can help elaborating an ethics of information on a solid foundation and argue against both rapid technophilia and old

technophobia. In this section, we show the uniqueness of the universal Catopticon and the possible coexistence of the universal Catopticon with multiple local Panopticons. Before providing a sketch of such demonstrations, let us indicate the logical framework that we deploy to formalize Catopticon and Panopticon. We first consider that both Catopticons and Panopticons are composed of locations, the L_i , and are inhabited by inforgs, I_j , which may be either humans or intelligent agents. Usually, each inforg I_j live in one or more locations L_i , which is characterized by the binary predicate $\text{location}(I_j, L_i)$. Each inforg can develop the locations he is living by placing there any content C_p he/she/it accesses and likes, which means that " $I_i I_j C_p \text{location}(I_i, L_j) \cup \text{access}(I_i, C_p) \cup \text{like}(I_i, C_p) \otimes \text{contain}(L_j, C_p)$ ".

Proof of the uniqueness of the Universal Catopticon

In Catopticon, everybody gets access to everything that is contained in the locations he/she knows, which means that: " $I_i I_j C_p \text{knows}(I_i, L_j) \cup \text{contains}(L_j, C_p) \otimes \text{access}(I_i, C_p)$ ". Let us now suppose that there exists two extended Catopticons that we call A and B for convenience. Either A and B are connected or not. If they were not connected, it would mean that there would not exist anyone belonging to A who had access to B and vice-versa, which is contradictory with the planetary extension of the universal Catopticons. If A and B were connected, then there would exist at least a location L_j that belongs simultaneously to A and B. Since they belong to A, all inforgs of A inhabiting L_j may add new contents to it. As a consequence, L_j may contain all the contents accessible to inforgs of A. As they simultaneously belong to B, all inforgs of B knowing it have potentially access to its contents, i.e. to the contents accessible to the inforgs of A. In conclusion, all inforgs of B have potentially access to all the contents accessible to inforgs of A and, by symmetry all inforgs of B have access to the contents accessible to inforgs of A. As a consequence, the extended Catopticons A and B cannot be distinguished any more, because they have naturally merged.

Since everything is accessible to everybody, communications are mainly public and transparent through exhibitions in the different locations L_j of the extended Catopticon. However, we don't live only in a public space: we exchange many private emails each others daily. Therefore, we have to take into account those private exchanges in our model. It can be done by adding a predicate $\text{send_message}(I_i, I_k, C_p)$ that characterize the private exchange of information between inforg I_i and I_k . This addition does not affect our demonstration by any means. As a matter of fact, if two extended Catopticons A and B, coexist simultaneously, they necessarily merged.

Proof of the Multiplicity of Extended Panopticons

As previously stated, in Catopticon, individuals may have access to every place they know. In Panopticon, things happen in a different manner since everybody is assigned to a fix location that belongs either to the periphery or to the watchtower. The watchtower inhabitants may freely communicate with every inhabitant of the periphery, but the inhabitant of the periphery can only communicate with watchtower inhabitants. As a consequence, the way the people communicate depends on their location. Watchtower inhabitants have access to all the contents accessible to the inhabitants of the periphery, i.e. "Ii"Lj"Cp"Wk"Im location(Ii, Lj) $\dot{\cup}$ access(Ii, Cp) $\dot{\cup}$ watch_tower(Lj, Wk) $\dot{\cup}$ location (Im, Wk) \otimes access(Im, Cp). In contrast, for the inhabitant of the periphery, the access to the center is submitted to a preliminary authorization of the inhabitants of the watchtower: "Ii"Lj"Cp"Wk"Im location(Ii, Lj) $\dot{\cup}$ watch_tower(Lj, Wk) $\dot{\cup}$ location (Im, Wk) $\dot{\cup}$ access(Im, Cp) $\dot{\cup}$ authorization(Im, Ii,Cp) \otimes access(Ii, Cp). As a consequence of this requirement, the union of two Panopticons A and B is submitted to the authorization of the inhabitants of both the watchtowers of A and B. Therefore, many Panopticons may subsist separately when the inhabitants of their respective watchtowers don't agree to merge.

Coexistence of the Universal Catopticon to Multiple of Extended Panopticons

The last point refers to the coexistence of the Universal Catopticon with multiple Panopticons. As a matter of fact, nothing prohibits the simultaneous existence of the Universal Catopticon with multiple extended Panopticons. Without going into details, some dictators, like Kim Jong-il in North Korea, succeed in isolating some areas from the rest of the world, which in this case physically prohibits any access to the network and consequently any participation to the great Catopticon. It may also happen that some states control and restrict the access to the network, as it is now the case in China. But, in parallel to those geographically circumscribed Panopticons, there exist some virtual Panopticons that exercise their influence on their members. Multinational companies or religious sects are examples of such networks that compel their members to secrecy and to a total obedience.

In conclusion to this section, note that, contrary to Steve Mann (Mann et al. 2003) who pretends that sousveillance allows to "surveil the surveillers", we claim that the emergence of some new hierarchies that may impose a local totalitarian power should always be feared. This power is not necessarily geographically localized, as it was in the past; now, multinational companies or mafia can perfectly do it with a delocalized global world. Simultaneously, Catopticon opens

some new perspectives from a political and social point of view that have to be explored.

The limitations of the universal catopticon

The extended Catopticon, where everyone exchanges to everyone, fits in to an ideal figure, which can't be fully achieved. There are several practical limitations that restrict the communications among the humankind. For instance, human cognitive abilities are imperfect and limited, and it is not possible for an individual to permanently access to the personal archives of the 6 billions humans. As a consequence, each of us restricts his or her attention to few person. Therefore, those who have the ability to capture the attention – and to become idols – get an incredible advantage. The opposite is also happening: while, in the modern bureaucratic time that was, according to Michel Foucault, characterized by the schema of Panopticon, the watchers got the power, now, in the information society, the lead comes to those who are watched.

As an illustration of these new relations of domination, let us consider the logic of the net economy: it is based any more on the utility of the provided goods, as was the old economy, but on publicity. In other words, no matter the goods and their utility; goods are not more than a pretext to make a site popular, because profitability is mainly based on popularity. To measure the efficiency of advertisements it is usual to evaluate the audience of the website where it is published. Therefore, the more a site is visited, the more advertisements that are published there are of value and the more the site yields profits. As a consequence, today, one of the most valuable activities is not to produce goods, but to artificially increase the search engine site index, i.e. to generate what is called “spamdexing”. Many techniques exist for this, for instance the creation of link farms that are clusters of highly interconnected websites. The site index being computed with respects to the number of references, it is increased by the number of hyperlinks that point to him. As a consequence, those who want to artificially increase their site indexes, pay for being mentioned in some link farms.

In addition, intelligent agents, avatars and other virtual beings are also members of the infosphere; consequently, we may also exchange with them, which adds again some supplementary confusion. Not only those inforgs increase the number of potential information sources, but also they are built to make illusion. It even appears that, for multiple reasons among which some are ethical, they may have either to hide information or to lie. Therefore, the total transparency, which appears to be a pre-requirement to Catopticon, cannot be totally achieved.

Conclusion and perspectives

As it was previously shown, the notion of universal Catopticon helps to discuss the fundamentals and the limits of privacy, which are undoubtedly deeply modified by the generalized use of information and communication technologies. It can also be used to understand some aspects of the modern political order, where politicians have less and less ability to decide by themselves and where people is submitted to new digital servitudes. Note that the principles on which Catopticon is based – i.e. the strict equality, the absence of hierarchies, the end of sovereignty, the perfect communication and the total transparency – seem to be currently accepted by almost everybody in the modern developed countries. Even if it does not totally correspond to the exact state of the present information society, it defines an ideal schematic structure that shapes the contemporaneous social and political order in the exact same way as Panopticon was a schematic model on which many aspects of the social institutions of the modern age, e.g. prison, asylum, hospital, school, e.tc., were shaped. Our goal, in this paper and in the future, is to explore our contemporary social order through the structure of the universal Catopticon.

The existence of the universal Catopticon raises many questions. Among them, one is relative to the separation between the public and the private sphere, on which the modern societies were based since the imposition of a legal state. Note that, before this separation, the person of the Sovereign had a total access to everything that concerned his subjects, including their family life, their thoughts and their beliefs. With the end of absolute monarchy, the legal state introduced separations between the private and the public sphere. Today, many people, especially young generations, would like to see the private spheres invade all the public space. The popularity of social networks is one of the symptoms of this evolution. The notion of *sousveillance*, theorized by Steve Mann is another sign of that transformation. It seems that this means the end of privacy, i.e. the end of the separation between the individuals and the society. The Catopticon that was presented along this paper tends to enlighten and to formalize those evolutions. Therefore the open question is: does the advent of the universal Catopticon means the end of the legal state and the beginning of a new political order?

As a conclusion, let us express our last point by reference to the famous George Orwell's novel: "Is "1984" behind or before us?"... Arithmetically 1984 seems to be behind. People like Steve Mann tell us that the *sousveillance* technologies offer guarantees against the surveillance society. To this respect, it is also definitively behind. But, from a strict logical point of view, the study of Catopticon

shows that nothing prohibits the coexistence of Catopticon with multiple Panopticons and the future subsistence of Panopticons in the 21st century.

REFERENCES

Bailey, J. and Kerr, I. (2007). The experience capture experiments of Ringley & Mann, *Ethics and Information Technology*, Springer Netherlands, Volume 9, Number 2 / July 2007, 129-139

Benjamin, W. (1934). *Selected Writings*, Volume 2, 1927-1934 Translated by Rodney Livingstone and Others Edited by Michael W. Jennings, Howard Eiland, and Gary Smith

Bentham, J. (1838). *Panopticon or the Inspection House*, The Work of Jeremy Bentham, volume IV, 37-172

Floridi, L. (2008). Information Ethics, its Nature and Scope, in: Jeroen van den Hoven and John Weckert (eds.), *Information Technology and Moral Philosophy*, Cambridge University Press, Cambridge

Foucault, M. (1975). *Surveiller et punir*, Gallimard, Paris, France, p. 252 – In *English Discipline and Punish*, trans. A. Sheridan. (1977) New York: Vintage.

Jimroglou, K. M. (1999). A camera with a view: JenniCAM, visual representation, and cyborg subjectivity, *INFORMATION COMMUNICATION AND SOCIETY*, VOL 2; NUMB 4, p.p. 439-453

Kerr, I. and Mann S. (2006). Exploring Equiveillance, http://www.anonequity.org/weblog/archives/2006/01/exploring_equiv_1.php

Mann, S. (1998). 'Reflectionism' and 'diffusionism': new tactics for deconstructing the video surveillance superhighway. *Leonardo*, 31(2): 93-102.

Mann, S., Nolan, J., Wellman, B. (2003). *Sousveillance: Inventing and Using Wearable Computing Devices for Data Collection in Surveillance Environments*, *Surveillance & Society* 1(3): 331-355, <http://www.surveillance-and-society.org> - <http://wearcam.org/sousveillance.pdf>

Weber M. (1969). *Economy and Society*. Edited by Guenther Roth and Claus Wittich. New York, Bedminister Press

Virtual Decisions: Just Consequentialism, Video game ethics, and Ethics on the fly

Don Gotterbarn*-James Moor**

East Tennessee State
Dartmouth University

Abstract

Applied ethics needs to address constantly emerging policy vacuums in information technology. In many cases the ethical issues that arise are easy to identify and address because of their similarity to previously identified and analyzed situations. For example some video games, such as the Grand Theft Auto series are notorious for their rewarding of racial, sexual, nationalistic discrimination, and savage behavior. There is no debate that the behaviors depicted are unethical. Video games are effective training tools for learning a variety of skills and approaches to problem solving. Studies have shown that both negative and positive behavioral and cognitive patterns are learned. Significant psychological studies of the negative impact of prolonged playing of such video games indicate an increase in aggressive thoughts.' (Anderson & Bushman 2001).

Gotterbarn (2008) identifies a more fundamental problem for ethical analysis and decision making.

A major family of video games trains gamers that their decisions are all and only about themselves. The most banal description of the problem is to say, in a non-pejorative way, that the ethical decisions in video games seem to be 'self-centered'. It is not that the gamer is encouraged to think of themselves and others but that the gamer is encouraged to think solely in terms of benefit to their character when making a decision. The decision making driver is not the impact of the decision on society but rather the quantity of rewards for their individual character. The focus is on the individual's success in winning the game, accumu-

* *Don Gotterbarn*, the Director of the Software Engineering Ethics Research Institute and a visiting professor at the Centre for Computing and Social Responsibility, has worked on numerous software projects. He was awarded the «Making a Difference» award by the ACM's SIGCAS and in 2005 received the Outstanding Contribution award from the ACM for promoting ethical behavior of professionals and organizations.

** *James Moor*, is the Daniel P. Stone Professor in Intellectual and Moral Philosophy at Dartmouth College. He writes widely on computer ethics and the philosophy of AI.

lating objects or 'successful' good; crimes. We are training individuals to make decisions without any attention to the effects of their decisions on others or on the larger society

The problem is that for so many of these games the only ethical obligation is to oneself. The "decision" standard is selfishness. Why not have more games that require some ethical decision making beyond egoism? The failure to engage in any ethical analysis is frequently justified in the virtual world by appealing to a perceived urgency for a decision (the need for a fast interactive game) so that an alternative rapid decision making process is substituted for a more careful ethical analysis.

This model of selfish egoism is being applied to decisions beyond the virtual world. In military situations in principle the use of lethal force is governed by ethical constraints called Rules of Engagement (ROE) and Laws of War (LOW) (Slim 2008). In the Surgeon General's study (Surgeon 2008) on ethical behavior by the military it was determined that the majority of soldiers using Patriot missile systems did not step through the ROE but merely fired without consideration of the ROE, LOW, and potential collateral damage. This is similar to the videogame ethical decision model described above and justified in a similar fashion, namely claiming that the time for reasoned decision making is too short to make intelligent informed decisions in modern warfare. (Surgeon 2008). An alleged external consequence of playing these video games is that they cause crimes such as the Columbine school massacre. Law suits have been filed against game companies charging them with partial responsibility for the 14 year old Michael Carneal's 1997 killing of three students at Heath High School in Kentucky.

The proven negative educational impacts of video games requires the development of a more appropriate model of ethical analysis that goes beyond this narrow approach and can be applied in a relatively rapid fashion. Video games should/could be designed to reward success at using a better ethical analysis model. A better decision making model is needed in video games so players can test and develop skills in rapid ethical analysis.

There are numerous approaches to ethical analysis. We want our decisions to be ethical, but what should we look for when constructing ethical decisions? When we turn to traditional ethical theories for help we discover a strong rivalry exists between the leading contenders – consequentialist theories that emphasize the consequences of actions and deontological theories that stress rights and duties. We look at more unifying theories which can be used in the kind of extreme situations described above requiring both rapid and effective ethical analysis.

Moor (1999) developed a theory of just consequentialism which coherently unifies deontological and consequentialist aspects of ethical analysis. The theory of just consequentialism emphasizes consequences of policies within the constraints of justice making just consequentialism a practical and theoretically sound approach to ethical problems of computer and information ethics.

In this paper we adapt 'just consequentialism' and apply it to very rapid ethical decisions, showing how it can be used to introduce values into video game design and how it can be used in decision making beyond virtual world.

Keywords: Computer ethics, Just Consequentialism, Video Game Ethics, E-Game, Ethical=align Egoism

Virtual ethics

Ethics matters. There are many reasons for this, but at the very least it matters because we do not want ourselves and those about whom we care harmed unnecessarily or treated unjustly. Living in an ethical society is likely to be more pleasant and satisfying than living in an unethical one. Numerous concerns have been raised about video games and their impact on how we address ethical issues. In the very complex interactions that take place between players, game designers and developers and the game, we believe there is a potential danger that has been overlooked but can be mitigated by the application of some philosophical theories in the development and design of these games.

The Good, the Bad, and ...the Questionable

The widespread use of video-games in education, training, entertainment and the use of video game technology in the design and control of medical, commercial, and military systems has a significant impact on the present and the future directions of society. As in other rapid advances in technology, the speed of the development and immersion of this technology in our lives has left numerous significant questions unanswered.

Unlike Pong generation games, the current video games attract attention because in many of them players can, and are sometimes encouraged, to perform actions that are morally proscribed in the real world. Many researchers have examined the alleged negative real world consequences of encouraging such actions while others focus on the positive aspects of video games.

Single user strategy or button mashing games, where the user plays against the machine, were transformed by advances in technology into multi-player competitive games. As the Internet developed so did the games. They transformed into multi-user team structured games where some games are played internationally

by thousands of players simultaneously. The group nature of some of the play has led to claims that these are socializing games.

The Good

Without question some video games, used in health care and education have positive effects. They assist patients' rehabilitation from surgery and injuries. Advanced technology makes it possible to develop educational games wherein players study historical events by replaying history and engage in 'what-if' analysis of major decisions. Because of their ability to improve hand and eye coordination video games are employed in a variety of sectors from flight training and surgery to military training. They are also used to improve critical thinking, creativity and encourage an exploratory approach to problem solving. The US Defense Intelligence Agency is using video games to train its spies and soldiers. These games are not haphazard single shooter games, but they place agents in situations "to teach them how to think under pressure, how to reason, and how to use violence only as a last resort." (Gamingsteve, 2008).

The Bad

In some games one can, and is sometimes encouraged to perform actions that are morally proscribed in the real world. Some video games, such as the Grand Theft Auto (GTA) series, are notorious for rewarding of racial, sexual, nationalistic discrimination and savage behavior. There are significant psychological studies of the negative impact of prolonged playing of such games, indicating an increase in aggressive thoughts. Most work on the ethics of video games seems to focus on the potential for reinforcement of specific behaviors. For example, after stealing a car in the real life (RL), Devin Moore shot three people, two of whom were police officers. His behavior has been attributed to his playing hundreds of hours of Grand Theft Auto.

The Questionable

The influence of these games is illustrated by this news account.

One blistering afternoon in Iraq, while fighting insurgents in the northern town of Mosul, Sgt. Sinque Swales opened fire with his .50-cal. That was only the second time, he says, that he ever shot an enemy a human enemy. "It felt like I was in a big video game. It didn't even faze me, shooting back. It was just natural instinct. Boom! Boom! Boom! Boom!" remembers Swales. The chief of Defense Modeling and Simulation officer for the army said. "When the time came for him" -- meaning Swales -- "to fire his weapon, he was ready to do that. And capable of doing that. His experience leading up to that time, through on-the-ground training and playing 'Halo' and whatever

else, enabled him to execute. His situation awareness was up. He knew what he had to do. He had done it before -- or something like it up to that point.” (Washington Post 2006).

Motivation, practice, rewards, and ever increasing tasks helped this soldier transfer skills learned from video games to the real world.

Claims versus Research

The debate rages about the accuracy of such claims. Some argue that the events in Grand Theft Auto are no more real than the events in a Flight Simulator. These are ‘internal’ events. They argue that there are no real-world violent ‘external’ consequences. Only ‘fictional violence’ is occurring. They admit some limited external events related to game playing such as hitting the keyboard in frustration and some extended consequences such as enhanced reflexes in critical situations. The primary concern with video games is that they may have real world external violent consequences.

There are ethical issues raised about video games in general and against violent games in particular. Unfortunately this same behavior impacting power can be used in non-positive ways to subtly manipulate opinion. “America’s Army is the first computer video game to make recruitment an explicit goal and the first well-known overt use of computer gaming for political aims. The game is used as a playable recruiting tool and critics have charged the game serves as a propaganda device. America’s Army was developed by the U.S. Army iteself (sic).” [Wikepedia] Like any device which can be used in training, video games can be used to mold ideas either in a positive or negative way.

The discussion of violent video games has drawn some interesting responses. Some have argued vigorously that calling video games ‘violent’ is mere ethical emotivism. Tavinor [2007] argues that the events which occur within videos games are not relevant to the ethical evaluation of the game.

The apparent violent, sadistic, and otherwise criminal events that occur within games cannot be factored into the consequentialist account for the very simple reason that the worlds and events of video gaming are fictional. Grand Theft Auto, for example, has repeatedly been condemned for allowing its players to perform acts of theft, assault, murder, and worse. But these apparent actions are fictional ones, and really there are no such things involved in the game. Grand Theft Auto, and similar games, might be thought of as crime simulators, in that similar to flight simulators, they allow their players to indulge in immediately non-consequential behavior that pursued in reality can be quite dangerous.

Tavinor goes on to claim that separating this game playing behavior from its normal consequences is a pre-requisite for a player's ability to enjoy it. "if what was fictionally occurring in the world of Grand Theft Auto was genuinely occurring, the player would not be enjoying it quite so much!" This is an interesting assertion which is easy to find exceptions to, see for example the account of the Jake Baker Case (Johnson 2001) where after describing the torture, rape, and murder of a student said, "[j]ust thinking about it anymore doesn't do the trick...I need to do it".

One can easily find empirical studies supporting either side of this issue. In one study, "the authors found that respondents who had more exposure to violent video games held more pro-violent attitudes, had more hostile personalities, were less forgiving, believed violence to be more typical, and behaved more aggressively in their everyday lives." [Psychology and Crime News 2008]. Compare this with the 2008 CNET News article on Harvard video game research concluding "that there's no data to support the notion that violent video games cause the kids who play them to act out violence in real life."

The Debate versus the Consensus

Settling the debates about increased aggression, game addiction, acting out horrific game scenarios are beyond the scope of this paper (and our skills). These claims about real world negative consequences of encouraging such actions and the positive aspects of video games in education and training need to be adjudicated by trained specialist in psychology, sociology, e.tc. These clearly contentious claims about the impact of video games have at least one common element; namely, that these games are able to change our ideas about the world and how we act in it.

The Consensus

Although there is significant debate about how the influence of the video game is instanced in particular behavior in real life (RL), there is no disagreement that they do influence behavior. Some have tried to make an absolute distinction between RL and the virtual world. It is interesting that while denying any evidence for the eternal extended effects of 'so-called' violent video games, Tavinor accepts evidence that "video games are beneficial in terms of learning and literacy." He cites work by Patricia Greenfield emphasizing the importance of "the necessity of induction in discerning the patterns and rules involved in game play, the tracking of the interaction of multiple variables, and the development of spatial skills." When talking about the 'non-violent games' Tavinor accepts that games do have at least limited external effects that seem to affect the way we approach what is good. This observation is consistent with Kant who rejects the abuse of

animals because it will lead the abuser to become accustomed to, or desensitized about extending that abuse to rational beings. In fact Waddington (2007) finds two Kantian grounds for concern with video games. Namely, when “video games involve acts of cruelty, those acts violate our duties to ourselves” and “second video game characters, like animals, may be analogues of humanity. If we do not treat human analogues with respect, it may make us less likely to perform our duties toward other human “beings.

Video games are frequently adopted as educational tools. The reviews about the relative effectiveness of learning by videogame or RL face to face interaction have been mixed but “evidence has shown that learning from e-games can change attitudes” (de Freitas 2008). There is no question of their impact in RL.

Computer ethics needs to address constantly emerging policy vacuums in information technology. (Moor, 1985) In particular, computer ethics now needs to address the ethical policy vacuums in the design and development of video games.

Real Virtual Reality – Addressing Ethics in Video Games

One approach to the variety of issues mentioned above and to ethics in general has been to design games whose specific goal is to teach ethical principles. In one game developed in Thailand, the sole purpose of the game is to teach Buddhist ethics (Kaewmorakot 2007). The “Ethics Game” aims to indirectly teach players about morals, doing good and the five Buddhist precepts. The game involves three children who have to follow a monk on a pilgrimage. There are many barriers they have to face during the journey were only intelligence, goodness and morals (not weapons and force) help get them past the barriers. The five precepts of Buddhism are: do not kill, do not steal, do not commit adultery, do not lie and do not drink alcohol. These are included in the final stage of the game when the four characters must teach villagers about the precepts and instruct them if their behavior has gone against any of them.

The attempt to teach ethics through videogames is not limited to teaching religious concepts. Thompson (2008) discusses several videogames designed to introduce ethics into video games. There are games like Modern Prometheus designed for use in ethics classes where students can discuss their ethical choices. Playing the role of Dr. Frankenstein’s assistant the player must address questions like should they steal brains out of the local cemetery to help the Doctor develop a cure for the plague or politely leave the dead to rest and soon die themselves from the plague? Other games like the third-person adventure game Fable define certain actions as good or evil and do not depend upon classroom discussion. In Fable the player is presented with situations requiring intervention and the player is then told the morality of those choices.

Although in some of these games the ethical choices are superficial; the programs encourage evaluating choices and the broad consequences are considered. But they do not present multiple models of decision making.

Self-focused decision making as the only kind

All of these discussions about ethics are important.. We do not minimize these issues but instead want to focus on an issue which at best is under addressed. This new issue is not one of intentional harm or propaganda designed into video games but more of oversight or lack of awareness and so the issue is harder to correct. This issue does not involve conscious manipulation through video game design. The problem is one of well trained practitioners designing a learning pattern into the games with impacts they may not have considered. We do not attribute any ill motives to the designers and developers.

It is about me and my stuff: Self-focused decisions

There is a family of video games based on rapid decisions that are related to the success of the game player. These include role playing games, real-time strategy games, games whose success is determined by the number killed, or to use the euphemism of the games “the number K.O.ed”. The Xbox and Game Boy generation of students raised on these games are being trained that decisions are all and only about themselves and the materials they can end the game with. He/she who has the most wins.

The most banal description of the problem is to say, in a non-pejorative way, that the development of e-games seems to be ‘self-centered’. The gamer is encouraged to think solely in terms of benefit to their character when making a decision. Thus in ‘Gears of War’ which mimics dangerous military situations, the only reason to save a wounded teammate is that the action will improve your chances of winning. The motivation is not loyalty or concern for the other soldier’s life. In Grand Theft Auto one is rewarded for the number of acts done in service of a master criminal. The driving question is not the impact a decision has on society but rather the reward impact on the individual character. The focus is on the individual’s success in winning the game, accumulating objects or successful good crimes. This single-focus approach is true of most button mashing shoot-up games as well. We are training a generation to make decisions without any attention to the consequences for others of their actions.

This ethical decision making standard is a form of ‘egoism’. We could define ‘egoism’ as a form of decision making where one considers their own interests. This is contrasted with ‘altruism’ where one considers the interests of others “rather than” one’s own interests. James Rachels calls this ‘social responsibility’ where

the outcomes of one's decisions are designed to benefit society rather than one's self. This approach to decision making justifies a sacrificing to give a wide variety of service to one's group, or video game team. (Rachels, 1999) It is important to note that 'egoism' does not preclude service to others. Friedman has argued "One can make egoism work in a society as long as one does not induce harm to others and follow "the basic rules (laws) of society." (Friedman 79).

There is a broad range of weights one can give to their own interests and weights one can give to the interest of others in decision making. For example some people help others because it will help others and because it makes the doer feel good. Our concern is that the 'egoist model of decision making fostered in these games is an extreme form of egoism where one consider ONLY their own interests rather than the interest of others- a selfish-egoism. The games encourage one to concentrate only on oneself. The games encourage an ethical arrogance in decision making. What is good for me is good. This becomes even more problematic in some games where a motivation of self-interest violates traditional ethical values. In the game Bioshock the player encounters very large generally benign beings called Big Daddys who need drugs to avoid feeling severe pain. They are accompanied by "Little Sisters" who supply them with drugs- Big Daddys protect little sisters. The player has 3 options to interact with this duo:

1. Kill sister to get drugs and the player gets more power – but this increases pain in the world for the BigDaddy
2. Free little sister from her spell so she no longer serves BigDaddy – player does not get as much power
3. Leave them alone, but you die since no drugs.

It is in your selfish-egoism interest to select the option that creates the most pain. It is not considered an ethical issue if you create more pain

Providing other models of decision making such as providing a potential for balancing self-interest and other-interest would mitigate the problem of mistraining in ethical decision making. This kind of balance is consistent with philosophical positions like just consequentialism. Just consequentialism in part is based on the view that we use a set of core human values with a conception of justice to make ethical decisions. However in many video games core values are ignored or worse yet contradicted and they are not addressed in decision making. For example Fallout 3, achieving different levels makes you eligible for 'Perks'. "With the "Bloody Mess" perk, characters and creatures you kill will often explode into a red, gut-ridden, eyeball-strewn paste. Fun! Oh, and you'll do 5% extra damage with all weapons." http://guides.ign.com/guides/882301/page_2.html

We do not think, in general, the player is aware of what is happening and how the repeated decision making style affects their non-game behavior. Players are not aware of the nature of the decision method - 'good' game design leads them to use particular standards defined in terms of the actions to win the game and the structure of the reward in the game.

This is not one of intentional harm or propaganda designed into video games; it is the more invidious issue of well trained practitioners designing a learning pattern into the games with impacts they have not considered or have been trained to recognize

The Virtuality Fallacy

A fallacious form of reasoning, which we call the Virtuality Fallacy goes as follows: Some X (a thing or activity) occurs in virtual space and virtual space by definition is virtual and therefore not real. Therefore, X or the effect of X is not real.

The virtuality fallacy is easy to commit because something that is virtual is likely considered intangible and ephemeral and hence something that cannot be real or have real effects. Of course, a little reflection shows that the line of reasoning is fallacious. A promise made in an e-mail is made in cyberspace but is a real promise. A cyberattack on a defense department computer is a real attack. Criticizing an immature, awkward teenager through a social network site may have a malicious and tragic real life outcome.

Decision Models in Some Best Selling Games

"Gears of War" is a third person shooter game where the player manipulates a character/avatar representing the player in the game in which players working in groups of up to 8 soldiers armed with assault rifles equipped with chainsaw bayonets try to save the inhabitants of a planet by wiping out the Locust Horde. In this game if a teammate is injured it is important to bring them along with you because it will increase your chances of winning. The basis for the decision is "will it help me win". You do not save your teammate out of compassion. There is no sense of personal sacrifice in the game to improve the lot of society.

In the "World of War Craft" groups form to go on quests. There is an approach to this game where a player assumes the role of a "griever" who will join a weaker group that is going on a quest under the pretense that they are out to help the group. The grievers real goal is to steal everything that the group gains on the quest. Thus armed and stronger the griever goes on to repeat this behavior. The griever eventually achieves the highest level of the game. This anti-social behavior has no consequences in the game.

As we have maintained the claim that game playing never affects the external behavior of the game player is dubious. We can argue whether in individual cases it has limited (immediately after game play) or extended (long duration effects) consequences, but even those who defend violent video games admit the causal influence of less violent educational games. The problem is that many of these games train the decision maker to base their decision solely on what is in their interest. This encourages them to ignore broader consequences. This is not merely the open question of how and whether games cause one to mimic particular horrific actions but it is the problem of learning how to make moral decisions; learning what is the central consideration for our moral decisions.

Brey (1999) for example talks about the impact of these games on moral development, but he like others focuses on the particular actions- torture, murder, etc - which are encouraged in the virtual world. There is another important element in moral development and that is learning how to make moral decisions; what are the central considerations for our moral decisions.

Some have argued that this is not really an issue because of the 'humor' in the decision making. The attempt at humor related to the violence does not lessen the single stakeholder decision. In the e-game Postal, the player, a serial killer trying to escape the police, is awarded extra points for killing different groups of innocent bystanders including school children and people at an anti-game violence protest. The killing of the anti-game group while escaping is claimed to be humorous and it gains you extra points.

What is the problem?

One could object that we have had war games like chess for centuries. What is the difference? A game is a game is a game.

We believe three differences make a difference in the video games. First, video games are much more graphic than chess and other traditional games. Capturing a rook in chess does not compare in terms of the audio and visual input from pil-laging, eviscerating, and murdering with all of their gory details in a video game. Video games will only become increasing realistic as computer techniques improve. The psychological impact of a realistic video game is much different than traditional games. Second, chess and most other traditional games have pieces with known capabilities. Players use the pieces within the constraints on action that are known to all. Pawns cannot suddenly jump four spaces. To the extent that video games are social activities using features whose powers may not be known, whole new kinds of action may arise that are not expected. Gamers take pride in making video games that have secret capabilities and powers that can be discovered by the players. This possibility is reminiscent of what happened in

a non-virtual electronic community LambdaMOO many years ago in which one character was able to take control of others to perform sexually offensive actions. In modern videogames players sometimes perform actions that are not anticipated even by the designers of the game themselves. Rape, as it was performed symbolically in LambdaMOO, may be tame compared to the actions that can and will be generated in contemporary video games. Third, as we have been arguing, virtual immorality is promoted in many violent video games. These are actions such that if humans did them in the real world we would regard them as immoral. For these three reasons we regard video games as different. It is not true that a game is a game is game.

Consider a particularly violent game series “Halo” in which one is encouraged to eliminate opponents. One description under “Pick on weakened opponents” says. “This is also known as “clean up”. Why try to kill someone who is at full strength when you can prey on the weak? Think of it as attacking someone who has their shields worn down, but someone else eliminated their shields for you. The advantage of doing this is obvious. Some players out there might consider this “kill stealing”. Don’t feel bad about doing it - it’s a perfectly legal tactic and the game doesn’t prevent anyone from doing it. If someone whines about it, then they should start doing it too.” (<http://www.halo3basics.com/#level4>). Noticed the standard is. There are no constraints against it so it is acceptable.

Explicit Ethics training Games:

There are video games which are designed to address ethics, but they are primarily games of ethics training or indoctrination.

The “Ultima” series of games for version 1 thru three were shooter games but version 4 changes the nature of the game. Your team of 8 still goes through dungeons but your goal is to perform good deed and increase your virtue/self. You have to follow the “virtues” specified in the game: Compassion, Honesty, Honor, humility, Justice, Sacrifice, and Spirituality. The game has fixed values assigned to specific actions. You get Compassion points: +2 giving to beggar, +2 fleeing from Non-evil enemy, +1 letting Non-evil creatures flee, but -5 for attacking non-evil creatures. The virtues are not of your choosing. There is no judgment necessary. Points are allocated to a variety of actions such as killing or fleeing from an enemy. In interesting ways this is still like the problem we are concerned with. The basis for the decision is purely winning the game and not based on judgment about the impact of your actions.

“The Ethics Game” developed in Thailand is used to bring youth closer to the five fundamental principles of Buddhism - not to kill, steal, commit adultery, tell lies or drink alcohol. “These principles are fundamental to reach the highest level in

the game, where they will then be called upon to teach farmers or peasants the fundamental rules of Buddhism.”

<http://www.buddhistchannel.tv/index.php?id=52,3826,0,0,1,0>

Again there is no real ethical analysis, the player learns the rules, follows them and they win.

Rethinking ethical theory

Traditional Ethics

How can we address this critical issue in video games? Traditional philosophical ethical theories offer ethical guidance. However, the three major ethical theories are inadequate when taken individually and are inconsistent when taken together. Consequentialism tells us to base our actions on promoting good consequences and avoiding bad ones. Deontology tells us to act so as to respect duties and rights. Virtue theory tells us to act virtuously and avoid vice. All are commendable and capture part of ethics. But, each is limited. None of the theories is dominant because each has at least one Achilles' heel. Traditional consequentialism generates results that lead to unjust outcomes. (For example, stealing someone else's computer if one could get better use out of it. In World of Warcraft it is acceptable for a Griefer to turn on his teammates killing them and taking their powers because it will help the griefer achieve a higher status.) Strict deontology ignores consequences that in some situations ought to be of overriding concern. (For example, refusing to tell a small lie to prevent a terrorist sabotage of a smart electric grid.) Strict virtue theory provides little guidance in ethical decision making in novel or complex situations. (For example, implanting an enhancement chip in a child of normal ability.) If we consider the approaches collectively, conflicts arise immediately as the theories are not mutually consistent.

A common misunderstanding or expectation is that an ethical decision making procedure should be in the form of an algorithm such that for every ethical problem the algorithm produces a unique correct answer. This is the problem we have seen in Ultima 3 and in Ethics. There are problems like this in mathematics. For every finite set of natural numbers an algorithm will find the greatest common divisor of those numbers. But as the Gödel Theorems show, this standard is much too high for all decision making even for mathematics. And in science as well there is no algorithm that will guarantee that for any set of data we can determine which is the best theory to explain it. For any interesting set of data points there will be multiple possible theories, usually an infinite number, consistent with it. We want to pick the true theory, but an algorithm will not tell us which one that is. If the expectation of algorithmic decision making perfection is too

great for mathematics and science, it is reasonable not to expect it to apply to ethics or for that matter most of our life decision making.

Ordinary Decisions

How then do we make ordinary decisions? Usually we have a set of standards on which the decision is based. Suppose we are considering buying a new computer. We likely have in mind such standards we wish to meet. Of course, some of these are so obvious we implicitly take them for granted. For example, the computer should work when turned on. And some of the standards are quite explicit and precisely defined. The computer should have so much memory and so much speed.. And it should not cost too much. And so on. When we are deciding on what computer to buy, we have many, sometimes competing standards to satisfy. Typically when applying our standards we will rule out many computers. Some will not have enough memory and some will be much too expensive. We might find one which is obviously better than all the rest, but often we will find ourselves in a situation in which we need to compromise. Given the demands of our standards there may be any number of computers that are acceptable. We can pick a computer, but it may have been reasonable for us to pick another one that we were seriously considering given our standards. In such a case there may be several justified decisions but no dominant decision. Someone else using the same standards could have picked another computer and still have been justified in their choice. This does not mean that any computer would satisfy our standards. It does not mean anything goes. If the computer store only had computers that did not work when turned on, we would have gone to another store. And we may reject all of the working computers as not being good enough, given our standards, to make a purchase. This approach is very different from the guidance free choices made in many video games; choices whose only basis is satisfaction of a narrow ego based goal.

This account of buying a computer is a kind of decision making that we do all the time. The method of deciding is not algorithmic but it is not random either; neither Grand Theft Auto nor Ultima 3 provides adequate models of deciding. The one regards the satisfaction of every desire as good and the other dictates what should be satisfied. We have a procedure in which we have standards of varying degrees of importance and we apply those standards to find acceptable possible decisions from which to pick. Our decision can be rationally challenged and sometimes is. Sometimes we may be persuaded that we made a poor choice given our standards or that we were using inadequate standards. But frequently we decide using defensible standards in judging many good choices so that a number of options emerge as acceptable and reasonable.

Now the question is why not regard ethical decision making in the same way? Rather than see it on the algorithmic model such as finding the unique greatest common divisor for a set natural numbers, see it on the procedural model that considers a range of choices that can be evaluated on a set of diverse and possibly competing standards.

Just Consequentialism

Just Consequentialism (Moor 1999) is an approach to ethical decision making that utilizes many of the insights of traditional theories of ethics but places them within a common sense model of decision making. Just Consequentialism is not merely consequentialism but consequentialism in conjunction with justice including concerns for rights and duties.¹

Consequentialists reflect on human nature and argue for maximizing what people regard as good. Traditional consequentialists pick out happiness in some form or the satisfaction of desires as the good which we all seek. Just Consequentialism takes the view there is a broad set of goods that everyone values but does not regard the satisfaction of every desire as good. Humans obviously do value happiness but also value life, even a life without much happiness. In addition, there are other such goods that at least some people may value more highly than happiness or life. These goods are at least instrumental for everyone but may be intrinsic, goods in themselves, for others. These goods include ability, security, knowledge, freedom, opportunities and resources. These are the goods that everyone ASKS FOR whatever their goals. These eight goods are core goods. Their loss is an evil. Together they make up the core values that all people hold. The claim is that in any culture these core values will hold for the humans in the population. The claim is not that everyone values them at the same level. A scholar may value knowledge or an athlete may value ability above the other core values, even above life and happiness. But the core values are important for everyone. Everyone in every culture values resources to some degree since they must eat and take shelter. Everyone will value security to some degree because they seek protection from danger. And so on. Within different cultures these core values will be articulated differently. One culture may articulate resources in terms of good fishing and another in terms of good wireless connections. The point is that there is a way of understanding human beings and their activities across cultures even when they differ in their manifestations of the values. These values are critical to any ethical decision

1. Many insights for Just Consequentialism have come from Common Morality as developed by Bernard Gert. See Gert (2007)

Just consequentialism puts extra weight on avoiding evil as opposed to doing good. A central just consequentialist maxim is do not harm others. Exceptions need to be allowed but as an initial directive not harming others is a good starting position. We know what the goods and evils are in terms of the core values. These then become one of set of standards in evaluating actions. For each choice of action, we can examine its consequences in terms of the goods and evils produced. In just consequentialism it is assumed that there will be some variation in the ranking of the importance of the goods among people. But there is agreement on what counts as goods and evils and that some rankings of the goods and evils are not rational.

Concerns for justice provide another set of standards for evaluating ethical decisions and actions. Respecting duties and rights is one of these standards. Duties and rights derive from different sources – sometimes from a constitution or a set of laws, sometimes from agreements such as promises and contracts, and sometimes from obligations of roles such as the role a parent, a mayor, or a software engineer. Another important standard of justice used in evaluating ethical actions, particularly in areas of social policy, is distributive justice. We expect a fair distribution of rights and duties as well as benefits and responsibilities in society. There are yet other standards of justice, such as procedural justice, we use to assess ethical actions and policies, but the one we wish to emphasize for our purposes here is impartiality. Impartiality requires that whatever ethical decision one proposes one should be prepared to advocate it as an acceptable ethical decision for anyone to make in a similar situation, not knowing who plays what roles in that similar situation. Thus, another simple ethical maxim is to be fair to others.

Thus, we have two sets of standards we typically bring to the table when making ethical decisions. Some of the standards involve our values used to determine the overall value of the consequences of various possible actions. Some of the standards involve justice to ensure fairness with respect to duties, rights, distribution, and impartiality. When making ethical decisions, not unlike when making other decisions, many factors need to be considered. These are easy to model in video games.

Some examples will be useful. Many of our ethical decisions are easy to make, so easy that we tend not to see them as ethical decisions. Given our two very general maxims we know we need to be fair to others and avoid harming others and given that as humans we have a developed sense of fairness and know what counts as harm, many ethical decisions are made effortlessly. No serious thought or calculation is required. We just do the right action. But sometimes it is not clear and some reflection is helpful. Suppose some one working as a chief programmer for

a company wonders whether to tell one's boss that the program, whose deadline is due, really has not been adequately tested. The program will be used in an important medical application but may not run well. No information that it will not run well is available and it has been tested somewhat, just not fully as required by industry standards. If something goes wrong, the blame could be put on one of the lower level programmers who worked on the part of the program that turned out to be faulty. The consequences of informing one's boss that the program had not been fully tested would lead to a poor report for the chief programmer in a time at which personnel in the company are being laid off. The consequences of not informing one's boss are not certain but that there is risk is clear. There is tension here because of the bad consequences for the chief programmer if he tells and the duty he has in his role to tell. The impartiality standard for justice requires us to ask whether anyone in similar situations would be allowed not to tell. Similar situations are not restricted to computer contexts. One could imagine an engineer deciding about whether to approve a new braking system that had not been fully tested and that the chief programmer would eventually be a buyer of that car. The impartiality standard puts the ethical issue in a broader perspective and requires that the chief programmer inform his boss or at the very least be sure the program is fully tested before approving it.

Looking at ethical decision making through the lens of just consequentialism assists one in making difficult ethical decisions by setting of a framework of standards to consider. It allows and indeed encourages rational discussion among those who may agree or disagree with a possible decision. Obviously, it does not provide an algorithm for deciding all ethical issues. There is no such algorithm. Sometimes consequences will be so severe they will override rights and duties. And sometimes duties and rights will override consequences that are beneficial. Sometimes people will disagree on what should be done. But, there will be large agreement on a huge number of cases. Difficult cases are difficult, and there is a temptation, especially among philosophers, to focus on those to the exclusion of others. But in fact there is major consensus on many ethical issues and consensus can sometimes be established on cases that are difficult.

Just consequentialism and video games

There are many simple ways consistent with just consequentialism to mitigate the potential danger of selfish egoism based decision making. Gamers will say that such inclusion will slow the game or lead to uninteresting games which no one will buy or no one will play. Even though such assertions are not directed specifically at our concern that there is a danger in the current design of games, responding to the challenge of both an interesting and ethically sensitive game will help clarify our position.

The view that we cannot act both ethically and quickly is a mistake. We do it all the time. Many of our everyday actions are both ethical and quick. They are so obvious that we take them for granted. We do not drive willingly into other people's cars or hit people capriciously with canned goods. We have built up habits of behavior that we perform automatically. We develop routine virtuous actions that allow us to operate ethically almost on automatic pilot much of the time. When we do act unethically, such as when lying to another, we typically know immediately that it is unethical. We may still do it, but most of the time it is not because we are lost in some ethical calculation. We consider our duties and the foreseeable consequences of our actions. Unlike traditional consequentialism, just consequentialism does not require infinite, unknowable calculations of the consequences. The basic ethical recipe is simple. Do not harm others and treat them justly. And ideally, help them when you can. We all know the core values from our own lives and so we know what causes harm. And we understand what counts as treating others unjustly because, if nothing else, we have likely suffered it in some form in the past. Therefore, ethical decision making often can be made in a flash. Of course, this is not to deny that sometimes ethics is complicated. Values can conflict as can duties and rights. Facts may be difficult to determine. In these cases details of an ethical theory can be brought to bear in more considered approach to the decision-making. Games that do not include some elements that illustrate this are not as ethically adequate as they could be.

Here are some simple and interesting ways to introduce just consequentialism into video games. It is relatively simple to get players to think of the principle of 'impartiality' and justice. In a game like *Gears of War* a random switch could be introduced that switches a player from one character/avatar to another. Thus a soldier who was deciding whether to help a wounded member of his group might have his role switched to some other avatar in the current game situation. The healthy soldier might end up changing places with the wounded soldier for a short period of time. The same randomizing function could be used in the way rewards are allocated. Helping a fallen comrade may not increase the likelihood of winning the game so there is some unrewarded risk in going to the wounded soldier. Helping others may not always result in a gain. The same random switching of roles in *Worlds of Warcraft* will generate a broader consideration of core values when deciding how to behave as a Griefer.

Another option is to introduce distinct consequences for actions with different impact. In the *Rainbow 6* series the goal is for the player, a member of a tactical swat team, to "take out" bad guys and save hostages. In saving the hostages there is no difference in the rewards for saving the hostages and innocent bystanders over saving the hostages and killing innocent people as a means of saving the

hostages. You still “successfully complete the mission” even if you lose innocent civilians. A reward could be introduced for actions with different consequences.

Make different actions have different consequences- introduce the notion of sacrifice. One could make games where sometimes you can break a lower priority duty to preserve a higher one, lie or steal to save a life. In some situations one might allow duties and rights to override good consequences.

The modeling of how decisions are made, even for a murderous event, is ethically important. Who are the stakeholder’s and how are they affected. Instead of programming for a particular result the system should encourage the consideration of different stakeholder sets and adjust rewards for the number considered. This would encourage but not mandate particular ethical decisions.

A player could have an ethics rating that is adjusted in terms of respect for others. The reward might not be a better situation for the player -- self-focused decision -- but the reward might be a better situation for the rest of society. Choices which only involved self-focused decisions would not result in as good of a situation at the end of the game. Here the goal of the game is to improve the situation at the end of the game. The winners are those who make the best worlds.

We believe it is important for developers to weigh the way decisions are made against the potential harmful societal effect a single mode decision mechanism might have. This involves different degrees of ethical reflection for different types of games. A major design element should be a consideration of the ethical implications, potential or expected positive and negative effects on players and their environment. Education in ethics is needed

Conclusions

Our point in this paper is not to argue that we should ban all video games, not even all the violent ones. It is an empirical question what the after-effects of a given game are. If there really are no harmful effects in any way, then we have no objection. But, we have argued that convincing evidence of harmful after-effects in some cases does exist and that common sense tells us to be cautious about what the widespread indiscriminate use of video games might be. This matter certainly deserves more careful empirical study.

We have argued that video games have properties that make us wary of their misuse. Unlike traditional games they have increasingly good graphic capabilities, and more importantly they give players open ended possibilities for kinds of action, and in the case of some video games they overtly promote and reward immoral actions.

Finally, we have argued that at least some video games be produced that encourage ethical reflection and action. We need to adopt the same variety of critical approaches to decision making in them that we apply in the real world. Just consequentialism has been offered as an account of ethics that allows for realistic ethical decision-making and may provide insights how to embed ethical content into interesting and even fast action video games.

REFERENCES

Anderson, C. & Bushman B. (2001) "Effects of Violent Video Games on Aggressive Behavior, Aggressive Cognition, Aggressive Affect, Physiological Arousal, and Prosocial Behavior: A Meta-Analytic Review of the Scientific Literature." *Psychological Science* 12(5).

Brey, P. (1999) "The ethics of representation and action in virtual reality", *Ethics and Information Technology* 1(1), 5-14.

CNET News "Harvard researchers: Violent video games OK for kids" http://news.cnet.com/8301-17938_105-9942041-1.html accessed 30 April 2008.

Dodig-Crnkovic, G. (2006) On the Importance of Teaching Professional Ethics to Computer Science and Engineering Students, in *Computing and Philosophy* ed. L. Magnani, Associated International Academic Publishers, Pavia.

De Freitas, S. (2008) "Emerging trends in serious games and virtual worlds." *Emerging technologies for Learning* v. 3 (2008).

Milton Friedman (1991) "The Social Responsibility of Business Is to Increase Its Profits" in Deborah G. Johnson (Ed.) *Ethical Issues in Engineering* NJ, Prentice Hall, 84-92

Gammingsteve 2008, <http://gamer.blorge.com/2008/04/25/united-states-teaches-spies-how-to-think-using-video-games/> accessed June 2008

Gert, B. (1998) *Morality: Its Nature and Justification*. Oxford, Oxford University Press, 1998

Gert, B. (1999), Common morality and computing, *Ethics and Information Technology*, v.1 n.1, p.53-60

Gert, B. (2007), *Common Morality*, Oxford, Oxford University Press, 2007

Gotterbarn, D. (2008) "Video Game Ethics: Mayhem, Death, and the training of the Next Generation," *Proceedings of Ethicomp*

Johnson, D. (2001) *Computer Ethics*, 3rd edition, Prentice Hall, NJ

Moor, J. (1985) "What is Computer Ethics", *Metaphilosophy* 16(4), 266-275

Moor, J. (1998) "Reason, relativity, and responsibility in computer ethics", ACM SIGCAS Computers and Society, v.28 n.1, p.14-21, March 1998

Moor, J. (1999) "Just consequentialism and computing," Ethics and Information Technology 1: 65-69.

Psychology and Crime New, <http://crimepsychblog.com/?p=1453> accessed 30 April 2009.

Rawls, J. (1971) A Theory of Justice, Beeknap, Cambridge: Massachusetts

Rachels, J. (1999) The Elements of Moral Society The Elements of Moral Philosophy (3rd Ed.) Boston, McGraw Hill), 70-95

Slim, H. (2008) Killing Civilians: Method, Madness, and Morality in War, Columbia University Press, New York

Smith, H. (2009) "The Future of Game Design: Moving Beyond Deus Ex and Other Dated Paradigms" International Game Developers Association, http://www.igda.org/articles/hsmith_future.php accessed 30 April 2009

Surgeon General's Office, (2006) Mental Health Advisory Team IV Operation Iraqi Freedom 0507 Final report, November 17,2006.

Tavinor, G. (2007), Towards an Ethics of Video Gaming, Future Play, November 15-17, 2007, Toronto, Canada.

Waddington, D. (2007) "Locating the wrongness in ultra-violent video games" Ethics and Information Technology 9:121-128

Developing artificial agents worthy of trust: “Would you buy a used car from this artificial agent?”

F. S. Grodzinsky*

Sacred Heart University
Fairfield, CT USA

K. W. Miller**

University of Illinois at Springfield
Springfield, IL USA

M. J. Wolf***

Bemidji State University
Bemidji, MN USA

Abstract

There is a growing literature on the concept of “e-trust,” and on the feasibility and advisability of “trusting” artificial agents. In this paper we review important recent contributions to this literature, and then re-examine these matters from the perspective of a software developer. Too often, the primary focus of research in this area has been on the artificial agents and the humans they may encounter after they are deployed. We contend that the humans who design, implement and deploy the artificial agents are crucial to any discussion of e-trust.

Keywords: artificial agents, trust, e-trust

* *Frances Grodzinsky* is a professor of Computer Science and Information Technology at Sacred Heart University where she is co-chair of the Hersher Institute of Ethics. She is also a Visiting Scholar, Research Center on Computer Ethics and Social Responsibility, Southern Connecticut State University, New Haven, CT and serves on the board of INSEIT (the International Society for Ethics and Information Technology).

** *Keith W. Miller* is a professor of Computer Science at the University of Illinois at Springfield. His research areas are computer ethics and software testing. He is the editor-in-chief of *IEEE Technology and Society*, and serves on the board of INSEIT.

*** *Marty J. Wolf* is a professor of Computer Science at Bemidji State University, where he also serves as Coordinator for the Computer Science Program. In addition to his computer ethics research, he has also published papers in graph theory.

Introduction

In her 1978 book *Lying*, Bok writes, “Whatever matters to human beings, trust is the atmosphere in which it thrives” (1978). Annette Baier (1986) opens her article on “Trust and Antitrust” with this same quote. Both the book and the article are about trust among humans. A central issue we explore in this paper is whether this same atmosphere is likely or even possible when discussing interactions between humans and artificial agents (AAs), and when discussing interactions among artificial agents.

Trust is commonly defined as “a: assured reliance on the character, ability, strength, or truth of someone or something b: one in which confidence is placed 2 a: dependence on something future or contingent ...” (Webster, 2008). While this relatively straightforward definition might apply easily to human to human interactions ($H \rightarrow H$), and might be the starting point for a discussion on the concept of trust among philosophers and sociologists, it raises a series of questions for the software developer who is deciding if it is even possible to model trust into an artificial agent. What parameters should be in place when the interaction is human to artificial agent ($H \rightarrow AA$), artificial agent to human ($AA \rightarrow H$), or artificial agent to artificial agent ($AA \rightarrow AA$)? Is the word “trust” appropriate when discussing interactions that include, sometimes exclusively, AAs? If so, what should AA developers do to create trust in these environments? All three perspectives, $H \rightarrow AA$, $AA \rightarrow H$, and $AA \rightarrow AA$, present different challenges to the developer of an AA, not the least of which is exactly what he/she is trying to model.

Recent analyses of trust

In his paper entitled “What Model of Trust for Networked Cooperation? Online Social Trust in the Production of Common Goods (Knowledge Sharing)” Massimo Durante (2008) explores trust in terms of a socio-cognitive model of limited rationality for $H \rightarrow H$ interaction. This model “is aimed at coping with the uncertainty of what remains beyond control. The idea of trust offers us some insightful elements to reduce uncertainty within cooperative relations” (Durante, 2008). Durante also explores the role of delegation in trust relationships and how we move from “control trust” in technology where mechanisms define the trustworthiness of a system to “perceived trust” where trust is based on the confidence that the trustor has in the trustee (2008). While Durante is primarily investigating online cooperation in terms commons-based peer production among humans, there is a strong and growing possibility that a user might be dealing with AAs and not other humans. Should it change our notion of trust if we don’t know if the entity we are dealing with is human? Do we consciously delegate responsibility to an AA because we trust it, or is this an unconscious decision that we make

when we delegate responsibility to our computer? In addition, while his situations focus on cooperation, there are times when the trustee is simply facilitating the realization of a goal of the trustor.

To take a practical example, when we use Google to search the web, we rely on Google to return a list of sites relevant to our search terms, but most of the time we don't make a conscious decision to place our trust in Google. At one level of abstraction, the user and Google share the goal of an information exchange about sites relevant to the search terms; at another level of abstraction, Google has no information about the ultimate goal of the entity doing the search. Note that AAs do Google searches as well as humans. For example, Howe and Nissenbaum (2009) produced TrackMeNot software to do random Google searches to obscure information about the user's actual searches.

Humans have some trust issues that are conscious and well thought out; for example, trust is an explicit issue when we vote in elections. We have some trust decisions that are initially thought out, but then become habitual; for example, "Should I buy from this website?" These different approaches to trust - largely automatic, explicitly deliberate, and shifting from deliberate to automatic over time - are all relevant to how a software developer must approach designing an AA to interact with humans and with other AAs.

The history of humans trusting humans (and not) offers a starting point to our discussion of these questions. We humans trust each other (and not) to different degrees. By and large, that trust works, but it also fails on occasion, sometimes disastrously. Our first attempt at conceptualizing trust that involves AAs is to apply whatever trust framework a person has for dealing with people as a basis for trusting AAs. This strategy has the strength of being "species-independent;" that is, we trust entities to the degree that their actions warrant regardless of whether the entities are carbon-based humans or silicon-based AAs. Modeling human trust in AAs after human trust in humans would be one way to move beyond what Gunkel (2007) describes as the "anthropocentrism" of traditional moral theory. However, the elegant simplicity of this approach may not be appropriate; humans and AAs are not identical, and therefore our approach to trust perhaps should be different in order to better take into account those differences where possible. (Identifying when we are dealing with a human and when we are dealing with an AA is itself problematic, as we will discuss later.) No matter what approach is taken to trusting AAs (or not), humans need to be aware that some AAs will fail in terms of violating trust, in ways similar to those that some humans we choose to trust ultimately fail. Mariarosaria Taddeo (2008) states that the issue for ICT is the management of trust, developing parameters for trust to emerge, and then finding the methods of assessment.

Taddeo's paper entitled "Modeling trust in artificial agents, a first step toward the analysis of e-trust" presents an analysis of how to build an assessment model to evaluate trust in distributive systems. She bases this model on the assumption that you can design an AA as fully rational (see Floridi and Sanders, 2004), and, therefore that agent would be able to choose the best option on who to trust based on specific information and the agent's goal.

We have an immediate concern at this point, because there is no guarantee that the developer of an AA has been successful at creating an AA whose behavior would be fairly characterized as "rational." Furthermore, if the AA is capable of changing its internal program after deployment, then its "rationality" is even further from guaranteed (Grodzinsky et al., 2008). Despite these reservations, we still think Taddeo's next point is important: how an artificial agent could be programmed to behave in a manner similar to how humans behave when they report that they have learned to trust someone or something.

In Taddeo's analysis, an AA can measure another entity's trustworthiness according to the ratio of successful actions divided by the total number of actions necessary to achieve a similar goal. The entity whose trustworthiness is being measured could be another AA or a human; in fact, the AA making the measurement may not know what kind of entity is being evaluated. Anything or anyone whose actions led to a success rate above a designated threshold would be deemed trustworthy (Taddeo, 2008). Entities whose measured performance was below the threshold would be deemed risky or "untrustworthy."

Clearly, there is an important risk analysis inherent when a threshold is determined. In one scenario, the developers of an AA could set the threshold before an AA is deployed, and it would remain at this value after deployment. In a significantly different scenario, the developers could establish an initial value for the threshold, and then the AA might adjust the threshold after deployment based on events that occur during the AA's interactions with other entities. No matter when or how the threshold is set, the ultimate goal is to have the benefits outweigh the costs of committing to a trust relationship.

Taddeo points out that if the trustor has sufficient confidence in the trustee, then the trustor will let the trustee act without supervision, increasing the benefit of the trust relationship for the trustor. If the AA trustee has a fixed program, we expect that the trustor should have higher confidence than if the AA trustee can change its program after deployment. However, because currently there is little information publicly available about AAs that we (or other AAs) might encounter, potential trust partners do not normally know anything about the programming of an AA, or in fact whether a potential trust partner is human or AA. In such a situation, we necessarily rely on reputation and past performance. For ex-

ample, we trust Google to return useful search results because it has for us in the past, and if there were suddenly a large problem with Google searches, we would expect to hear about it in short order from many disgruntled users.

Trust involving artificial agents

In her paper “Defining trust and e-trust: from old theories to new problems,” Taddeo (2009) does an analysis of several different definitions of trust and e-trust that have been suggested in the past twenty years, and presents several problems that remain. We will not replay the arguments behind this analysis here; interested readers should see Taddeo’s paper, as well as ideas she criticizes from Luhmann (1979), Gambetta (1998), Nissenbaum (2001), Weckert (2005), and Tuomela and Hofmann (2003). Despite the remaining controversies and questions that Taddeo identifies, we require at least an outline of trust and e-trust to accomplish our goal of ethical advice to software developers involved in AA projects. To that end and following Taddeo’s analysis, we will assert the following principles about trust and e-trust. First, about trust:

1. Trust is a relation between a (the trustor) and b (the trustee). NOTE: a and b can be human or artificial. A relation (certainly in the mathematical sense, but also in the sociological sense) can involve both.
2. Trust is a decision by a to delegate to b some aspect of importance to a in achieving a goal. NOTE: We rely on the notion that an artificial entity a includes “decisions” (implemented by, for example, IF/THEN/ELSE statements), and we assume that a’s decisions are designed and implemented with the assumption that there is a high probability that b will behave as expected.
3. Trust involves risk; the less information the trustor a has about the trustee b, the higher the risk and the more trust is required. NOTE: this is true for both artificial and human entities. In AAs, we expect that risk and trust are quantified or at least categorized explicitly; in humans, we do not expect that this proportionality is measured with mathematical precision.
4. The trustor a has the expectation of gain by trusting the trustee b. NOTE: With respect to AAs, “expectation of gain” may refer to the expectation of the AA’s designer in moving toward a particular goal, or it may refer to an explicit expression in the source code that identifies this expected gain, or both.
5. The trustee b may or may not be aware that trustor a trusts b. NOTE: If b is human, circumstances may have prevented b from knowing that a trusts b. The same is true if b is an AA, but there is also some possibility that an AA trustee b may not even be capable of “knowing” in the traditional human sense.

6. Positive outcomes when a trusts b encourage a to continue trusting b.
 NOTE: If a is an AA, this cycle of trustàgood outcomeàmore trust could be explicit in the design and implementation of the AA, or else it could be implicit in data relationships, as in a neural net.

Second, about e-trust:

E-trust occurs in Cyberspace, where physical contact is not required between a and b, and where there may or may not be social norms.

“Trust needs touch” is not a requirement.

Referential trust (based on recommendations) is often important in e-trust.

In addition to these assertions, we will assume a definition of artificial agents that we used in an earlier paper (Grodzinsky et al., 2008): An “artificial agent” is a nonhuman entity that is autonomous, interacts with its environment and adapts itself as a function of its internal state and its interaction with the environment. There are numerous objections to this (or any other) definition of AAs. One important objection is that an entity might act as an agent without the ability to adapt itself. Although we think that is a possibility, we also think that the adaptable AA is the far more interesting case in ethical analysis, and it is the one we will focus on here. We note that the adaptation could be as simple as changing the value of a variable or as complex as “self-modifying code,” by which an AA can change its programming after deployment.

The extent to which the AA can adapt is a factor that we found crucial in our previous paper. If the adaptation can negate the initial design, the behavior of the AA is far less predictable than if the possible adaptations are strictly limited. As we will see, the predictability of an AA has important consequences in issues of trust.

Finally, as background to several of the issues we will consider, we assert that at some point in the future it will not always be easy or convenient during an interaction to discern if an entity is artificial. This is already true for some small subset of interactions mediated by computers. (For example, emails can be written by AAs that appear to be written by a particular human to a particular human.) We expect that some phone interactions in the future will be done by AAs in a way that will be hard to distinguish from phone interactions between humans. The use of humanoid robots that are indistinguishable from humans in personal interactions is much further away in time, but we do not foresee any decisive reasons why such future developments are impossible. For this reason, unless otherwise noted, our discussion about AAs interacting with humans and with each other include the possibility that the participants are physically proximate, are communicating via a phone, are interacting over the Internet in real time, or are interacting over the Internet asynchronously.

Initial analysis suggests that the four kinds of trust ($H \rightarrow H$, $H \rightarrow AA$, $AA \rightarrow H$, and $AA \rightarrow AA$) are sufficiently similar so that the principles noted previously are shared by all four. Except to occasionally point out contrasts with the other three possible trust interactions, we will not explore $H \rightarrow H$ trust interactions beyond noting that they share the six characteristics above with the three other types of trust interactions, whether the interaction takes place in person or online. Next, we will examine the remaining categories of trust interactions in more detail. Using these principles of trust we identify similarities as well as significant differences. The differences can stem from the fact that AAs can go through rapid and abrupt changes that may be due to software updates from the developer, or from some additional information, gathered by the AA that then modifies its transition table (Grodzinsky, et al., 2008); that is, we assume that some AAs are capable of something analogous to what is called “learning” in humans. Again, we prefer not to argue here how close such learning in AAs is to learning in humans; suffice it to say that we expect that AAs of the type we are considering can change future behaviors based on the effects of past actions.

$H \rightarrow AA$

The notation $H \rightarrow AA$ represents a human trusting an AA. As humans, our interest in $H \rightarrow AA$ is not merely in describing it. It is vitally important that we explore the issue of whether trusting AAs is a strategy that is likely to lead to positive results. The most simple-minded analysis of that issue leads to the vague answer “it depends on the AA in question.” Fair enough, but we need to go further. What characteristics of an AA will lead to it being worthy of human trust? As we look at this question from the perspective of the humans who design, implement and deploy an AA (a group we will call the AA’s developers), predictability is a central theme that we wish to emphasize.

Predictability is an important attribute from which to draw important distinctions between humans and AAs. AAs are distinct in the sense that we expect that they are capable of much faster changes than humans. Also, the discrete nature of binary encoded programs increases the likelihood of abrupt and dramatic changes; we expect slower, more gradual changes in processes that at least appear to follow laws described with continuous values and mathematics. (That is, in general we expect binary processes to appear more “jumpy” and analog processes to appear “smoother.”) Because software moves at speeds that are beyond the perception of humans, AAs can go through a dramatic self-modification process multiple times during a relatively slow interaction with a human. This sort of change can be disruptive to any existing trust relationship that relies on predictability and that grows out of past experience with that AA.

Distinguishing between types of implementations of AAs is important here. The simplest situation is when the AA is software that is run on a computer that the human in the trust relationship controls. When the hardware and the software of this computer are completely secure, and the software does not self-modify its code, then the human can be relatively confident that any trust that has been established is still valid.

A slightly more complicated case occurs when numerous people have developed a trust relationship with a single instance of an AA. Without a mechanism by which the humans are alerted that its programming has changed, there is the potential that the trust knowledge held by the human can be exploited by the AA. These changes can come about either through self-modification or through developer directed upgrades. Regardless, developers need to consider the impact of those changes on everyone who has developed a trust relationship with the AA. The fundamental question is whether the people involved deserve to know that the AA has changed.

This observation raises another important issue: trust may be tied to whether the human can identify the actual AA he/she is working with. Building trust over multiple transactions with an AA demands that the human be in a position to identify the AA. This raises an important ethical consideration for AA developers. At what point does an AA stop being the original AA and become a different one? At what point has the AA changed so significantly, that any reasonable human who had interacted with the AA in the past should no longer rely on the trust history that has been built up? The answers to these types of questions are further complicated by the possibility that some humans may have had multiple recent interactions with recent versions of the AA and other interactions may have been with much older versions. If a human is unable to identify the AA or identify that the AA has been modified, then that human is at risk if he/she trusts the AA.

Even in the simplest AA interactions, there is always an element of risk. The saying "no one is perfect" originally referred to humans, but it is no less applicable to AAs. Artificial agents are becoming increasingly sophisticated, and the software size and complexity necessary for that sophistication increases the likelihood of software faults (Clarke and Wing, 1996). The problems of validation and verification for such systems are particularly acute for adaptive systems such as neural nets (Schumann and Nelson, 2002). The potential for making an individual AA more useful to humans encourages increasing its complexity. The lure of making humans more comfortable with AAs also encourages this complexity, at least when "more human-like" is equated with increased comfort (DiSalvo et al., 2002). However, it is important to note the likely increase of risk (and the decrease of trustworthiness) that increased complexity can entail.

This aspect of the relationship of AA software reliability to human trust in AAs is part of a much larger issue of the impact of software quality on humans (Wolf and Grodzinsky, 2006). We cannot explore this issue in any depth here, but note that the many arguments for simplifying software artifacts in order to increase their reliability are particularly relevant to the issue of H_àAA trust. As AAs proliferate, and as they become increasingly responsible for important aspects of human lives, the reliability of those AAs is a direct responsibility of the AA developers. The seemingly autonomous behavior of sophisticated AAs may mask the role of the developers who launched the AA; but this does not absolve the AA developers from their responsibilities for the immediate and the future consequences of the AA's deployment.

There is a revealing example of H_àAA trust that started in the early days of the web: the textfile called "robots.txt." A website owner who wants web-crawling programs to ignore all or parts of the website pages, places a file named "robots.txt" in the top directory where the web files are located. Using a voluntary standard called the "Robot Exclusion Protocol," the robots.txt file indicates what web pages the owner would like excluded from any searches done by an automated web searching bot. According to a 2007 study (Sun et al., 2007), more than 38% of the websites they examined included a robots.txt file. Subsequent research reported 2.2 million robots.txt files (Sun et al., 2008). These numbers document an enormous demonstration of human trust that at least some web crawling bots (which fit our definition of an AA) will honor this voluntary protocol. Interestingly, some web page owners have sought to shame people who deploy web bots that violate the robots.txt protocol. (For example, see Kloth, 2007). The existence of millions of robots.txt files, and the attempt at retribution against AAs and their developers who do not honor the voluntary protocol, are a singular and contemporary example of a trust relationship between humans, AAs, and the developers of AAs.

Whenever data mining is used, ethical issues arise. These issues can become particularly complex if data mining is used to select and weight indicators (Fule and Roddick, 2004). But we assert that these issues are particularly troubling if AAs collect and analyze data mining results and then apply the result of that analysis to influence AA actions. Such actions can have significant effects on humans without possible interventions from other humans. It is one thing to program an AA to discern patterns based on fixed criteria in data available to the AA after deployment; it is quite another to allow the AA to adjust or augment those criteria based on its own analysis of data-mined information after deployment.

Imagine being denied a loan because the decision was made by an AA using data mining information collected during a deep recession. In the best case, the hu-

man might be able to appeal the decision; in the worst case, the AA might have the final say based on a neural net “analysis” that would be difficult if not impossible to explain to a human.¹

Transparency at best, and traceability at least, is a theme we raise in this section, and reprise in the next. If humans are to trust AAs, then AA developers should produce systems whose criteria and processes for making decisions are accessible to humans. If these systems’ decision making processes are obscure or hidden, humans are less likely to trust AAs over the long run, and we assert that humans should not trust such systems.

The formal nature of software makes transparency possible for AAs in a way that is not possible for humans. In the area of transparency, it may be possible (while acknowledging the potential for deception) to exhibit trustworthiness in AAs more readily than in humans.

AA→H

This notation represents an AA “trusting” a human. We place the term “trusting” in quotes because we do not want to defend the claim that a computer program “experiences trust” in a way that is identical to humans. Instead, we want to explore (without a protracted debate about the nature of an AA’s “experience”) the behavior of an AA that would appear to an outside observer to be based on a relationship of trust. In this respect we are using levels of abstraction in a way consistent with Floridi and Sanders (2004), but without conceding that an AA should be declared a moral agent.

In the previous section we explored humans making judgments about AAs and their trustworthiness. In this section, we explore the process and implications of AAs that act according to some measure of human trustworthiness. For example, an effective software agent used to buy goods offered on the web can not merely look for the lowest price. Although price is surely a factor, the reliability of the seller is also important. According to our definitions, such an AA is deciding which entity can be trusted sufficiently to risk a buy.

When an AA developer is programming an AA to perform a task that requires such a “judgment,” what indicators can the AA use to decide among possible trading partners? And among the possible indicators, which indicators should be selected by the AA developers? Issues of justice and fairness are clearly at stake here. In this section (as well as the next), we note that a software developer deal-

1. There are both theoretical and practical reasons why neural net decisions are unlikely to be easily explained to humans. For example, systems that can give a comprehensible explanation to a human of why a decision was reached are more far resource intensive than systems that are less expressive about their reasoning (Greiner et al., 2001).

ing with these AA→H trust behaviors will have to make explicit notions about trust that may be vague and amorphous in normal human trust relationships. Human might report emotions and intuitions as important elements in establishing trust; but software developers will be hard pressed to program such notions, at least with currently available AI techniques.

As AA developers create, deploy and gain experience with AAs that explicitly model AA→H trust, they will be able to collect and analyze data about which AA decisions and protocols for trust succeed and which fail according to some objective criteria about what constitutes a successful AA→H trust relationship. By analyzing this data, psychologists and philosophers may be able to make new hypotheses about the mechanisms of AA trust, which may offer insights into human trust.

AA→AA

When the need for trust decisions arises in an AA to AA interaction, what criteria should AA1 use to decide if AA2 is trustworthy? The details of the criteria will necessarily be application-specific, but we present some general principles to guide AA developers producing AAs that will need to be trustors or trustees in their interactions with other AAs.

As in the H→AA section, we contend that transparency is vital when determining the trustworthiness of decision-making procedures and sources of data. Transparency can be difficult when a commercial AA's decision-making details constitute a trade secret; even in such cases, the AA's details could be made known to designated third parties (regulators or consultants) dedicated to keeping AA interactions fair. In addition to this overall transparency, individual decisions should be traceable, so that disputed decisions in AA→AA interactions can be investigated ex-post facto for purposes of undoing any injustices, and for analyzing what went wrong in order to improve future interactions.

At least initially, AA developers necessarily make explicit the criteria the AA will use in making decisions. (If the AA can self-modify, those criteria can change.) This formalizing activity is a difficult one requiring delicate judgments and ethical sensitivity. It is also an opportunity to explore methods of making such decisions justly and efficiently. When AA developers take these ethical challenges seriously, their deliberations, the resulting AA programs, and data about the results of their programming as the AAs are deployed and used, are likely to be useful to philosophical and political debates about both human and AA decision-making policies.²

2. The learning that can take place studying AA trust decisions will be facilitated if the software is available widely. Thus source code available AA software (including Free Software) will be of particular interest.

Assessment methods and risk analysis need to be developed to evaluate parameters and the trustworthiness of AAs. While developers of unmodifiable AAs (see Grodzinsky et al., 2008) can set their parameters and test them, how can anything but initial parameters be relied upon in a modifiable AA? The testing of modifiable AAs is far more complicated than the testing of unmodifiable AAs. Unless the AA includes safeguards that are effectively shielded from future modifications, the possibility of modifications in response to unforeseeable future circumstances make testing of modifiable AAs at the very least impractical and probably impossible in any reasonable amount of time. Even “protected” safeguards may be vulnerable to future modifications (Grodzinsky et al., 2008).

Testing any software of even modest complexity is a major challenge in software engineering; testing software that can modify itself makes that challenge unmanageable. We contend that modifications after deployment that could affect an AA’s behavior should be severely limited by effective and well designed safeguards; it may be advisable to avoid self-modifying AAs altogether because of the inherent risks. Great caution is required in this area if humans (and AAs for that matter) are expected to rely on the trustworthiness of AAs.

Conclusions

Some issues surrounding trust and AAs are likely to be controversial for the foreseeable future. For example, will computers ever be capable of the kind of emotional attachments that humans associate with close trust relationships like close friends and family? We suspect that such questions will not be settled in our lifetimes. The existence of such questions should not deter us from wrestling with more pressing, more practical considerations such as those we have discussed above.

In the previous sections, we differentiated among $H \rightarrow H$, $H \rightarrow AA$, $AA \rightarrow H$, and $AA \rightarrow AA$. That seems appropriate, since humans and AAs are clearly not identical, nor do we expect they will ever be identical. However, as AAs become more sophisticated, we expect that it will be increasingly difficult for a human to determine if an entity with which they are interacting is human or artificial. Unless the speed of the interaction is obvious (artificial entities are capable of faster interactions), it may also become commonplace for AAs to interact in a similar way to humans and other AAs; this will be especially true of asynchronous interactions.

When this discrimination problem becomes commonplace (that is, when AAs routinely pass the Turing Test), there are at least two obvious strategies for both humans and AAs to deal with the situation in which the interaction partner is not known to be human or artificial. One strategy is that humans and AAs adopt

a common protocol for interactions, a protocol that does not discriminate between artificial and human partners. This strategy is elegantly simple, but it precludes taking advantage of some possibilities that are only possible if partners in an interaction know more about each other. For example, an AA \leftrightarrow AA interaction can proceed much faster if that is known to be the situation. A less technical advantage of knowing is that humans may be more comfortable interacting with a human, or at least knowing whether an interacting partner is of the same species.

The issue of trust is directly related to development of more human-like robots and automated voices because it is assumed that humans will trust robots more when the robots become more like humans (Bruemmer et al., 2004). However, our analysis above suggests that a desire for more human-like characteristics may result in complexity and unreliability that will decrease rather than enhance our trust in AAs.

One useful and distinctive characteristic of humans is our capacity to adapt. However, this adaptability can also lead to capriciousness and unpredictability. Some AA developers are attempting to make AAs more human-like by programming them to be more adaptable to their environment, by allowing them to self-modify their programs. We contend that the potential gains of this strategy are not sufficient to justify the enormous risks, especially when the adaptation process is poorly understood by the developer, and not easily recognized by humans who have trust relationships with the AA. We prefer that AAs be boringly predictable. We are far more concerned about the trustworthiness of AAs, and far less concerned that they mimic human's adaptability. In almost all situations (with the possible exception of computer gaming), we think that AA developers have a duty to the safety of the public that should restrict their use of self-modifying code to implement AAs, including limitations on the use of neural nets in AAs.

REFERENCES

- Baier, A.** (1986) Trust and antitrust. *Ethics*, 96, No. 2 (Jan 1986) 231-260.
- Bok, S.** (1978) *Lying*. New York: Pantheon Books.
- Bruemmer, D., Few, D., Goodrich, M., Norman, D., Sarkar, N., Scholtz, J., Smart, B., Swinson, M. L., and Yanco, H.** (2004) How to trust robots further than we can throw them, in CHI '04 Extended Abstracts on Human Factors in Computing Systems (Vienna, Austria, April 24 - 29, 2004). CHI '04. ACM, New York, NY, 1576-1577.

Clarke, E. M. and Wing, J. M. (1996) Formal methods: state of the art and future directions. *ACM Computing Surveys*, 28, 4 (Dec. 1996), 626-643.

DiSalvo, C. F., Gemperle, F., Forlizzi, J., and Kiesler, S. (2002) All robots are not created equal: the design and perception of humanoid robot heads, in *Proceedings of the 4th Conference on Designing interactive Systems: Processes, Practices, Methods, and Techniques* (London, England, June 25 - 28, 2002). DIS '02. ACM, New York, NY, 321-326.

Durante, M. (2008) What model of trust for networked cooperation? Online social trust in the production of common goods (knowledge sharing), in *Ethcomp 2008, Conference Proceedings*, University of Pavia, Mantua, Italy, 211-223.

Floridi, L. and Sanders, J.W. (2004) On the morality of artificial agents. *Minds and Machines*, 14, no.3, 349-379.

Fule, P. and Roddick, J. F. (2004) Detecting privacy and ethical sensitivity in data mining results, in *Proceedings of the 27th Australasian Conference on Computer Science - Volume 26* (Dunedin, New Zealand). Estivill-Castro, Ed. ACM International Conference Proceeding Series, vol. 56. Australian Computer Society, Darlinghurst, Australia, 159-166.

Gambetta, D. (1998) Can we trust trust?, in *Trust: Making and Breaking Cooperative Relations* (ed. D. Gambetta), 213-238.

Greiner, R., Darken, C., and Santoso, N. I. (2001) Efficient reasoning. *ACM Computing Surveys* 33, 1 (Mar. 2001), 1-30.

Grodzinsky, F. S., Miller, K. and Wolf, M. J. (2008) The ethics of designing artificial agents. *Ethics and Information Technology*, 10, 115-121.

Gunkel, D. J. (2007) Thinking otherwise: ethics, technology and other subjects. *Ethics and Information Technology*, 9, 165-177.

Howe, D. and Nissenbaum, H. (2009) TrackMeNot. <http://mrl.nyu.edu/~dhowe/TrackMeNot/>, accessed April 7, 2009.

Kloth, R. (2009) List of bad bots. <http://www.kloth.net/internet/badbots.php>, accessed March 31, 2009.

Luhmann, N. (1979) *Trust and Power*. Chichester: John Wiley

Merriam Webster Dictionary (2008) <http://mw1.m-w.com/dictionary/trust>. Accessed October 3, 2008.

Nissenbaum, H. (2001) Securing trust online: wisdom or oxymoron. *Boston University Law Review*, 81(3), 635-664.

Schumann, J. and Nelson, S. (2002) Toward V&V of neural network based controllers, in Proceedings of the First Workshop on Self-Healing Systems (Charleston, South Carolina, November 18 - 19, 2002). (eds. D. Garlan, J. Kramer, and A. Wolf), WOSS '02. ACM, New York, NY, 67-72.

Sun, Y., Councill, I. G., and Giles, C. L. (2008) BotSeer: An automated information system for analyzing web robots, in Proceedings of the 2008 Eighth international Conference on Web Engineering - Volume 00 (July 14 - 18, 2008). International Conference On Web Engineering. IEEE Computer Society, Washington, DC, 108-114.

Sun, Y., Zhuang, Z., and Giles, C. L., (2007) A large-scale study of robots.txt, in Proceedings of the 16th international Conference on World Wide Web (Banff, Alberta, Canada, May 08 - 12, 2007). WWW '07. ACM, New York, NY, 1123-1124.

Taddeo, M. (2008) Modeling trust in artificial agents, a first step toward the analysis of e-trust, in Sixth European Conference of Computing and Philosophy, University for Science and Technology, Montpellier, France, 16-18 June.

Taddeo, M. (2009) Defining trust and e-trust: from old theories to new problems. International Journal of Technology and Human Interaction 5, 2, April-June 2009.

Tuomela, M., and Hofmann, S. (2003) Simulating rational social normative trust, predictive trust, and predictive reliance between agents. Ethics and Information Technology, 5(3), 163-176.

Weckert, J. (2005) Trust in cyberspace, in The Impact of the Internet on Our Moral Lives (ed. R. J. Cavalier), Albany: University of New York Press, 95-120.

Wolf, M. J. and Grodzinsky, F. S. (2006). Good/fast/cheap: contexts, relationships and professional responsibility during software development, in Proceedings of the 2006 ACM Symposium on Applied Computing (Dijon, France, April 23 - 27, 2006). SAC '06. ACM, New York, NY, 261-266.

Can the “Contextual Integrity” Model of Privacy Be Applied to Personal Blogs in the Blogosphere?

Frances S. Grodzinsky* & Herman T. Tavani**
Sacred Heart University Rivier College

Abstract

In this paper, we analyze some controversial aspects of blogging and the blogosphere from the perspective of privacy. In particular, we focus on personal blogs and the case of the “Washingtonienne” blogger. We then apply Helen Nissenbaum’s theory of privacy as “contextual integrity” to that case. We next ask whether personal blogs that are not password protected can be considered “normatively private contexts” according to Nissenbaum’s principles of privacy. We argue that they cannot. Using Nissenbaum’s model, we conclude that privacy expectations for those who disclose personal information in such blogs are unrealistic. We also suggest that Nissenbaum’s theory can inform the contemporary debate about privacy and blogging in a wide variety of contexts, in addition to personal blogs, and we encourage researchers to apply Nissenbaum’s model in those contexts.

Keywords: blogs, contextual integrity, privacy

INTRODUCTION

In a previous work, we examined Helen Nissenbaum’s theory of “Privacy as Contextual Integrity” (Nissenbaum, 2004) in light of the RIAA (Recording Industry Association of America) v. Verizon and the MGM v. Grokster cases.¹ There, we used her privacy model to illustrate why P2P networks could be considered

* Frances Grodzinsky is a professor of Computer Science and Information Technology at Sacred Heart University where she is co-chair of the Hersher Institute of Ethics. She is also a Visiting Scholar, Research Center on Computer Ethics and Social Responsibility, Southern Connecticut State University, New Haven, CT and serves on the board of INSEIT (the International Society for Ethics and Information Technology).

** Herman Tavani is Professor of Philosophy at Rivier College and President of the International Society for Ethics and Information Technology (INSEIT). He is the author, editor, or co-editor of five books on ethical aspects of information technology. In 2008, he received the ACM SIGCAS Outstanding Service Award.

1. See “Online File Sharing: Resolving the Tensions between Privacy and Property Interests” (Grodzinsky and Tavani, 2008).

normatively private contexts.² This theory has clarified, for us, some of the conceptual muddles surrounding privacy and technology. In the present paper, we ask whether Nissenbaum's theory can also be applied to privacy concerns that arise in the "blogosphere," where expectations and assumptions about privacy are quite distinct from those affecting users in both the Verizon and MGM cases. The present paper is organized into two parts. In Part I, we briefly review the theory of privacy as contextual integrity. In Part II, we apply that theory to a specific case involving personal blogging: the "Washingtonienne" controversy that occurred in 2004. We consider three key questions: (1) Is a personal blog a normatively private context? (2) Were Jessica Cutler's assumptions about privacy protection of her blog correct? (3) Were Robert Steinbuch's expectations of privacy (in the Washingtonienne case) unrealistic?

In her essay "Privacy as Contextual Integrity," Nissenbaum (2004) expands upon the core concerns affecting (what she calls) "the problem of privacy in public," which she introduced in two earlier essays (Nissenbaum 1997, 1998). Her theory (as expressed in Nissenbaum 2004) is based on two principles:

- (i) The activities people engage in take place in a "plurality of realms" (i.e., spheres or contexts)
- (ii) Each realm has a distinct set of norms that govern its aspects.

Nissenbaum argues that norms affecting these two principles both shape and limit our roles, behavior, and expectations by governing the flow of personal information in a given context.³

There are two types of informational norms in Nissenbaum's privacy scheme: (a) norms of appropriateness, and (b) norms of distribution. The first of these determines whether a given type of personal information is either appropriate or inappropriate to divulge within a particular context. The second set of norms restricts the flow of information within and across contexts (Nissenbaum, 2004: 125). Nissenbaum argues that when either of these norms is "breached," a violation of privacy occurs. Her theory also illustrates why we must always attend to the context in which information flows, not the nature of the information itself, in

2. This expression was introduced by James Moor (1997), but applies to Nissenbaum's privacy theory as well.

3. The contextual integrity model proceeds on the assumption that there are "no areas of life are not governed by norms of information flow" (Nissenbaum 2004, 137). Our analysis of this model draws from an examination of the privacy-as-contextual-integrity theory in Tavani (2008a, 2008b).

determining whether normative protection is needed.⁴ Before seeing how Nissenbaum's theory can be applied to privacy concerns affecting blogging and the blogosphere, we briefly examine a controversial case of blogging that raises some key privacy concerns.

The washingtonienne case

In 2004 Jessica Cutler (aka "The Washingtonienne"), who was employed as an assistant to U.S. Senator Michael DeWine (R-Ohio), authored an online diary posted on blogger.com. This diary described a number of personal details about Cutler, including her annual salary as a Washington, DC Congressional Staff Assistant or "DC staffer" (\$25,000). Her diary also disclosed that most of her living expenses were "thankfully subsidized by a few generous older gentlemen." Additionally, Cutler's diary described her sexual relationships with these men, one of whom was married and an official in the Bush Administration. Cutler had assumed that her blog was being read only by a few close friends. However, in May 2004, an editor at Wonkette: DC Gossip, a popular blog in the DC area, discovered "The Washingtonienne" diary. When Cutler learned that her blog had been discovered, she deleted its contents from blogger.com (on the same day, ironically, that it appeared in Wonkette). When Cutler was eventually "outed" in Wonkette as "The Washingtonienne," she was fired as a staffer for "misuse of congressional computer resources." Shortly after her termination, Cutler received a book contract to publish details of her encounters and experiences as a DC staffer. She was also sued by one of the men implicated in her original blog (Leiby, 2004).

A source of considerable controversy and discussion, the Cutler case has received national attention in the media since 2004. Before analyzing the Washingtonienne case vis-à-vis Nissenbaum's privacy theory, however, we believe that it is important first to ask: What, exactly, is a blog?

Blogs

According to the (online) Merriam Webster Dictionary, a blog (or Web log) is "a Web site that contains an online personal journal with reflections, comments, and often hyperlinks provided by the writer; also: the contents of such a site" (<http://www.merriam-webster.com/dictionary/blog>). Note that in this paper, we do not discuss the history of blogs; interested readers are instead referred to Anton Zuiker's Web site (www.unc.edu/~zuiker/blogging 101/) for historical information

4. In the Verizon case, we saw that rather than focusing on the nature of the information included in a P2P situation – i.e., asking whether or not it should be viewed as private – we can ask whether P2P situations or contexts (in general) deserve protection as "normatively private situations".

on this topic. We do, however, believe that it is important to note that there are several different types of blogs. For example, there are political blogs, personal blogs, topical blogs, health blogs, literary blogs, travel blogs, and so forth. As a type of online personal dairy, Jessica Cutler's blog would seem to fall under the category "personal blog." So, in this paper, we limit our analysis of blogging vis-à-vis the "privacy as contextual integrity" model to personal blogs.

Personal Blogs

Blogging has become popular because it is an easy way to reach many people. Subscribers to sites such as Blogger.com can post their thoughts and opinions on whatever topic they choose. The personal blog has virtually replaced distribution list emailing to family and friends. A friend whose son is teaching in China for the year reflects on his experiences, shares his photos and gives his family accounts of his experiences. Another, whose husband is overseas in the military posts pictures of the children and recounts the daily happenings at home. The personal blog is an easy way to keep in touch. Others use personal blogs as a forum for opinions. In January, four students in the Philippines received a 10 day suspension from high school for insulting the principal on a personal blog (Sans appeal, 2009). The administrators at the school claimed such a posting impacted the school's reputation. Should personal bloggers be held to the same standards as online journalists? The debate centers on the motivation of the blogger. Personal bloggers argue that they are not journalists and their blogs are sites of personal reflection. Others say that irrespective of motivation, "responsible bloggers should recognize that they are publishing words publicly, and therefore have certain ethical obligations to their readers, the people they write about, and society in general" (A Bloggers' Code of Ethics, 2003). On the Web site CyberJournalist.net there is a Bloggers' Code of Ethics that is modeled on the Society of Professional Journalists Code of Ethics. The site explains that "these are just guidelines – in the end it is up to individual bloggers to choose their own best practices" (A Bloggers' Code of Ethics, 2003).

It is interesting that those who advocate for this code also appeal to the integrity of the bloggers in the "practice of ethical publishing". Yet, many who write personal blogs do not consider themselves "authors" in the field of publishing. Engaging in personal blogging is instead a form of social networking for the blogger, and the blog becomes a venue to reveal details of the blogger's personal life. For this generation, it has become a part of his/her social identity. In an article written for the New York Times "When Information Becomes T.M.I.", Warren St. John writes, "Through MySpace, personal blogs, YouTube and the like, this generation has seemed to view the notion of personal privacy as a quaint anachronism. Details that those of less enlightened generations might have viewed as embar-

assing — who you slept with last night, how many drinks you had before getting sick in your friend's car, the petty reason you had dropped a friend or been fired from a job — are instead signature elements of one's personal brand. To reveal, it has seemed is to be" (St. John, 2006).

There seems to be a misconception among some personal bloggers, Jessica Cutler included, that anonymity somehow protects the blogger and that only "friends" will be interested or know the identity of the writer of personal information. In their minds it is analogous to writing a personal diary and only allowing certain friends access to the information. Given the nature of the Internet, however, this is often not the case, and personal bloggers can lose control over their information as did Jessica Cutler. To retain some semblance of control, some bloggers choose to password protect their site; others do not list their blog, making it more difficult for search engines to find it. The most open blog is completely unprotected. Why would a blogger sharing personal diary-like entries unprotect his/her blog? Often it is simple naiveté about the media or indifference. Nardi et al report that "In theory, about 900 million people (if current estimates of Internet connectivity [14] are correct) could read any blog that is not password protected. How did bloggers in our study feel about this? Responses varied. A common response was indifference" (Nardi, 2004). In an interview with the Washington Post, Cutler demonstrates this behavior, "But I thought that was, like, too much trouble for my friends to have to type in a password, I thought there are so many people with their own blogs, mine is not even going to come up on the radar" (Witt, 2004). Cutler pointed out that she wrote her blog for three friends.

Nardi et al indicate that for some, blogging is a social activity involving others. "We learned that blogs create the audience, but the audience also creates the blog. This linkage happened in a number of ways: friends urging friends to blog, readers letting bloggers know they were waiting for posts, bloggers crafting posts with their audience in mind, and bloggers continuing discussions with readers in other media outside the blog" (Nardi, et al, 2004). We can observe this in Jessica Cutler's behavior. The Washington Post reported, "By the first week of May, she was having flings with so many guys that reporting them all to her girlfriends was starting to feel like way too much work. 'So I sent a mass e-mail out: 'You guys, should I have my own blog or what?' I was kidding, 'Jessica says. "But they were all, like, 'Yes, if anyone should have a blog it's you, because you have the most interesting life'" (Witt, 2004). Cutler's intention to protect the privacy of her lovers is supported by the claim that in writing the blog, no one would be named. Instead, initials would be used to distinguish her sexual partners (Witt, 2004).

Is there any other way of protecting privacy in personal blogs? Many blog sites that allow individuals to post personal blogs (in the form of online diaries) have

privacy policies. For example, WordPress and BlogHer have explicit privacy statements available on their Web sites (<http://wordpress.org/about/privacy> and www.blogger.com/privacy-policy, respectively). However, these policies primarily state how personal information of the blogger is or is not collected via mechanisms such as cookies. These policies do not guarantee that any personal information entered into a blog is prevented from being seen by the public. It is up to the blogger to set up that constraint by password protecting his/her site, which Cutler decided not to do.

Privacy as contextual integrity

We now focus our application of the contextual-integrity model of privacy on personal blogs. In particular, we examine the application of this model in the Washingtonienne case vis-à-vis certain norms of behavior and expectations of privacy for those interacting in that blog.

Norms of Appropriateness in a personal blog

With respect to norms of appropriateness in the context of a personal blog, we first need to differentiate between password-protected and non-password-protected personal blogs. In the case of non-password-protected blogs that function as online diaries, we ask whether the creator of the “diary” has a responsibility to the people with whom she/he interacts in real life? Consider that in a traditional (i.e., physical) diary, the diarist shares intimate details not meant to be public; that is the nature of “the diary.” Personal information is under the control of the diarist who may or may not allow the diary to be read by others. Often the diary is key locked with the diarist having the only key. What happens, however, when that diary moves to an online forum or venue, and when its content can be more easily read by others? Also, what are the expectations of privacy of those people who interacted with the diarist? Do they expect to read about their private interactions in her blog, especially when personal details are more open and thus more accessible to others? Assuming that the parties involved have not consented to such information about them being included on the blog, it would seem that norms of appropriateness have been violated.

Norms of Distribution in a personal blog

Did Cutler’s behavior also violate the norms of distribution in sharing intimate details of her sexual relationships with the blogosphere? Again we can ask the question about consent from the parties involved. Were they asked, and did they consent to their personal information being shared? If not, would they have any reasonable expectation with regard to privacy? Robert Steinbuch, one of the men implicated in her blog, has since sued Cutler, stating that her blog’s content was

sufficient to reveal his identity and thus violate his privacy. Not only did he not consent to having information about him distributed, but Steinbuch apparently had no idea that he, via an alias (initials), was even in Cutler's online diary. But what kind of recourse does Steinbuch have given the public context of online personal blogs? There is no clear precedent here from a legal point of view. "His case hinges on a century-old privacy tort claim known as 'public disclosure of private facts.' In theory, the tort provides a remedy when one publicizes private, embarrassing, non-newsworthy facts about a person in a manner that reasonable people would find highly offensive" (McClurg, 2005). The courts, however, have tended to defer to the "free speech" argument against privacy if the information published online is true (McClurg, 2005). However, the theory of privacy as contextual integrity can inform this discussion.

It would seem that Steinbuch's privacy was violated both from the contextual norms of appropriateness and distribution. It is not clear, however, whether those same norms would have been violated in the *Washingtonienne* case had the blog had been password-protected and limited to the three people that Cutler mentions. It would then seem to be no different than instances where three friends share some intimate details about office affairs with one another – the only difference being that there would be a log of the discussion as opposed to the whisperings at the bar. But the fact that Cutler took no precautions to protect the diary from the public makes the information available to anyone and thus violates both contextual norms. The same would seem to hold for all non-password-protected blogs where all the parties mentioned have not consented to having their names included or consented to having intimate details about them chronicled.

Conclusion

In this paper, we have examined some privacy issues affecting blogging and the blogosphere. We focused in particular on personal blogs, as opposed to blogs in general (i.e., the wide range of alternative types of blogs currently used). And we further limited our discussion to personal blogs that function as online diaries, such as in the *Washingtonienne* case, and which were not password-protected. We then applied Nissenbaum's theory of privacy as contextual integrity to the *Washingtonienne* case; we saw how both Nissenbaum's norms of appropriateness and norms of distribution can be applied. On the basis of our analysis, we concluded that (1) users who participate in such blogs have no reasonable expectation of personal privacy, and (2) the privacy of unconsenting parties whose names appear in those blogs can be violated. Our analysis has not, however, examined privacy concerns affecting the wider blogosphere, including more recent controversies involving twitter and "reblogging." While an examination of privacy issues in these scenarios is beyond the scope of this paper, we believe they

would also make excellent case studies for the application of Nissenbaum's privacy theory in future research projects.

REFERENCES

A Bloggers' Code of Ethics (2003) <http://www.cyberjournalist.net/news-/000215.php> accessed February 21, 2009.

Grodzinsky, F. S., and Tavani, H. T. (2005) "P2P Networks and the Verizon v. RIAA Case: Implications for Personal Privacy and Intellectual Property," *Ethics and Information Technology*, Vol. 7, No. 4, pp. 243-250.

Grodzinsky, F. S. and Tavani, H. T. (2008) "On-line File Sharing: Resolving the Tensions between Privacy and Property", *Ethcomp 2008 Conference Proceedings*, University of Pavia, Mantua, Italy, September, 2008, pp. 373-383. Reprinted in *Computers and Society*, Vol. 38, No. 4, 2008, pp. 28-39.

Leiby, R. (2004) "The Hill's Sex Diarist Reveals All (Well Some)," *The Washington Post*, May 23, p. D03. Also available at <http://www.washingtonpost.com/wp-dyn/articles/A48909-2004May22.html>. Accessed July 27, 2008.

McClurg, A. (2005) "Online Lessons on Unprotected Sex", *The Washington Post*, Available at <http://www.washingtonpost.com/wp-dyn/content/article/2005/08/14/AR2005081401034.html>. Accessed November 20, 2008.

Merriam Webster Dictionary. <http://www.merriam-webster.com/dictionary/blog>. Accessed Nov. 2, 2008.

Moor, J. H. (1997) "Towards a Theory of Privacy for the Information Age," *Computers and Society*, Vol. 27, No. 3, pp. 27-32.

Nardi, S., Gumbrecht, (2004) "Blogging as Social Activity, or, Would You Let 900

Million People Read Your Diary?" *CSCW* (November, 2004), Vol. 6, No. 3, pp. 222-231, New York: ACM Press.

Nardi S. et al (2004) [14] endnote; NITLE Blog Census. <http://www.blogcensus.net/?page=Home>

Nissenbaum, H. (1997) "Toward an Approach to Privacy in Public: Challenges of Information Technology," *Ethics and Behavior*, Vol. 7, No. 3, pp. 207-219.

Nissenbaum, H. (1998) "Protecting Privacy in an Information Age," *Law and Philosophy*, 559-596.

Nissenbaum, H. (2004) "Privacy as Contextual Integrity," *Washington Law Review*, Vol. 79, No. 1, pp. 119-157.

“Sans appeal, QC high school students’ suspension starts Monday” (2009). <http://www9.gmanews.tv/story/144297/Sans-appeal-QC-high-school-students-suspension-starts-Monday>, Jan 15, 2009. Accessed, February 21, 2009.

St. John, W. (2006) “When Information is T.M.I.” New York Times, September 10, 2006, http://www.nytimes.com/2006/09/10/fashion/10FACE.html?_r=1&scp=1&sq=When Information Becomes T.M.I. 20sept 2006 Accessed February 21, 2009.

Tavani, H. T. (2007) “Philosophical Theories of Privacy: Implications for an Adequate Online Privacy Policy,” *Metaphilosophy*, Vol. 38, No. 1, pp. 1-22.

Tavani, H. T. (2008a) “Informational Privacy: Concepts, Theories, and Controversies.” In K. E. Himma and H. T. Tavani, eds. *The Handbook of Information and Computer Ethics*. Hoboken, NJ: John Wiley and Sons, pp. 131-164.

Tavani, H. T. (2008b) “Floridi’s Ontological Theory of Informational Privacy: Some Implications and Challenges,” *Ethics and Information Technology*, Vol. 10, Nos. 2-3, pp. 155-166.

Witt, A. (2004) “Blog Interrupted”, *The Washington Post*, <http://www.washingtonpost.com/wp-dyn/articles/A48909-2004May22.html> Feb 7, 2009.

Zuriak, A. “Blogging 101.” Available at www.unc.edu/~zuiker/blogging 101. Accessed 11/16/2008.

ICT and financial services: learning lessons from the environmentalists

Mike Healy*

Westminster Business School
University of Westminster

N. Ben Fairweather**

Senior Research Fellow
Centre for Computing and Social Responsibility
De Montfort University

Abstract

Information communications technology has been instrumental in enabling the growth of the globalised financial system particularly with the deregulation of the sector in the 1980's. Yet the relationship between information communications technology (ICT), financial services and ethics has been under-researched. The transformation of the credit crisis into a global recession in many countries and an economic depression in others raises questions about the ethical and societal responsibilities of financial services computer professionals. This paper seeks to open an investigation into these issues. It starts with a review of the relationship between global finance and ICT and goes on to refer to previous work on ethical responsibilities of computer professionals. It then looks at the activities of ICT specialists involved with the green agenda consider how they approach their work to provide a comparison with the financial services sector. The paper concludes by arguing that financial services computer professionals can learn much from their colleagues in the green movement and highlights a number of possible research avenues to be explored.

* *Mike Healy* is a senior lecturer at the Westminster Business School where he teaches the ethical and societal implications of ICT, intranet design and management, and corporate identity. He has a number of publications concerning ICT and business ethics. He was recently the UK project manager of a pan-European project on age diversity within the ICT industry. He currently is leading the UK team associated with a European project focused on the use of ICT by teachers in their teaching.

** *Ben Fairweather* is Senior Research Fellow in the Centre for Computing and Social Responsibility at De Montfort University, where he spends much of his time supervising research students. He has been at De Montfort University since completing his own PhD on Freedom and Green Moral and Political Theory in 1996. He is co-editor of the *Journal of Information, Communication and Ethics in Society*.

Keywords: Credit crisis, financial ethics, ICT, banking, green computing

Recent developments in green and sustainable computing underline the growing interest by ICT (Information Communications Technology) developers and users in addressing the ethical and social issues associated with ICT (eg Wang 2007). This paper seeks to explore the extent to which debates about ethical and green computing could inform discussions about the use and abuse of infrastructural technologies that underpin the globalised financial system. The paper opens by looking at the growing concern of business ethicists about decisions and consequences of activities in the financial sector. It then goes on to describe how ICT operates within the financial sector and shows that ICT has been a critical element in the workings of international financial services. In doing so a range of problematic areas are identified such as the concentration and centralisation of decision making; the narrow parameters within which decision making is undertaken, and the pronounced tendency for national financial systems to automatically operate in tandem.

The paper then goes on to draw upon the experience of green or sustainable computing to argue that practitioners in the ICT industry have to be more vigorous in applying ethical standards when developing software models, software and hardware for the financial sector. By drawing upon the rich and growing discussions concerning green and sustainable computing, as well as corporate social responsibility, the paper argues that financial services computer professionals (FSCPs), be they software or hardware professionals, should be concerned with the ethical aspects of their work. The paper concludes by identifying future areas of research.

Dobson recognizes the difficulties linked to bringing ethics into the arena of financial activity and believes that financial professionals would question whether there is a role at all for ethics in finance given that the overarching theoretical construct motivating behavior in this sphere is personal wealth maximization (Dobson 1993, see Friedman, 1970). Boatright notes there are a further number of problems associated with financial ethics in that the sector covers many areas of work and a range of professions and activities from stock broking through to corporate financial regulation (Boatright 1999). However, the pressure of external events, such as the corporate scandals at the start of the millennium, provided a framework within which there was a vigorously renewed debate and research focusing on the role of the company directors in corporate governance. As was argued at the time, the “growing interest and concern is not surprising, given the significant financial and social harm these scandals have caused society” (Schwartz, et al. 2005: 79). A key conclusion of their research, and that of others (Baum 2003), was to see the need for a “code of ethics and ethics training specifically for directors, based on their unique role” providing a style and tone of

leadership throughout an organisation (Schwartz, et al 2005: 96). More recently, editions of the *Journal of Business Ethics* are, as is to be expected given the remit of the journal, liberally peppered with research papers concentrating on the role of ethics in the financial sector in general.

However, while the discussion on financial ethics has tended to focus on the role of managers, management and organisations, the recent economic crisis has brought the relationship between finance and of ICT, and the ethical implications of this relationship, into sharp relief. The fallout from the international credit crunch continues to develop both in terms of depth and breadth. While it was initially considered that the ICT sector would not be adversely affected by the crisis, recent evidence indicates that the sector is not immune from the detrimental consequences of the crisis. Here a number of snapshots emphasise the point. The Australian government has dramatically curtailed its spending on ICT both as a response to the economic crisis and in attempt to shape the future of ICT procurement. Hodgekinson of Ovum (an Australian subsidiary of Data-monitor) reports that the Australian government is seeking to reduce its number of ICT contractors by 50% by 2011 (Hodgekinson, 2008). Interactive Data Corp (a market intelligence group) estimates that 2009 will witness a deceleration in global technology spending with world shipment of PC's increasing by only 3.8% resulting in a falling value of 5.3%. This will make 2009 the first year of reduced growth since 2001 (Associated Press, 2008). The UK GDP figures suggest that the service sector, including computers, will slump by 1.2% in 2009, the worst figures since 1979 (Elliot and Seager 2009). Forrester Research estimates that technology spending by U.S. business and government will drop 3.1 percent in 2009 (Hoffman 2009).

A catalogue of events such as the international credit freeze, volatile exchange rates, the banking failures, a collapse in share prices, the increase in personal debt, and, particularly in the United Kingdom the United States, and countries such as Spain (where property prices fell 4.6% in 2008 and further, greater falls are expected in 2009 and 2010 [Spanish Property Insight 2009]) the dramatic decline in property prices has pushed economies into recession. The threat of depression looms large. The crisis has also thrown into doubt the whole globalization project with national economies slipping into protectionism and governments in public disagreement over the best policies to follow. As a result, the lives of millions have been dramatically affected by the crisis as so called toxic debt erodes economic confidence. The very notion of "toxic" debt evokes a parallel with environmental issues.

ICT has provided the infrastructure upon and through which the globalised financial system has developed, particularly since the deregulation of the sector in the mid 1980's (Yeomans 2004). As Wolf (2007) argues:

“No less important has been the revolution in computing and communications. This has permitted the generation and pricing of a host of complex transactions, particularly derivatives. It has also permitted 24-hour trading of vast volumes of financial assets. New computer-based risk management models have been employed across the financial sector. Today’s financial sector is a particularly vigorous child of the computer revolution” Wolf 2007: 2).

Software systems, using pre-programmed algorithms drawing upon weak risk assessment models which are based on false assumptions, running on powerful hardware systems, link national financial sectors to create a vast interlocking financial network often operating on auto-pilot. The underlying perspective influencing this activity is the realization of short-term stockholder value. This network leads to national financial sectors responding in similar ways to various events and for events in a national market to create an almost immediate reaction worldwide (Hurlburt et al 2008). The software and hardware developed for the financial sector, while not creating the current economic crisis, has been a contributing factor and this paper seeks to raise the ethical issues for the ICT sector arising from the current international banking crisis. Even before the current crisis, as Kumar and Hillegersberg (2004) note, “the financial services industry (was) in the early phases of a major transformation. The combined forces of deregulation, globalization and ICT usage are key components of this transformation” (Kuman and Hillegersberg 2004: 30). Before the onset of the recent crisis, there were concerns about problems arising from the expanded use of ICT within the financial sector (Schoenmaker and Oosterloo 2005, Chandra, 2007). Although some research on issues such as trust in financial services was undertaken before the economic crisis, this was primarily focused on online banking with the mainstream financial sector significantly under-researched (Tyler and Stanley 2007). Before the recession, few questions were raised concerning the ethical issues related to the use of ICT within financial systems and the discussion here tended to revolve around notions of sustainability and corporate social responsibility (Scholtens 2006). However, recent conditions have created an environment within which criticism about the use of technology in this area is now beginning to be articulated.

Stark (2008), member of the Executive Board of the European Central Bank, argues that innovation of both financial processes and products have benefited from advances in ICT in that these technologies have helped in creation, valuation and exchange of complex financial products. It has also affected the core of financial inter-mediation. Jennings (2008) also discusses problems with computers operating within financial systems but makes a more stinging criticism than Stark. “Computers have enabled financial products to proliferate. Computers amplify information and rumour, a vital market engine, to thunderous proportions.

They turbo-charge greed, stupidity and lemming-like behaviour, never in short supply, and the cream prudence gets dispersed, pushed to the bottom" (Jennings, 2009).

The pivotal role of that financial services play, particularly the banking and venture capital sectors that provide the bulk of external finance in most countries means that financial services have an immediate and dramatic impact of our lives in a way that no other economic activity is able. The ethical responsibilities are increasingly the concern of researchers looking at professions concerned with finance, such as accounting, and in doing so some question the value of professional codes of ethics and/or conduct (Flanagan and Clarke 2007). The importance of ethics associated with financial decision-making within public sector bodies has also received significant attention (Weiss and Gilman 2005). Research such as this and the current outcry at the way in which United Kingdom Members of Parliament claim and are paid expenses are indicative of a growing public concern that the processes used to undertake and the products resulting from financial decision making need to be made more transparent and accountable.

The scale, scope and range of the current economic crisis should therefore provide the impetus for research examining the roles of ICT in the financial sector from an ethical perspective. However, seeking to examine this area requires confronting a conundrum. Are we dealing with the private sector ethics, public sector ethics and/or individual ethics (see Schultz, 2004)? This paper is premised on the basis that because of the massive adverse affects of the financial crisis on the public globally, we are forced to address this issue from the public ethics perspective and to relate the general to the specific, i.e. individual, arena. Of course the problem then arises about how to develop ethical perspectives for ICT professionals working in the financial sector. Apart from needing to consider how far one should question the range of assumptions that underpin the activities in the area, consideration needs to be given to the ethical responsibilities of financial institutions and how these relate to the ethical responsibilities of individuals working with the industry. Further, there is also a need to examine the mechanisms that would allow organisations and individuals to advance and champion ethical perspectives and what protection should be given to those who choose to expose (or in the current jargon, whistle blow on) unethical practices in the sector. Some of the additional issues that are worthy of investigation are: the role of professional bodies in this discussion; the function of company compliance officers; the value of non-governmental organisations, including trade unions, in the process of developing financial ICT ethics; who should be held accountable for the adverse consequences of unethical decisions making; and what should be the compensation for the outcomes of such decisions.

It is obviously not possible in a paper of this length to engage in a discussion in all of these aspects. However, at this point it is worth noting that research and debate on how far ethical standards should influence the work of computer professionals has a long and vigorous history (ACM 1992, Rogerson and Gotterbarn 1998, Gotterbarn 2001). This tradition could provide signposts on how to explore the relationship between ethics, ICT, the ICT professional, and the financial sector. However a possible problem here could be the temptation to engage in the computer ethics and information ethics debate and/or to drift into discussions of an interesting but abstract nature.

It could be argued that since the recent actions of banks and financial institutions regarding risk management have been likened to gambling and it is therefore conceivable that exploration of this subject could draw upon work looking at ethics and e-gambling (Bull and Mclean 2007). However, the problem with such an approach is investigations focused on ethics and e-gambling are relatively under-researched; lack any firm theoretical underpinning; and the impact of e-gambling is relatively limited. Attempts have been made to apply ethical values to global commerce with an emphasis on trust, privacy, sharing, communication, security, anonymity, and learning (Deakin and Zutshi 2009). While such an approach is valuable, the problem is that the discussion is focused on making the activity more efficient rather than undertaking a more profound exploration of the activity itself to reveal the deeply embedded frailties that have led to the crisis and the role of FSCPs in this process.

The approach favoured here has been prompted by the view that the murmurings of criticism of the role of ICT in the financial sector reprise the doubts expressed years ago in another field: the environment. The economic crisis has been likened to a global financial tsunami (Greenspan in Crawford and Young 2009) adversely affecting the lives of millions of people with communities facing destruction; mass migration; development programmes suspended; increased levels of poverty; and increased child mortality rates. The United Nations Development Programme has characterized the crisis as a “human development catastrophe....This crisis is really a matter of life or death for many people in the poorest countries and it may take years to get back to the same level of economic growth, school enrollment and mortality rates” (UNDP 2009). The parallels with environmental issues are strikingly evident.

Taking their cue from and being influenced by the wider ecological movement, computer professionals have increasingly been conscious of the relationship between environment and ICT. Green computing focuses on the more efficient use of computing resources and attempts to address “the so-called triple bottom line of people, planet, profit” (Find White Papers 2009). This approach seeks to move be-

yond the economic aspects of computer systems with the result that the impact of extended computer use on the environment has increasingly been seen as an area of important and fruitful research. Yi and Thomas (2007) provide an overview of the literature concerned with the environmental impact of ICT and e-business and they note that there has been an increase in the number of journals, conferences, papers, projects and research activities in this field. While much of this research is directed towards measuring the positive or negative benefits of ICT and little tackles the more specific ethical or philosophical aspects related to this theme the evidence presented by Yi and Thomas clearly demonstrates a growing concern by environmentalists, academics and ICT practitioners of the environmental aspects of ICT. Murugesan (2008) argues that green ICT indicates a significant move in the IT industry in that it is increasingly concerned with infrastructure issues and the wider environmental implications of ICT development. He also notes that the IT sector as a whole has a responsibility to help create a more sustainable environment. One key theme notable within this discussion is the relationship between ICT, environment and corporate social responsibility. The maturity of the debate linking ICT to the environment could provide an inspiration guiding a critical appraisal of ICT within the financial sector. The approach advocated here is not unique and attempts have been made to link ethical consumerism to ICT (Chatzidakis and Mitussis 2007).

Mankoff et al (2008) argue that computer scientists have a major role to play in reducing carbon emissions in four ways, the first two of which: “involve mitigating the direct negative impact of computers—their power consumption as well as the economic and social costs associated with the manufacturing, maintenance, and disposal of components. The other two relate to the indirect positive impact of computers- their ability to increase energy efficiency by changing systems and ways of being; to potentially reduce world emissions by as much as 15 percent by 2020,.... and to help provide answers to important scientific questions” (Mankoff et al 2008). This call for action is reflected by research on ethical and sustainable computing across a range of sectors such as industrial and chemical processing; the development of interaction design; health care; urban development and transport planning; e-governance; and pharmaceutical research (e.g. Fermeiglia et al, 2009, Blevis, 2007, Haux, 2006, Bhattacharya and Vashistha, 2008, McGowen and McGowen, 2006). The growth of activity in this area has lead to computer professionals to talk of a green computing movement (Kurp 2008) with its own conferences, workshops, publications and influencing US government procurement policy (PR Newswire 2007).

As is to be expected in a broad movement, there are differences of opinion about the objectives of green computing. Some (British Computer Society 2008) focus on making existing ICT processes and activity, such as data centres, more sensitive to environmental concerns. Others (Fuchs 2008) have developed a more

direct criticism of the movement as a whole and call for a significant re-alignment of priorities. Notwithstanding the breadth of areas covered by the green ICT movement and the differences of view about the what goal of green or sustainable ICT should be, the starting point for much of the work undertaken by ICT professionals in the field is a shared recognition that the outcomes of their endeavors should not, at the minimum, lead to a deterioration in the environment, nor undermine peoples' living conditions. Indeed the shared core driver for the green computer movement is the enhancement of the human condition both in the short and long terms.

There are then three starting points for a discussion about the ethical position of ICT professionals working in the financial sector. These are the adverse impact of computer driven international financial activity on the lives of millions of people; the general discussion surrounding ICT professional practices and the extensive practical application of computer ethics as expressed in green computing. Gotterbarn (2001) refers to two forms of positive responsibility for computer professionals built upon a technical and value foundation and he calls for the computer professional to "use their skills for the good of society and not merely act as agents for the client" (Gotterbarn 2001: 229). Those computer professionals working in the green computer movement apply this perspective as part of their professional activity, as a matter of course.

That the global financial system operates the way it does is a testimony to the high level of technical skill shown by computer professionals working in the sector. No doubt they would subscribe to the view that it is important to strive to adhere to some notion of positive technical responsibility. However, the outcomes of their endeavors demands that they should look beyond this rather narrow view of responsibility (cf Fairweather, 2004) to consider the much wider implications of their products and processes. There is a further impetus for such a move because of the difficulties in involving stake holders the nature of the activity emphasizes the responsibility of the ICT professional. The weaker the link between the creator of an event and those impacted on by the event, the harder it will be to develop real and effective mechanisms of accountability, control and transparency (see, for example, Mellema 2003). These are very much the conditions within which FSCP's work. Thus FSCP's require the ballast of an ethical grounding to evaluate the ethical and societal implications of their work and to provide the intellectual framework to support challenges to the development of systems and processes that could be likely to have significant adverse results, such as the assumptions underlying principles guiding the construction of computerized risk models.

One result of the current economic crisis has been, quite rightly, the demand for tighter regulation over financial institutions, markets and systems and a call for greater transparency and accountability in these areas. This can, hopefully, be

useful at the macro level. At the micro level however, an argument has to be conducted with those who are charged with the day-to-day running of these systems about the wider implications of their activity. The very real difficulty here, as is much appreciated, is that while it might be relatively straightforward to identify what is the right thing to do: doing the right thing can be extremely problematic. Drawing upon the work of computer professionals working on the green agenda can be of help here since in undertaking this work they are adopting a perspective described by Gotterbarn (2001: 229) as the “broader sense of responsibility” and embracing the view that they “use their skills for the good of society and not merely act as agents for a client.”

At this moment such an approach requires that at three least types of conversation occur between various groups: ethical researchers and FSCP’s to investigate the range, depth and inhibitors of ethical influence in decision making; FSCP’s and green computer professionals to explore the how the difficulties of doing the right thing can be overcome in a concrete, practical manner; and FSCP’s and computer professional bodies to consider how far professional codes of conduct can have a meaningful influence in these circumstances.

Building on Moor (1985: 266), Tavani has argued (2002: 53) that “not only will the further development and use of ICT continue to create new policy vacuums and new conceptual muddles involving our moral notions at the level of specific problems such as privacy and free speech, but I believe that future uses of this technology may also present challenges to some of our more general and fundamental moral categories as well. For example, it is very likely that we will need to revisit core moral notions such as autonomy, agency, and moral responsibility.” The current economic crisis has created one of these new policy vacuums and, as Floridi and Sanders (2002) highlight, computer ethics should be concerned with the consequences of ICT on society and since, as has been seen above, the technology is intimately interwoven into the fabric of financial transactions, by extension it is appropriate to examine the ethical implications for ICT professionals working in this field. It can be further argued that given the global consequences of the crisis, such a project requires a degree of urgency.

REFERENCES

ACM (1992) ACM Code of Ethics and Professional Conduct available online at: <http://www.acm.org/about/code-of-ethics> accessed 15 December 2008

Associated Press (2009) IDC sees decline in US PC market next year available online at: <http://www.google.com/hostednews/ap/article/ALeqM5gsCobUx-4dBXw19Lg4Z0Od-Nb7WqAD94RBNKG0> accessed 4 December 2009

Baum, S. (2003) 'A Board's Top Job: Watching the CEO', Business Week Online, available at: http://www.businessweek.com/bwdaily/dnflash/may2003/nf-2-003056_5899.htm. accessed December 2008

Bhattacharya, J. Vashistha S. (2008) Utility computing-based framework for e-governance ICEGOV '08: Proceedings of the 2nd International Conference on Theory and Practice of Electronic Governance ACM

Blevis, E. (2007) Sustainable interaction design: invention & disposal, renewal & reuse Proceedings of the SIGCHI conference on Human factors in computing systems San Jose, California, USA

Boatright J.R. (1999) Ethics in Finance, Blackwell, London

British Computer Society (2008) Cleaning up the carbon footprints ITNOW available online at: <http://itnow.oxfordjournals.org/cgi/reprint/50/3/12> accessed 12 November 2008

Bull, C. and McLean, R. (2007) Computer and Information Ethics: The Neglect of Moral Issues Associated in E-Commerce The 5th European Conference on Computing and Philosophy (ECAP) Twente, The Netherlands available online at: <http://e-space.openrepository.com/e-space/handle/2173/23554> accessed 20 February 2009

Chandra C.A.I. (2007) Auditing in Core-Banking Environment-Some Special Considerations Chartered Accountant available online at: http://www.icaai.org/resource_file/96841404-1411.pdf. accessed 10 January 2009

Chatzidakis, A. and Mitussis D. (2007) Computer ethics and consumer ethics: the impact of the internet on consumers' ethical decision-making process Journal of Consumer Behaviour Volume 6 Issue 5, Pages 305 - 320

Ciborra, C. (2006) Imbrications of Representations: Risk and Digital Technologies Journal of Management Studies 43 1339-1356

Crawford P., Young T. (2009) The Credit Crunch: The Roller Coaster Ride Continues Journal of Business & Economics Research Volume 7, Number 1

Creed, A., and Zutshi, A., (2009) Relational Ethics in Global Commerce Journal of Electronic Commerce in Organizations, Volume 7, Issue 1

Dobson, J. (1993) The Role of Ethics in Finance Financial Analysts Journal, Vol. 49, No. 6 57-61

Elliot, L. and Seager A. (2009) Alistair Darling backs economic recovery despite slump in GDP available online at: <http://www.guardian.co.uk/>

politics/2009/apr/25/alistair-darling-budget-gdp-output accessed 26 April 2009

Fairweather, N.B. (2004) No PAPA: Why Incomplete Codes of Ethics are Worse than None at All: in Bynum, TW and Rogerson, S eds *Computer Ethics and Professional Responsibility* (Malden, MA: Blackwell)

Fen, W. and Cameron, K.W (2007) The Green500 List: Encouraging Sustainable Supercomputing *Computer Vol 40 No2* 50-55

Fermeglia, M., Longo, G., Toma, L. (2008) Computer aided design for sustainable industrial processes: Specific tools and applications *AIChE Journal Volume 55 Issue 4, Pages 1065 – 1078*

Flanagan, J. and Clarke K. (2007). Beyond a Code of Professional Ethics: A Holistic Model of Ethical Decision-Making for Accountants *ABACUS, Vol. 43, No. 4*, 488-518

Floridi, L. and Sanders J.W. (2002) Mapping the foundationalist debate in computer ethics *Ethics and Information Technology 4*: 1-9

Friedman M. (1970) The Social Responsibility of Business is to Increase Profits. *New York Times Magazine* (Sept) 13.

Fuchs, C. (2008) The implications of new information and communication technologies for sustainability *Environment, Development and Sustainability Volume 10, Number 3*

Gotterbarn, D. (2001) Informatics and Professional Responsibility *Science and Engineering Ethics Volume 7 Issue 2 pp. 221-230*

Haux, R. (2006) Individualization, globalization and health – about sustainable information technologies and the aim of medical informatics *International Journal of Medical Informatics Volume 75, Issue 12, 795-808*

Hilty, L.M. et al (2006). The relevance of information and communication technologies for environmental sustainability - A prospective simulation study *Environmental Modelling & Software 21 pp. 1618-1629*

Hodgekinson, S. (2008) Australian government agrees to implement Gershon's ICT Review available at: <http://www.ovum.com/news/euronews.asp?id=7535> accessed 4 December 2008

Hoffman K. (2009) Unisys Reviews Debt Options Amid Expiring Credit Line (Update3) available online at: <http://www.bloomberg.com/apps/news?pid=20601109&sid=aIT7vYasAVwg&refer=home> accessed 29 April 2009

Hurlburt G.F., Miller K.W. and Jeffrey M. (2009) An Ethical Analysis of Automation, Risk, and the Financial Crises of 2008 IT Pro January/February available on line at: http://www.computer.org/portal/cms_docs_itpro/itpro/14-19.pdf accessed 31 March 2009

Jennings, K. (2009) The banker: once a financial anchor available online from: <http://business.smh.com.au/business/the-banker-once-a-financial-anchor-20081128-6mym.html?page=fullpage#contentSwap2> accessed 1 December 2009

Kobayashi, H. et al (2007) Green Behaviour Generation: A Digital Approach to Reduce Consumption Systems, Man and Cybernetics 2007 available online at: <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=4414155&isnumber=4413561> accessed 15 November 2008

Kumar, K. and Hillegersberg, J. (2004) New Architectures for Financial Services Communications of the ACM Vol 47 No 5 27-30

Kumar, K. and Hillegersberg, J. (2008) Bank in a Box: An ICT architecture for enabling agile transformation of financial services Managerial Finance Vol 34 No 6 413-422

Kurp, P. (2008) Green computing Communications of the ACM Volume 51, Issue 10

Logutenkova, E. and Baumgaertel, C. (2008) Credit Suisse to Eliminate 5,300 Jobs After Losses available online at: http://www.bloomberg.com/apps/news?pid=20601102&sid=afGPN._nqmiU&refer=uk accessed 4 December 2009

Mankoff, J., Kravets, R. & Blevis, E., (2008) Science Issues in Creating a Sustainable World Computer available online at: <http://research.stepgreen.org/site/uploads/Footprints/ieee08-preprint.pdf> accessed 10 January 2009

McGowan M.K. and McGowan R.J. (2006) The Ethics of Computing in Pharmaceutical Research in The Ethics of Computing in Pharmaceutical Research John Wiley & Sons

Mellema, G. (2003) Responsibility, Taint, and Ethical Distance in Business Ethics Journal of Business Ethics 47: 125-132

Moor, J. (1985) What Is Computer Ethics? Metaphilosophy 16(4), 266-275

Murugesan, S. (2008) Harnessing Green IT: Principles and Practices IT Professional Volume 10, Issue 1, 24 - 33

PR Newswire (2007) FOSE Rolls Out the Green Carpet for 'Green Computing' available on-line at: <http://www.prnewswire.com/cgi-bin/stories.pl?ACCT=109&STORY=/www/story/03-12-2007/0004544255&EDATE=> accessed on 21 February 2009

Rogerson, S. and Gotterbarn, D., (1998), The Ethics of Software Project Management, in COLLSTE, G., (editor), Ethics and information technology, New Academic Publishers, Delhi, India, pp137-154, reprinted in COLLSTE, G., (editor), Ethics in the age of information technology, Linkopings Universitet, Sweden, 2000, pp 278-300

Schoenmaker D., Oosterloo Sander (2005) Cross-Border Issues in European Financial Supervision available online at: [http://staff.feweb.vu.nl/dschoenmaker/Cross-border%20issues%20\(BoF%2021-2-2005\).pdf](http://staff.feweb.vu.nl/dschoenmaker/Cross-border%20issues%20(BoF%2021-2-2005).pdf). Accessed 24 February 2009

Scholtens, B., (2006) Finance as a Driver of Corporate Social Responsibility Journal of Business Ethics 68:19–33

Schultz, D., (2004) Professional Ethics in a Postmodern Society Public Integrity 6(4): 279–297.

Spanish Property Insight (2009) Spanish property prices will fall 25% by 2011 says Spanish bank BBVA available online at: <http://www.spanishpropertyinsight.com/buff/2008/12/spanish-property-prices-will-fall-25-percent-say-bbva/> accessed 27 April 2009

Stark, J. (2008) Growth and productivity of the financial sector: challenges for monetary policy available online at: <http://www.edubourse.com/finance/actualites.php?actu=47770> accessed 24 November 2008.

Schwartz, M.S., Dunfee, T.W. and Kline, M.J. (2005) Tone at the Top: An Ethics Code for Directors? Journal of Business Ethics 58: 79–100

Tavani H.T. (2002) The uniqueness debate in computer ethics: What exactly is at issue, and why does it matter? Ethics and Information Technology 4: 37–54.

Tyler, K. and Stanley E., (2007) The role of trust in financial services business relationships Journal of Services Marketing Vol 21 No 5 334-344

UNDP (2009) Poor countries threatened with higher death and school drop-out rates available online from: <http://content.undp.org/go/newsroom/2009/march/higher-death-and-school-drop-out-rates-threaten-poor-countries-facing-the-global-economic-crisis.en> accessed 5 April 2009

Van Veen-Groot, D.B. and Nijkamp P. 1999 Globalisation, transport and the environment: new perspectives for ecological economics Ecological Economics Volume 31(3): 331-346

Wang, D. (2007) Meeting Green Computing Challenges International Symposium on High Density packaging and Microsystem Integration, 2007. HDP '07. DOI 10.1109/HDP.2007.4283590

Wang F. and Tang S. (2004) Artificial societies for integrated and sustainable development of metropolitan systems *Intelligent Systems, IEEE* Volume: 19, Issue: 4 82- 87

Weiss L.C. and Gilman S. (2005) *The ethics challenge in public service: a problem-solving guide* Edition: 2, John Wiley and Sons

Wolf, M. (2007) The new capitalism *Financial Times* available on line at: http://www.cep.cl/Cenda/Cen_Documentos/Varios/0706_FT_Wolf.pdf accessed 11 November 2008

Yeomans, J. (2004) Centralisation – the route to controlling payment costs *Card Technology Today* Volume 16, Issue 5 10-11

Yi L. and Hywel T.R. (2007) A review of research on the environmental impact of e-business and ICT *Environment International* Vol 33

Moral Luck and Computer Ethics: Gaugin at the keyboard

David Sanford Horner*
University of Brighton
School of Computing, Mathematical
and Information Sciences

Abstract

I argue that the problem of 'moral luck' is an unjustly neglected topic within Computer Ethics. The concept of moral luck bears down most heavily on notions of professional responsibility and the identification and attribution of responsibility. It is the immunity from luck that conventionally marks out moral value from other kinds of values such as instrumental, technical, and use values. The paper describes the nature of moral luck and its erosion of the scope of responsibility and agency. I discuss four types of moral luck and their relevance to Computer Ethics: consequential luck (luck in the way things turn out); constitutive luck (luck in character - the kind of person you are in terms of inclinations, capacities, and temperament); circumstantial luck (luck in the problems faced and decisions that have to be made); and finally luck in antecedent circumstances. Moral luck also poses a challenge to the kinds of theoretical approaches often deployed in analyzing moral questions arising from the design and implementation of information and communication technologies: consequentialism; virtue theory; and duty ethics. I claim that by considering the concept of moral luck we arrive at some valuable insights into the scope and limits of ethical theory and professional practice.

Keywords: Computer Ethics; Moral Value; Moral Luck; Risk; Testing; Professional Codes; Consequentialism; Deontology; Virtue Ethics; Paradox.

* Currently *David Horner* is a principal lecturer and Head of the Research Student Division in the School of Computing, Mathematical and Information Sciences, University of Brighton, UK. He teaches courses in research ethics, information and media ethics, and film and philosophy. He is currently working on two book projects: *Understanding Media Ethics* (with Sage) and another on *Prediction and Computer Ethics*.

Introduction

The idea of responsibility is central to Computer Ethics. Simon Rogerson points out that computer professionals "...must be aware of their professional responsibilities, have available methods for resolving non-technical ethics questions and develop proactive skills to reduce the likelihood of ethical problems occurring" (Rogerson, 2002, p.159). Herman Tavani writes that: "...It would seem reasonable to hold the manufacturers of unreliable computer systems legally liable for harms caused by faulty design interfaces or by 'buggy' software. But we must also ask to what extent the computer professionals, especially the software engineers who design and develop 'safety-critical' and life critical applications, bear responsibility for the harmful consequences that result from unreliable computer systems?" (Tavani, 2004, p. 87). It is clear that an important assumption of conventional ethical theory is that moral judgments about an agent or an action are only appropriate when we can attribute responsibility. On the standard account in order to deserve praise or blame I must be the voluntary author of my actions. I bear responsibility for the harmful consequences I have deliberately brought about and not those arising by chance and beyond my control.

If this is the case, then an important pre-condition for Computer Ethics must also be that morality must be immune from luck; luck cannot determine or influence moral assessments. As Fearn (2005, p. 172) points out: "...Philosophers have spent most of the discipline's history denying that any such thing can take place, arguing that although luck can impinge on our physical and mental well-being, it cannot enhance or discolour our inherent virtue or lack thereof." It is this immunity from luck that partly marks out moral value from other kinds of values such as instrumental, technical, and use values. In this paper I want to explore the relationship between the idea of moral luck and Computer Ethics because of the threat it poses to our usual understanding of moral judgement. In particular the concept of moral luck bears down most heavily on notions of professional responsibility and how we identify and attribute such responsibility. If there is any mileage in the idea of moral luck then it seems important to get clear what it implies for moral judgement in professional matters. We often think of ethical theory as supplying justifications of how to make things go right but it is equally, if not more important, to consider, from the point of view of moral judgment, what happens when things go wrong and who is to blame, if anyone?

The Idea of Moral Luck

One can be lucky in many areas of life in one's friends, parents, children, health etc, but moral worth is an area which is meant to be uniquely free of the influence of luck. I don't deserve praise if I happen to do good by accident. This ideal

of freedom from contingency is most strongly proposed in Kant's concept of 'the good will': "....A good will is good not because of what it performs or effects, not by its aptness for the attainment of some proposed end, but simply by virtue of the volition, that is, it is good in itself, and considered by itself is to be esteemed much higher than all that can be brought about by it in favour of inclination, nay even of the sum total of all inclinations. Even if it should happen that, owing to a special favour of fortune, or the niggardly provision of a step-motherly nature, this will should wholly lack power to accomplish its purpose, if with its greatest efforts it should yet achieve nothing, and there should remain only the good will (not, to be sure, a mere wish, but summoning of all means in our power), then, like a jewel it would still shine by its own light, as a thing which has its whole value in itself." (Kant, 1873, pp. 12 – 13) In this way moral value is a special kind of value (as distinct from the 'provisions of step-motherly nature', accidents of birth, the favours of fortune e.tc.) and is supposed to be potentially accessible to all rational persons. The condition for the existence of this special (supreme) kind of value, if it is not in Bernard Williams' (1981, p. 21) phrase to be '...merely a last resort, the doss-house of the spirit', must be its immunity to luck. In addition our exercise of a good will provides for our own partial immunity to luck. On this view: "...Both the disposition to correct moral judgement, and the objects of such judgement, are ... free from external contingency, for both are, in their related ways, the product of the unconditioned will." (Williams, 1981, p. 20)

The moral philosopher, Bernard Williams, has challenged this Kantian ideal with the idea of moral luck. He argues that moral value may, like other values, be subject to luck. Williams (1981) goes on to try and show that moral judgment may also be conditioned and subject to contingency. This creates a paradox which poses a threat to our conventional understanding of morality and moral responsibility. If the moral decisions and the consequences of the actions of software designers, for example, are to a great extent subject to luck how then can we attribute responsibility? We appear to be morally responsible and not morally responsible at the same time. The paradox of moral luck is that our usual assumption is that I ought not to be judged on the basis of things and events not in my control but, on reflection, we are frequently held responsible for actions which may be greatly influenced by matters beyond our control. In fact an analysis of luck and risk seems to lead to the conclusion that very few outcomes are within our control. Williams illustrates this paradox in the case of Paul Gauguin, which we will examine below, arguing that only success ultimately provides justification for a (moral) decision making, in effect, such decisions moral gambles.

Williams focuses on the implications of moral luck for deontological ethics but also stresses the importance of outcomes; the 'good will' doesn't shine without action and my action may depend on a host of contingencies. Thomas Nagel (1979)

provides a more nuanced account of moral luck but again based on a critique of the Kantian assumptions of what makes moral value special. Nagel defines moral luck in the following way: it can be called moral luck "...Where a significant aspect of what someone does depends on factors beyond his control, yet we continue to treat him in that respect as an object of moral judgement...Such luck can be good or bad" (Nagel, 1979, p. 356). He develops a fourfold categorisation of moral luck: consequential luck (luck in the way things turn out); constitutive luck (luck in character - the kind of person you are in inclinations, capacities, and temperament); circumstantial luck (luck in the problems faced and decisions that have to be made); and finally luck in antecedent circumstances. Donna Dickenson in *Risk and Luck in Medical Ethics* (2003) develops both Williams' and Nagel's insights and applies them to various problems in medical ethics. She shows how an appreciation of moral luck sheds light on such problems as: withdrawal of life-sustaining treatments and assisted suicides; the allocation of health care resources; reproductive ethics; psychiatry and risk; genetics and moral character and proposes a synthesis of feminism, global ethics and moral luck! In applying the concept of moral luck to medical ethics Dickenson shows that its implications are not restricted to ethics in the Kantian tradition. Luck threatens also the notions of agency and action that are embedded in consequentialism and in virtue ethics. I hope to show similarly that moral luck has implications for theory and practice in Computer Ethics.

The intervention of luck raises problems equally for attempts to provide duty based, consequence or virtue based foundations to Computer Ethics. The threat that moral luck poses is to our understanding of what it means to be an agent at all and what it means to engage in action. And this, as we will see, goes to the heart of ideas on the nature of responsibility for computer professionals. Even if we can counter the threats of moral luck, its achievement may be to show that consideration of moral judgement, decision-making and the attribution of responsibility turns out to be a much more complicated affair than we have assumed. It has much more scope for the forces of risk and luck to intervene and muddy the waters. We may need to pay more attention to disentangling the 'favours of fortune' or the 'niggardly provision of step-motherly nature' than we have hitherto done. Much of what we do depends on things and events which are beyond our control. Donna Dickenson draws attention to an incompatibility between standards used to judge right actions and those to judge good character; this tension creates the paradox of moral luck: "...In practice we regard actions as right or wrong, and moral character as good or bad, partly according to what happens as a result an agent's decision. That is, we make responsibility hinge to some extent on things outside the agent's control. Yet at the same time we think that

people should not be held responsible for matters beyond their control” (Dickenson, 2003, p.1).

The Case of Paul Gauguin

Williams (1981) develops the application of the idea of moral luck through a discussion of the painter Gauguin. Williams begins his exploration with Gauguin’s decision to abandon his wife and children to pursue his artistic endeavours in the South Seas. This is clearly a failure to fulfil his duties to his wife and children in favour of a putative duty to his art. In the event Gauguin’s enterprise was successful in that he produced a set of highly valued works of art. Williams suggests that we say that Gauguin’s decision was vindicated by the fact that he was able to realise his talent and produce works of artistic genius. But this judgment is only possible in hindsight. Gauguin could not justify his decision in advance because he did not know that he would be successful. Suppose on arrival in the South Seas he had cut his foot, suffered an infection and quickly died before producing the works for which he is now justly famous. We might then imagine a footnote in some history of art commenting on the immorality of a minor painter who had abandoned his wife and children in favour of an easy life in the sun.

In mitigation we might, with Williams, want to say that here Gauguin was the subject of ‘brute’ bad luck. If this was the case then that we might be reluctant to hold Gauguin culpable for his failure. However, suppose Gauguin fails not because of any extrinsic factor such as illness but because in the first place he simply misjudged the potential of his own talent. In other words there was an intrinsic failure on his part to assess correctly his own abilities. In this case we should, I think be more inclined to hold him responsible for not fulfilling his duties to his family. But again the judgement is retrospective in that only success justifies or falsifies the original decision. In either case we cannot escape the fact that Gauguin’s decision was risky. It could not be justified in Kantian terms, i.e. as the act of a good will, precisely because at the time of making the decision Gauguin could not know what his duty to art was. If anything the pull of duty was towards his wife and children rather than some unproven talent. Similarly, if we try to approach this from a consequentialist position then, yes, the decision was justified but again only in retrospect. To justify the decision to leave his family from a broadly consequentialist point of view Gauguin would have to have known the consequences of his actions; he would have to have the foreknowledge that he would paint great works of art in the South Seas. But he just couldn’t have known this with anything like the required degree of certainty. Williams argues that the Gauguin example shows that the outcomes depend on luck and the outcomes cannot be known in advance. So if things had gone awry would Gauguin have been culpable or not? My argument is that similar questions begin to arise when we

consider Computer Ethics and professional responsibility. Similar vulnerabilities to risk and luck are exposed.

Codes and Responsibility

Before pursuing the concept and types of moral luck I want to look briefly at professional responsibility to understand why the notion of moral luck might be perceived to be threatening to the whole enterprise. Simon Rogerson argues that professional codes should address a number of levels of ethical obligation embracing core duties that professionals have in their working relationships. These include: a duty to share with the common stock of moral values of humanity; a higher order duty of care to those affected by the professional's work; and specific duties specific to being a computer professional. (Rogerson, 2002, p. 177). Don Gotterbarn distinguishes a range of such codes: codes of ethics, codes of conduct or codes of practice. He points out that these three types of code parallel levels of professional obligation (Tavani, 2004, p. 95).

By implication whatever else they may do codes lay out standards by which professional performance might be judged from a moral point of view. Remember that we are in the domain of 'non-technical ethics questions'. Now those standards might be derived from common standards of morality or might be derived from specific activities of the profession. Nevertheless they provide a means of assessment providing a standard by which we can reach a moral judgement, attribute praise or blame and identify right and wrong professional behaviour. Central to this process is the belief that professionals are autonomous, rational agents who could always do otherwise than what they in fact do; I could choose to fulfil my duties to my employer or I could cheat her (conduct); I could design well or negligently (practice); I can promote social benefit or social harm (ethics). The importance of moral luck in all this is that the making of an ethical decision and how things turn out may be a contingent matter and beyond control.

Let's look at some examples: members of the ACM commit themselves to a general responsibility to contribute to society and human well-being (1.1) in the Code of Ethics and Professional Conduct: "This principle concerning the quality of life of all people affirms an obligation to protect fundamental human rights and to respect the diversity of all cultures. An essential aim of computing professionals is to minimize negative consequences of computing systems, including threats to health and safety. When designing or implementing systems, computer professionals must attempt to ensure that the products of their efforts will be used in socially responsible ways, will meet social needs, and will avoid harmful effects to health and welfare" (Tavani, 2004, p. 315). Now let us suppose that things go wrong and the system that I have been involved in designing is used in socially

harmful ways, for example, say in the distribution of personal credit information to corporate agencies that I don't approve of. I may not have known this when I was working on the system. As far as I'm concerned this is an immoral but an unlucky outcome. How culpable am I – if I couldn't have known? What weight do we put upon the 'must attempt' in the code? Nevertheless if we judge on the basis of success or failure – it certainly looks like I've failed in my commitment here. The outcome here may cause me, or others, to revise the moral assessment of my original actions.

Take a further example which focuses on a more specific obligation (2.6) which lays out duties about honouring contracts, agreements, and assigned responsibilities: "...A computing professional has a responsibility to request a change in any assignment that he or she feels cannot be completed as defined. Only after serious consideration and with full disclosure of risks and concerns to the employer or client, should one accept the assignment. The major underlying principle here is the obligation to accept personal accountability for professional work. On some occasions other ethical principles may take greater priority...A judgement that a specific assignment should not be performed may not be accepted. Having clearly identified one's concerns and reasons for that judgment, but failed to procure a change in that assignment, one may yet be obligated by contract or by law, or proceed as directed. The computing professional's ethical judgement should be the final guide in deciding whether or not to proceed. Regardless of the decision, one must accept the responsibility for the consequences" (Tavani, 2004, pp. 318 - 319).

It seems to me that the assumptions that underlie such obligations are those of the classical rational Kantian agent '...the computer professional's ethical judgement should be the final guide'. As I hope to go on to show such assumptions are vulnerable to risk and moral luck. Equally, from a consequentialist point of view you just have to know the outcomes in order to justify the decision. But if success is the criterion then you only know retrospectively. Well you might say I just thought it was likely – there was a risk – so a change was requested. Suppose the change is accepted and it turns out badly? The fact that it turned out badly may be a matter of luck. But I was instrumental in requesting the change. Is it a case of better luck next time or to what degree am I responsible? I believe we too easily fall into the trap of underestimating the role of luck and assuming neat chains of causality will settle clearly the matter of responsibility.

Ethics and Testing

Professional responsibility, as we have seen, goes beyond the formal and technical aspects of system development. For example, information systems developers should consider the consequences of implementing a system on people, organisations and society generally (Rogerson, 2002). Now if we are to take these injunctions seriously it must imply that developers may be blamed if things go wrong when they do not take wider societal issues into consideration. However, we can again see the possibilities for a role for moral luck. Suppose a project team does not seriously consider the consequences of their design for the client, the users and the ultimate success of the system. But suppose, nevertheless, (and luckily) everything turns out more or less to everyone's satisfaction. The ethicist still might want to say to the team, or the team leader: "In spite of the fact that everything turned out well you were at fault; you ignored risk assessments, ethical protocols, e.tc..." If we judge by consequences then the reply might be that success is the test and we have, by that test, increased rather than decreased social utility. Again this suggests the paradoxical nature of moral luck.

It is the case that being sensitive to all the duties of professionalism, including ethical considerations in the design process, and considering wider consequences by no means guarantee that everything will turn out well. The range of possibilities I believe is something like this: (a) I can deliberate well over moral matters and things can go well; (b) I can deliberate badly over moral matters and things still may go well; (c) I can deliberate well over moral matters and things may go badly; (d) I can deliberate badly over moral matters and things can go badly. There is no necessary connection between the quality of deliberation and outcomes. But isn't it more likely that we will get good outcomes if we deliberate well? Similarly, if we test exhaustively shouldn't that reduce the chances of things going wrong? But testing software always involves some trade-off between the costs of testing and the technical possibilities of testing. This involves a combination of practical (financial), technical and moral decisions. In theory the more we test the more we should be confident that we can accurately predict the reliability of the software's behaviour. This assumes a linear relationship between error detection and the amount of testing i.e. the outcomes and effort in testing. However, there is an asymmetry between justification and falsification. Many successful tests do not necessarily 'prove' the soundness of the system. But one failure certainly falsifies the claims of success. The larger the scope and complexity of the software's behaviour the greater are the possibilities of failure. The point is that we cannot exclude the operation of chance and hence of moral luck.

The defence against unforeseen or unintended consequences may be a Kantian one: "Oh that doesn't trouble us unduly. Even though things turned out badly we

acted with the right intention and did as much as might be reasonably be expected to ensure success. It's just bad luck that the system you commissioned doesn't work as intended. Statistically our company usually gets it right. Better luck next time!" This intuitively doesn't seem a right or plausible response. But lest you may think this is just fanciful consider the history of failures of major public IT projects in the UK (Booker and North, 2007). Again what arises from this discussion is the paradox that we are urged to treat computer professionals and their behaviour as objects of moral judgement whilst at the same time we recognise that significant aspects of what happens in system implementation and performance may depend on factors beyond their control i.e. the paradox of moral luck. To resolve this apparent paradox we could just absolve professionals of responsibility for outcomes beyond their capacity and ability to control or predict. This might mean the level of professional responsibility would tend towards zero. In the next section I review Nagel's attempt to provide a more refined analysis of the nature of moral luck.

Four Types of Moral Luck

Like Williams, Nagel (1979) accepts that moral luck raises problems for conventional ethical theory. He focuses on the tension between responsibility and lack of control leading to the erosion of moral assessment. Nagel produces a typology of moral luck on the basis on the criterion of the kinds of lack of control associated with each type. For example, in 'circumstantial luck' we lack control over the kinds of moral decisions we have to face.

Consequential Luck

Nagel specifies consequential luck as luck in the way an agent's actions and projects turn out. We often lack control over the factors that determine the consequences of the decisions we take. In this category Nagel plays with a number of variations where the luck involves either negligence or uncertainty. Things may turn out badly even without any negligence being involved. Suppose I have designed a semi-autonomous life support machine and then a patient dies because the machine has ceased to function. We discover that the machine has been accidentally disconnected through electrical work in some other part of the building. I might feel regret about the outcome since my role had been to provide a machine that technically did not need supervision. Inadvertently I have created the opportunity for the accident to happen to the patient. In other words let's assume that no negligence (on my part as the designer) was involved here - I might have some regret but this would not on Nagel's account be a case of moral bad luck.

But the case is altogether different if we consider negligence. Suppose there is a bug in the system such that the machine automatically shuts down after twenty one days without warning. If this happens and a patient dies then we clearly have a case of negligence in the design and testing of the software. Suppose we have the same situation but, largely by chance, no patients ever have been supported by the system beyond twenty days. No one dies. In other words things seem to turn out well, patients are treated satisfactorily, and I win an award for my work. But in this case things don't turn out badly only by chance. However, in both these instances the negligence is the same. As the designer I'm morally lucky that no patients have died but, it can be argued, I'm equally negligent as in the case where the patient died; it's just that my negligence has not yet come to light. The paradox is that in the first instance the designer is to be blamed but in the second case is praised.

Nagel argues that consequential luck is particularly evident in cases where decisions are made under uncertainty or ignorance. This should come as no surprise in that our understanding of what may happen in the future is at best fragile and at worst a matter of ignorance. The uncertainties under which we labour are a function of our inability to foresee the outcomes of our actions and our inability to control the actions of others. Prior (1956) argues that this imposes the logical condition that we can never define what the content of our duty might then be from the perspective of a prospective assessment of consequences. Forecasts cannot be known to be true at the time they are made and both their factual and moral evaluation must follow as a matter of hindsight.

If we consider, for example, the macro-level forecast made by Ray Kurzweil (2005, p. 9) that by around 2050 we will be achieving what he calls 'the Singularity' in which through the accelerating pace of technology we will transcend '... the limitations (sic) of our biological bodies and brains'. Kurzweil's claims are so empirically extensive that they are equally extensively vulnerable to falsification. So suppose his predictions turn out to be wildly wrong. For example, some intrinsic scientific and technical limits are reached in the development of computer technologies and, what's more, collectively we come to value simpler and less technologically driven lifestyles. But in the meanwhile much public money has been spent on the fruitless pursuit of his general programme. I have argued elsewhere that indeed forecasters should be held to be morally responsible for their predictions (Horner, 2007). Would Kurzweil be then morally culpable for the bad prognostications he gave and the waste of public money given that the ultimate outcomes were a matter of luck? If things don't turn out as Kurzweil expects is he off the hook in the sense that we can say this is just another case of bad luck and his intentions were honourable?

If we take, for example the case of the so called 'millennium bug', the Y2K problem, even with hindsight it is far from clear that the public and private money spent on remedial action was spent to good effect (Booker and North, 2007). The predictions of catastrophe, the national and global breakdown of computing infrastructure, were not realised. Was this result of the remedial action or a matter of luck? If we could show that much of the money poured into national programmes to stave off catastrophe were unnecessary and, in effect, wasted, would the forecasters and policy makers be morally culpable? In the absence any comprehensive empirical studies that might establish one way or another whether the remedial action was responsible for staving off global catastrophe the suspicion remains that the outcome might just have been a product of bad prediction and consequential good luck.

The strong argument for the paradox of consequential moral luck has been formulated in this way (Dickenson, 2003, p. 13.):

A person P is morally responsible for an event e's occurrence only if e's occurring was not a matter of luck.

No event is such that its occurrence is not a matter of luck.

Therefore no event is such that P is morally responsible for its occurrence.

I reformulate this to take account of the morality of forecasting:

A person P is morally justified (and hence praiseworthy) by the correct prediction of an event e's occurrence only if e's occurring is not a matter of luck.

No event is such that its occurrence is not a matter of luck.

Therefore, no event is such that P is morally justified (and hence praiseworthy) by its occurrence.

The choice presented by these arguments seems to be either to absolve agents of any outcomes that they cannot predict or control or to hold them blameworthy or praiseworthy for action for which they are not responsible. Neither horn of this dilemma seems intuitively acceptable. But then how ought we to attribute moral responsibility and liability in the absence of foreknowledge? Should we be held responsible for the contributions of fate as well as for what we ourselves do?

Of course it is open to the sceptics to cast doubt on the second premise and defend the possibility of being able, with reasonable accuracy, to predict the outcome of actions. It seems plausible to resist the idea that because there is some uncertainty then we should abandon all hope of reasonably inferring consequences and predicating our decisions on the likelihood of beneficial consequences. It would be logically odd just to say I'm going to act without thinking of con-

sequences at all (but isn't this the Kantian position?). However, the question is about moral responsibility in the light of success or failure. "...If only success can justify a decision, but success is not certain at the time the agent makes a choice, there will turn out to be no basis but hindsight for judging whether an action was morally right or wrong" (Dickenson, 2004, p. 15). If things go unexpectedly wrong then the decision isn't justified, and if the decision isn't justified at the time it was made then the agent was culpable.

In the light of all this we might simply want to play the realist card and accept the uncertainty in moral decision making and accept, as in other areas of our lives, we are to some extent 'playing the odds'. We can't be certain about outcomes of our professional behaviour but we make a reasonable assessment of the possibilities on which to found a moral decision. But the problem with this approach is simply that it dethrones moral value from its special status amongst other values. Morality is not meant to be like other parts of our lives. There is no problem in accepting the role of chance and uncertainty in our worldly achievements. But we began this paper by asserting that it was precisely morality's immunity to luck which gave its special status. At the very root of Computer Ethics is the assumption that moral values are precisely more important than instrumental, technical or use values. We want to be able to say, for example, that although the technical challenge of building ever more sophisticated and undetectable surveillance systems is exciting yet for moral reasons we are just not going to work on such systems.

In summarising the paradox of consequential luck Baggini and Fosl (2007, p. 223) write that: "...To deny that moral luck exists at all, however, one needs to deny that actions become better or worse depending on what their consequences are, since what actually happens is almost always beyond anyone's full control. But this option also seems counter intuitive: surely it does matter what actually happens. To judge people purely on the basis of their intentions or on the nature of the act itself seems to diminish the importance of what actually happens."

Constitutive Luck

Nagel's second form of moral luck, 'constitutive luck', is luck in our inclinations, capacities and temperament where the kind of person we are is not dependent on what we deliberately do. The idea here is that our moral worth may depend on a chance formation of character beyond our control. A kind person has not fully chosen to be kind but that's just how she grew up. Constitutive luck undercuts both a virtue theory of ethics and the Kantian tradition. Our vices and virtues may be given rather than chosen and there may be limits within which they may be cultivated or re-shaped. Genetic and environmental may be crucial in shaping, if not entirely determining, our character: "... People are morally condemned for

such qualities, and esteemed for others equally beyond the control of the will: they are assessed for what they are like" (Nagel, 1979, p. 361). We don't choose our parents, genes, culture of origin e.tc. Our ability to respond morally may depend to a large extent on the constitution of our character for which we are largely not responsible.

In the Kantian tradition the domain of moral responsibility is restricted to 'the inner world' of the pure, good will. Outcomes may be relevant but ultimately do not count; success is irrelevant to moral evaluation. The only thing that matters is the exercise of the good will. Our inwardness in the exercise of moral judgement immunizes us from the contingencies of the outer world. Events may be out of our control in the workings of the world but, in contrast to Williams, this should not matter. In addition Nagel points out that, for Kant, even qualities of temperament and personality that are not directly under the control of the will are morally irrelevant: "...a person may be greedy, envious, cowardly, cold, ungenerous, unkind, vain or conceited but behave perfectly by a monumental effort of will" (Nagel, 1979, p. 361). He argues that to possess such traits, on which we are often judged, is simply a matter of 'constitutive bad fortune'. Even if by the exercise of will I control such qualities they are nevertheless my qualities. I may have qualities that I do not display or are not manifest in my behaviour; I may be envious of your success but do not show it. Such qualities may arise as a result of previous choices I have made, and may be ameliorated by present behaviour, but still largely a matter of constitutive bad fortune. And if certain dispositions in my character are not within the control of the will then I'm not in that sense responsible on the standard account.

Patterns of human personality and dispositions are deeply ingrained and change little over time. What we still do not know is the extent to which character might be the outcome of genetics. Whilst not suggesting here some crude genetic determinism it does seem to be the case that genetics suggests that human character is less malleable than we often believe. At one level that make it more immune to luck in terms of formative experiences and past choices but at the same time plunges back into the genetic lottery. In other words whether or not as a computer professional I'm inclined to pay attention to codes of ethics or not may be a matter of constitutive good or bad fortune.

The idea of constitutive luck can be illustrated by examples of whistle-blowing. Tavani gives us the example of David Parnas in the 1980s who was a consultant on the Strategic Defence Initiative (SDI). Parnas became convinced that it was not possible to design software for the project that had any likelihood of working. He gives three reasons why he thought that the SDI was not viable: "... the specification of the software could not be known with any confidence; the software could not undergo realistic testing; and there would not be sufficient time dur-

ing an attack to repair and reinstall failing software (no 'real-time' debugging)." (Tavani, 2004, pp. 103 - 105) For Parnas these reasons justify his decision to walk away from the project (and a lucrative consultancy contract) and go public with his criticisms. His detractors, supporters of SDI accused Parnas of being disloyal - a negative judgement on his character.

From an ethical point of view a key aspect of the debate about whistle-blowing concerns the criteria which might be used to determine when we are permitted or required to 'blow the whistle' (Tavani, 2004, pp. 102 - 105). The points of contention hinge on the relative stringency or leniency of such criteria. Should whistle-blowing be 'required' only in exceptional circumstance or be required more generally when there is a demonstrable *prima facie* moral obligation to expose wrong-doing that we cannot directly stop. Now it seems to me that the idea of constitutive luck cuts across this type of debate. We might suggest, in the tradition of virtue ethics, that this is more a matter of intrinsic strength of character rather simply a calculus of probabilities based on external criteria. If we consider the implications of our luck in our inclinations, capacities and temperament then it may be that some people, like Parnas, just have the disposition that enables them to be able to take a strong stand, subordinate other considerations, and make sacrifices for a greater social good. But from the perspective of constitutive luck - well they are just lucky! In contrast we would tend normally to praise the whistle-blowers because they act according to conscience. But suppose 'moral heroes' are born and not made (or self-made) then according to our conventional reasons for attributing praise or blame they are not particularly deserving. If I am only obligated to do what I am able to do ('ought implies can') but I simply don't have the character then ought I to be blamed for carrying on when Parnas leaves? Again there seems to be something logically odd about all this. The puzzle arises, I think, because we cleave to the Kantian model of the rational agent whose good will is immune both from the contingencies of character or the contingencies of circumstance or to the theory of virtue ethics which assumes we can bend our character to our will.

Circumstantial Luck

Circumstantial luck is luck in the problems and decisions that have to be faced. The kinds of moral tests we may face are largely the result of circumstances beyond our control. (Remember according to the standard account I should be held responsible only for what is within my control). To a certain extent then whether I'm tested or not is a matter of luck. If I'm never tested how do I know that I act morally? The classic example is that of resistance to a totalitarian state (Nagel, 1979, p. 361). If I have never lived in a country governed by a tyrannical and brutal dictatorship (it's not hard to think of examples here) then I don't know

how I would respond to such a situation; would I respond to the call of resistance that might entail dire consequences for both for myself and my family? Consider the whistle-blowing example again. If I'm never in the situation where I believe that the work I'm involved with, and the organisation that I work for, is engaged in projects that might produce social harm, then I will never have the chance to behave in a brave or cowardly fashion. In this way my moral record, the way I might be judged from a moral point of view, is dependent on whether morally challenging situations arise or do not arise. This will be a matter of luck. Well, as some do, I might go in search of moral challenge but again the extent to which I'm tested may still be a matter of luck. It must be the case that we judge people, if we can, by what they do or what they have not done and not according to what they would or would not have done.

Richard Norvin (1986) has criticised this notion of circumstantial luck. He argues that in a sense the luck does not reside in the circumstances. Take for example two negligent drivers. In the first case the driver is negligent but does not actually run any one down. In the second case a child runs out in front of the driver's car and he kills the child. For Norvin the function of the circumstances in the second case is merely to make clear the negligence of the driver but both drivers are equally culpable for their negligence. So called circumstantial moral luck does not alter moral assessment. Circumstances only render good or bad characters transparent to others. If we've not been confronted by a moral dilemma we remain untested and our moral character is just not revealed. Norvin's response then is that the role of luck is more a question of our uncertainty and ignorance, our 'epistemic position', than of what does or does not determine an agent's deserts.

Luck in Antecedent Circumstances

Nagel's fourth category of moral luck is luck is how one is determined by antecedent circumstances. Again the argument hinges on the diminishing area of agency under the pressure of the combined influence of factors that are antecedent and posterior to action. (Nagel, 1979, p. 363) The extent to which our acts of will and moral decision making are merely the products of antecedent circumstances further undermines our responsibility. If we are not responsible for antecedent circumstances how then can we be responsible for the results which arise from those circumstances. What we are able to achieve is severely circumscribed.

For example, let's consider the well known case in 1988 of the shooting down of an Iranian passenger aircraft, killing 230 people, by the USS Vincennes. The ship was equipped with the Aegis Radar System which had been developed by the United States Navy. The system enabled warships to monitor the airspace around them. The accidental shooting down was blamed by some commentators on a

flawed design of the system's user interface. In response some computer ethicists (Tavani, 2004, p. 111) argued that system designers should be more aware of the need to design into the interfaces of safety critical systems features that take account of human limitations and abilities. However, the Aegis Radar System had met all the requirements that the customer and developer had specified. In spite of this 230 airline passengers died. Who was responsible here? We might locate responsibility in the chain of command on the Vincennes. We equally might want to implicate the software designers. Don Gotterbarn, for example, argues that such software failures are not inevitable but arise primarily from: (a) a narrow conception of risk; and (b) a limited notion of system stakeholders. This is no doubt the case.

However, let's look at this from the point of view of antecedent luck. If the surrounding circumstances had been different then there may have been a different outcome. If a different set of surrounding circumstances had prevailed then the judgement of the action would have been assessed differently. Suppose, for example, the aircraft had been a military aircraft, threatening the Vincennes and not a civilian aircraft. The moral assessment of the case might have been different as well as the technical assessment of the system! Previous decisions determined that the Vincennes should be at that particular position and at that particular time. Similarly if the civilian aircraft had been on a different flight path and at a different time then it would not have encountered the Vincennes. So if the accident had not occurred then there would have been no debate. There would have been no case for moral blame or regret by the ship's command or the software designers. Such considerations appear to shrink the scope of genuine agency. (And remember in this case the designers fulfilled their brief!) The argument is that the accident resulted from an unlucky confluence of events – a product of antecedent circumstances. "Everything seems to result from the combined influence of factors, antecedent to and posterior to action, that are not within the agent's control..." (Nagel, 1979, p. 363).

A critic might reasonably reply: "That's all very well but the circumstances were not other than they in fact were, and the act was what it was, and 230 people died. Responsibility must lie in some combination of the decision to shoot down the aircraft and the imperfections in the system design that led to such a decision. We might concede that had some previous decisions been different then the shooting down may not have taken place. But the software was as it was and the decision to bring down the aircraft in fact was taken. We can argue that even if the outcome was the result of many factors nevertheless it still makes sense to talk of responsibility". However, I believe Nagel would still want to argue that: "...This form of moral determination by the actual is also paradoxical, but we can begin to see how deep in the concept of responsibility the paradox is embedded.

A person can be morally responsible only for what he does; but what he does results from great deal that he does not do; therefore he is not morally responsible for what he is and is not responsible for. (This is not a contradiction, but it is a paradox)" (Nagel, 1979, p. 362).

Conclusions

The question of luck and moral luck seems to me an unjustly neglected topic within Computer Ethics. The aim in this paper is a relatively modest one of trying to describe the nature of moral luck and to see how the concept of moral luck might raise problems in the field. It seems that these problems relate to the kinds of ethical theories that are typically employed in understanding moral questions in Computer Ethics. I have argued that moral luck challenges some basic assumptions about agency, action and responsibility. For example, is it right that the moral assessment of an action should rest on its consequences? I have argued also that it has implications for the attribution of moral responsibility in the professional behaviours of computer professionals. Baggin and Fosl restate the paradox of moral luck in this way "... Screening out those dimensions of a situation attributable to luck may leave little left to praise or blame. So, however one looks at it, accepting the role of luck presents a major challenge to judgments of moral praise and blame—but perhaps something essential, too." (2007, p. 223) The dilemma we are presented with is that on the one hand the role of luck and risk cannot be denied but the concept of moral luck seems oxymoronic. Moral value is the one kind of value that we have assumed is detached from the vicissitudes of fortune. We might be content to accept the idea that to some extent the fact of moral luck does lead to decisions being a kind of gamble. The problem is the resulting implication that moral value is no longer the kind of special value that it is assumed to be in conventional ethical thinking. If nothing else, I hope this analysis suggests that the problems of theory and practice in moral decision making in Computer Ethics are perhaps more complex and uncertain than we normally assume.

REFERENCES

- Baggini, J. and Fosl, P.S.,** (2007). *The Ethics Toolkit: a Compendium of Ethical Concepts and Methods*. Blackwell, Oxford.
- Booker, R. and North, R.,** (2007). *Scared to Death from BSE to Global Warming: why some scares are costing us the earth*. Continuum, London.
- Dickenson, D.,** (2003). *Risk and Luck in Medical Ethics*. Polity Press, Cambridge.

Fearn, N., (2005). *Philosophy: The Latest Answers to the Oldest Questions*. Atlantic Books, London.

Horner, D.S., (2005). Anticipating ethical challenges: Is there a coming era of nanotechnology? *Ethics and Information Technology*. 7, 127 – 138.

Horner, D.S., (2007). Digital futures: promising ethics and the ethics of promising. In: L. Hinman et al. eds. *Proceedings of CEPE 2007: The 7th International Conference of Computer Ethics: Philosophical Enquiry*. University of San Diego, July 12 – 14, 2007. Center for Telematics and Information Technology, Enschede, The Netherlands, 194-204.

Kant, I., (1773). *Fundamental principles of the metaphysic of morals*. In: T.K. Abbott, trans. *Kant's Theory of Ethics or Practical Philosophy*. 1st ed. Longmans, Green, Reader and Dyer, London.

Kurzweil, R., (2005). *The Singularity Is Near: When Humans Transcend Biology*. Gerald Duckworth, London.

Nagel, T., (1979). Moral luck. In: Louis P. Pojman, ed. *The Moral Life: An Introductory Reader in Ethics and Literature*. 2nd ed. Oxford University Press, New York, 2003.

Norvin, R., (1986). Luck and desert. *Mind* 65, 198 – 209.

Prior, A.N., (1956). Symposium: the consequences of actions. *Proceedings of the Aristotelian Society – Supplement*. 30, 91-99.

Rogerson, S., (2002). Computers and society. In: R.E. Spier, ed. *Science and Technology Ethics*. Routledge, London, 159 – 179.

Tavani, H.T., (2004). *Ethics and Technology: Ethical Issues in and Age of Information and Communication Technology*. John Wiley & Sons, Hoboken, N.J.

Williams, B., (1981). *Moral luck: Philosophical Papers 1973 – 1980*. Cambridge University Press, Cambridge.

Intellectual Property issues for digital libraries in the Internet networked public sphere

Dionysia Kallinikou*

Professor of Law, Athens Law School,
National & Kapodistrian University of Athens

Marinos Papadopoulos**

Attorney-at-Law (Athens Bar Association)

Alexandra Kaponi***

Attorney-at-Law (Athens Bar Association)

Vassiliki Strakantouna****

Librarian at the Civil Law School Library,
Athens Law School,
National & Kapodistrian University of Athens

* Dionysia Kallinikou. She is Professor of Law in Athens Law School at the National and Kapodistrian University of Greece, and an Attorney-at-Law. She has established a reputation as an Intellectual Property expert, and authored many books and scientific articles upon Intellectual Property. She served as the Director of Intellectual Property Organization in Greece, and as a Project Leader in European Community programs related to a variety of Intellectual Property issues.

** Marinos Papadopoulos. He is an Attorney-at-Law registered in Athens, Greece. He holds a law degree from Athens Law School and a graduate degree Master of Science from Boston and Harvard Universities. He has, also, graduate studies at Stanford University upon Information Technology and Law as well as The George Washington University upon Management. He is an active participant in international fora upon issues of Information Technology & Law as well as Information Society. He is a Legal Lead for Creative Commons in Greece. (Further info at URL: <http://www.marinos.com.gr>)

*** Alexandra Kaponi. She is an Attorney-at-Law registered in Athens, Greece. She holds a law degree from Athens Law School and a graduate degree Legum Magister from Heidelberg University. She is a Judge-Arbitrator of the European Arbitration Court on .eu domain names, and an Attorney-at-Law working for the National Telecommunications & Post Commission in Greece on .gr domain names. She is an active participant in international fora upon issues of Information Technology & Law as well as Information Society. (Further info at URL: <http://www.marinos.com.gr>)

**** Vassiliki Strakantouna. She is a Librarian at the Civil Law School Library of the National & Kapodistrian University of Athens. She holds a degree on Psychology from the Department of Psychology, Faculty of Philosophy, Pedagogy & Psychology of the National and Kapodistrian University of Athens, and a graduate degree Master of Information Science from the Ionian University, Department of Archives & Library Science. She holds also a degree on Drama (School of Drama 'Eugene Hatzikou') and a degree on Administration & Economics (Technological Education Institute of Athens, Department of Library Studies). She is an active participant in Greek fora upon issues of Libraries & Information Society.

Abstract

The development of Digital Libraries and repositories, a worldwide vision with enormous political and ideological importance for humanity, in an effort to approach cultures and preserve plurality and diversity, is directly affected by the provisions of Intellectual Property Law and is subject to the consideration of innovation through legislation. Legal issues such as these related to software use, database protection, the collection, digitization, archiving, and distribution of protected works are of outmost importance for the operation and viability of Digital Libraries and repositories. In this whitepaper, we focus upon some of these legal issues and consider an alternative proposal in respect of Intellectual Property law for open access to creative works furnished to the public through Digital Libraries and repositories.

Keywords: Intellectual property, digital libraries, collecting societies, peer-to-peer, copyright, file-sharing, copyright right-holders, Law 2121/1993, three-step-test, openness, open access, open access journals, digital rights of copyright, Creative Commons

Introduction

Information technology enables the reproduction, saving, and distribution of culture, arts, and sciences as well as the recording of the collective memory of humanity. Not only does information technology make almost all kinds of human creativity available in the Internet networked public sphere, but also it enables the creation of new forms of art, creative expression, and distribution of knowledge. More often than not, the rapid pace in the evolution of information technology causes friction with Law in as much as Regulators' foresightedness could not have ruled to cope with new social trends, socio-political and economic phenomena in the market. Thus, the evolution of information technology is frequently seen as a factor that sets at stake the legal rights of creators and right-holders, as a cause for stricter Intellectual Property Law and legal protection for the initial and subsequent right-holders. The instantaneous reaction towards making the Law stricter regarding the availability, use, and distribution of creative works via the Internet networked public sphere is—in most cases—a hazard to the evolution of digital libraries (hereinafter, DL) and repositories, i.e. organizations which leverage upon the radical changes caused by information technology and the tremendous capabilities that it has inferred upon the availability, use, and distribution of culture, arts, and sciences to the people. Making the Intellectual Property Law stricter in consideration of the rights of creators and right-holders, and most commonly the financial interests of them associated with the exploitation of their works in the markets, is not an option that satisfies at least to the point that it

does not sufficiently cater for the interest of people and/or DL and repositories in having access and making use of creative works leveraging upon the dynamic of new Internet networked media. We have reached a point at which the challenge to amend the Intellectual Property Law with the aim to consider favourably the changes in society caused by information technology, and to balance the conflicting interests of all the involved parties, i.e. creators, right-holders, and the people, in the creation, use of, distribution, and re-creation (remix) of protected works, is bigger than ever, at least in the history of Intellectual Property Law.

The development of DL and repositories, a worldwide vision with enormous political and ideological importance for humanity, in an effort to approach cultures and preserve plurality and diversity, is directly affected by the provisions of Intellectual Property Law and is subject to the consideration of innovation through legislation. Legal issues such as these related to software use, database protection, the collection, digitization, archiving, and distribution of protected works are of outmost importance for the operation and viability of DL and repositories. In this whitepaper, we will focus upon some of these legal issues and consider an alternative proposal in respect of Intellectual Property law for open access to creative works furnished to the public through DL and repositories. Open access for DL is a *sine-qua-non* prerequisite for their viable operation. It's like oxygen for human beings.

DL & ethics

DL acquire, organize and secure life-long access to creative works which are the building blocks of our civilization. Libraries are the repositories of human knowledge; they are our past, our present and our future.¹ Aside from the focus, the special collections and the added-value services which a DL may encompass, the main reason for the existence and development of any DL is the need to serve people in their quest to access knowledge. DL have always been gates through which people could access knowledge hosted in the premises of these organizations. Libraries have adopted internal regulations, abide by national laws and international practices with the aim to achieve their goals, i.e. satisfy people's quest for knowledge and creative works. While almost all libraries and librarians acknowledge the need to abide by Intellectual Property law that protects the interests of authors and creators, initial and subsequent right-holders, they do also show unwillingness to transform from organizations that enable access to knowledge into gatekeepers of locked-in, inaccessible knowledge.

1. Mason, Moya K. (2009) The ethics of librarianship, available at URL: <http://www.moyak.com/papers/ethics-librarianship.html> [last check, April 10, 2009].

The founding principle in accordance of which libraries ought to operate in a way that considers the interests of all, i.e. creators and authors, initial and subsequent right-holders, and the general public, and manage the works and collections hosted to them in a way that does not deviate from the framework of law creates legal and ethical obligations for librarians. The ethical and/or professional conduct of librarians ought to depict their commitment to society to acquire, organize and secure access to the elements of civilization, and should stress the important role that librarians could play in the wide distribution of knowledge.² The most profound ethical issues or dilemmas facing librarians concerns censorship, selection of materials and intellectual freedom, copyright, patron privacy, computer use, the Internet and plagiarism (Dole, 2000).³ Also, they do have a call upon issues such as the way that individuals make use of the copyrighted works and the extent that creators of such works control theirs use and dissemination. During the last years ethical challenges presented to librarians and information workers have increased dramatically and the ethical dilemmas faced are numerous. Technological innovation, for example, enables us to create "brave new worlds." But automated environments are unfamiliar worlds. Our old intuitive habits of evaluation, which are adequate for determining what is best in traditional worlds, are inadequate in new and different settings (Severson, 1995).⁴

The widely respected principles of ethical and professional conduct in librarianship could greatly assist in the development of a wide network of organizations allowing for access to, use of, and distribution of knowledge in society irrespective of factors such as financial power that keeps many members of it deprived of knowledge for sale. The codes of ethical and professional conduct in librarianship usually have the form of a set of rules for self-regulation⁵ and describe the principles of conduct that govern the librarians of a certain organization and/or professional librarians of a wider group of peers. Recognizing the importance of having a code of ethics, library associations have a long history of developing and

2. Rubin, R. (2000) *Foundations of Library and Information Science*, New York, Neal-Schuman publ., p. 265-296.

3. Dole, W.V. and Hurych, J. M. and Koehler, W. C. (2000) Values for Librarians in the Information Age, *Library Management*, 21(6), 285-286.

4. Severson, R. (1995) The recovery of ethics in librarianship, *Journal of information ethics*, 2(2), 11-15.

5. By the term 'self-regulation' we mean the adoption of regulation and a framework of obligations and rights which are created through voluntary commitment of all members of a certain community. Self-regulation is a kind of self-governance in cyberspace. See Mitrou, L. (2002) *The Law in the Information Age*, Sakkoulas.

promoting ethics for the profession (Hoffman, 2005).⁶ The promotion of open access to knowledge and information, the protection of privacy, the respect for copyright and right-holders, and the freedom of speech are some of the founding principles that govern most of these codes of conduct for professionals in the industry of libraries.

Copyright is a legal concept part of a broader notion of intellectual property with critical legal and ethical issues for the librarianship community. Copyright compliance is both a legal and an ethical issue. An information professional need to ensure that his activities remain on the right side of the law and that his conduct is ethical.⁷ The principle of information professionals to recognize the rights of creators and copyright holders of copyright-protected library and information material is explicitly stressed in most librarian's codes of conducts.⁸ To name a few, the codes of conduct of the German Association of Libraries and Information Scientists (BID)⁹ and the American Library Association (ALA)¹⁰ are notable examples of such texts of self-regulation in librarianship. In the UK the CILIP's ethical principles and code of professional practice make clear that the conduct of members should be characterized by "Respect for, and understanding of, the integrity of information items and for the intellectual effort of those who created them" and members has the responsibility to "defend the legitimate needs and interests of information users, while upholding the moral and legal rights of the

6. Hoffman, K. (2005) Professional ethics and Librarianship, Texas library journal, available at URL: www.txla.org/pubs/tlj81/Ethics.pdf,

7. Pedley, P. (2007) Digital Copyright, Facet Publications.

8. See International Federation of Library Associations (IFLA), the Codes of Conduct adopted by Librarians Associations of various countries, most of which stress the importance of balancing the conflicting interests of library-users and copyright-holders. See Codes of Conduct through URL: <http://www.ifla.org/faife/ethics/codes.htm>, [last check, April 10, 2009].

9. See IFLA/FAIFE, BID's code (Bibliothek und Information Deutschland), stand March 19, 2007, according to which is made clear that the information professionals recognize the rights of creators and copyright holders of copyright-protected library and information material. BID's code is available at URL: http://www.ifla.org/faife/ethics/germany_code_of_ethics-de.htm [last check, April 10, 2009].

10. See IFLA/FAIFE, ALA, Code of Ethics, 1995, available at URL: <http://www.ifla.org/faife/ethics/alacode.htm> [last check, April 10, 2009] & at URL: <http://www.ala.org/ala/aboutala/offices/oif/ifgroups/cope/Code%20of%20Ethics%202008.pdf> [last check, April 10, 2009]. In Section IV it states explicitly that ALA members "recognize and respect intellectual property rights" recognizing that authors and creators of works have the right to benefit from their creativity.

creators and distributors of intellectual property.”¹¹ In Greece, the Hellenic Association of Librarians and Information Scientists promotes the adoption of a code of conduct for all its members. Among other issues said code of conduct reports that “the librarian ought to make sure that a user has access to information and works available to the public without any restriction that is not necessary in consideration of law or other regulation.”¹²

While libraries’ approaches to copyright and intellectual property can be understood as legal compliance or as an ethical concern in various countries of the world, according to Schachaf and Rubenstein’s survey, that compares institutional policies as they appear on the Web sites of academic libraries in Israel, Russia, and the United States, is an ethical concern. For one reason, the appearance of intellectual property in the code of ethics indicates that the professional association considers it to be an ethical concern.¹³ It is possible that the professional code of ethics addresses copyright and intellectual property concerns only in countries where the law, for one reason or another, is not perceived to be sufficient by the librarians. It is in these countries that the issue becomes an ethical concern. Compliance with copyright laws in a country is likely to reflect the general level of individuals’ and organizations’ (such as academic libraries) compliance with the laws in this country. These differences among the countries may be further explained by political, social, technological, and economic factors.

In the new environment of information and communication systems, the protection of legal rights can hardly be achieved through the application of codes of conduct, only. Ethics and professional principles of librarianship are not enough to cope with conflicting interests and rights such as people’s the right to have access to information and knowledge from one side and the author’s right for communication to the public of his work including the making available to the public of his works in such a way that members of the public may access these works from a place and at a time individually chosen by them. Ethics and professional

-
11. See IFLA/FAIFE, The Library Association Code of Professional Conduct, available at URL: <http://www.ifla.org/faife/ethics/lacode.htm> [last check, April 10, 2009]. The Chartered Institute of Library and Information Professionals (CILIP) has developed a set of ethical principles and a Code of Professional Practice for Library and Information Professionals available at URL: <http://www.cilip.org.uk/policyadvocacy/ethics/code.htm> [last check, April 10, 2009].
 12. See The Hellenic Association of Librarians and Information Professionals at URL: <http://www.eebep.gr> [last check, April 10, 2009].
 13. Schachaf, P. and Rubenstein, E., A Comparative Analysis of Libraries’ approaches to Copyright: Israel, Russia and the U.S. available at URL: <http://dlist.sir.arizona.edu/2117/01/approachesToCopyright.pdf> [last check, April 10, 2009].

principles of librarianship are useful to the point that they can regulate the relations among librarians as well as between them and users of libraries. It is questionable, though, whether these codes of conduct can regulate upon issues such as copyright protection and/or people's right for access and use of copyrighted works. In most societies, the rule of law is not a matter of professional organizations but rather of legislators and the Parliament.¹⁴

The need to overcome the weakness of law as well as of self-regulation regarding the enforcement of rules upon all the stakeholders of the librarian community may lead to the application a hybrid model for regulation, i.e. the model of co-regulation which is a remix of self-regulation and legislation. It's a model that caters for consultation processes among members of the librarian community with the aim to (re)consider new trends in librarianship and produce consensus upon them of all interested members.¹⁵ The State through its legislative power remains the ruling player in this hybrid model. However, this hybrid model emphasises upon the State's role to amend legislation once new trends and status-quo in the market allows for such an amendment.¹⁶ After all, the State's legislative power in librarianship is limited by the social developments which alter the forms and means of libraries' various operations and services as well as the central role that libraries play in saving, archiving, and disseminating knowledge, arts, science, and culture in society.

While most librarians have been engaged in a relentless effort to secure proper protection for copyright holders for works deposited in DL, yet the vagueness in regulation regarding permitted uses of works in the environment of DL inevitably causes tension among right-holders of copyrighted works, librarians, and the end-users of works. A tension that as some scholars suggest is analogous to war, i.e. the copyright war. But what is the role of a librarian in this war? As Pnina Schachaf and Ellen Rubenstein, suggest librarians can take an active role by joining creators and right-holders or users and engage in fighting. They can serve as moderators who promote peace or maintain cease-fire between each side. They can also play more passive roles by observing and reporting the war or by remaining uninvolved. Whatever the role of librarians is, it is clear that they should follow their ethical guidelines and comply with copyright law.¹⁷

14. Strakantouna, V., Piskopani, A-M. and Mitrou, L. (2007). Personal Data and Libraries, Private Law Chronicle (Xronika Idiotikou Dikaïou) (Z), 281-288.

15. Mitrou, L. (2002) *ibid.*, p. 69.

16. Mitrou, L. (2005) Self-Regulation in Cyberspace, Sakkoulas, pp. 22-24.

17. Schachaf, P. and Rubenstein, H. *ibid.*

DL & p2p

The DL of the 21st century is a hybrid form of a library that deviates from the traditional book-keeping library of the past. The term “Digital Library” was coined because of the Internet and refers to an evolved new form of a library that could cover a wide range of information services.¹⁸ The DL of the 21st century is not merely a host of digitized books and collections, but rather it’s an integrator of information management systems, that consists of important elements such as data and metadata, human contribution (creators, users, managers), IT infrastructures (computers, networks, software) which are all orchestrated with the aim to organize, manage, and make available to, i.e. open access to, knowledge and information to library-users.¹⁹ The DL of the 21st century is a borderless organization much like the Internet is a borderless network of networks. Access to DL does not depend any more upon proximity to the local physical premises of the organization. In addition, access to the contents of a DL does not require ownership of an item that becomes available through it. Instead of “owning” the publication, DL are “leasing” it under a license agreement. DL’s focus has turned from the quest how they will digitize materials, store them and make them available to the quest how they will manage the rights along with the materials.²⁰

18. Bokos, G., (2001) Introduction to Information Science, Papassotiriou p. 168.

19. Atkins, D. E. (1997) Report of the Santa Fe Planning Workshop on Distributed Knowledge Work Environments: Digital Libraries, Report Version September 20, 1997, in which it is stated that “the concept of a «digital library» is not merely equivalent to a digitized collection with information management tools. It is rather an environment to bring together collections, services, and people in support of the full life cycle of creation, dissemination, use, and preservation of data, information, and knowledge.” Available at URL: <http://www.si.umich.edu/SantaFe> [last check, April 10, 2009]. See also Griffin, St. M., NSF/DARPA/NASA Digital Libraries Initiative, A Program Manager’s Perspective, available at URL: <http://www.dlib.org/dlib/july98/07griffin.html> [last check, April 10, 2009], and L. Candela, D. Castelli, N. Ferro, Y. Ioannidis, G. Koutrika, C. Meghini, P. Pagano, S. Ross, D. Soergel, M. Agosti, M. Dobrev, V. Katifori, H. Schuldt, The DELOS Digital Library Reference Model available at URL: http://www.delos.info/files/pdf/ReferenceModel/DELOS_DLReferenceModel_0.98.pdf [last check, April 10, 2009], which defines a digital library as: An organization, which might be virtual, that comprehensively collects, manages and preserves for the long term rich digital content, and offers to its user communities specialized functionality on that content, of measurable quality and according to codified policies.

20. Coyle, K. (2004) The rights in the Digital Rights Management, D-Lib Magazine, 10(9) available at: <http://www.dlib.org/dlib/september04/coyle/09coyle.html> [last check, April 10, 2009].

And that is because the architecture per se of a digital library is different than what we've been used to. It is a peer-to-peer (hereinafter, P2P) architecture.²¹ P2P technological networks are of vital importance for the evolution of DL.²² Actually, on the eve of DL, P2P technological networks evolve as technological infrastructure that is an important architectural element for DL's networking with peers and competitive advantage. During the '90s, the Internet consisted mainly of client/server models which are uncomplicated methods to manage and control the distribution of content. During the last years, however, several aspects of IT developments such as the widespread penetration of broadband Internet, more connectivity, mobility, the evolution of compression technology, the demand for more storage capacity, more CPU power, and a large amount of content residing on the personal computers of end-users, have changed the way in which users and prospective DL stakeholders connect to the Internet and make use of the content available online.²³ The combination of IT developments makes it difficult for DL to gain profit via the client/server model. DL infrastructure cannot scale based on the client/server model; but it can if it leverage on the P2P architecture.²⁴

The widespread penetration of the Internet causes content providers to explore new distribution platforms that provide solutions for the disadvantages of the client/server models. DL, publishers, the media industry and end users are exploring systems and platforms to publish and distribute online services and content. P2P have demonstrated the opportunities of this disruptive technology regarding

-
21. Kallinikou, D. Papadopoulos, M. Kaponi, A. & Strakantouna, V. (2009), Alternative system for non-commercial use of intellectual property in consideration of free P2P file-sharing, pp. 3-7, available at URL: <http://www.marinos.com.gr/bbpdf/pdfs/msg77.pdf> [last check, April 10, 2009].
 22. See Ioannidis, Y., Schek, H.-J. and Weikum, G., (2005). Future Digital Libraries Management Systems: System Architecture and Information Access, 8th DELOS Thematic Workshop, Schloss Dagstuhl, Germany, available at URL: http://dbis.cs.unibas.ch/delos_website/D1.1.2%20-%20Workshop%20II%20on%20DL%20Access%20and%20Architecture%20jointly%20with%20WP2%20FUTURE%20Digital%20Library%20Management%20Systems%20System%20Architecture%20and%20Information%20Access.pdf [last check, April 30, 2009].
 23. De Boever, J. Peer-to-Peer Networks as Distribution and Publishing Model, available at URL: http://elpub.scix.net/data/works/att/128_elpub2007.content.pdf [last check, April 30, 2009].
 24. Krishnan, R., Smith, M.D., Tang, Z. and Telang, R. (2006) Digital Business Models for Peer-to-Peer Networks: Analysis and Economic Issues, available at URL: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=917899 [last check, April 30, 2009], Travis, H. (2005) Building Universal Digital Libraries: An Agenda for Copyright Reform, Pepperdine Law Review. 33, 761-829, available at URL: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=860784 [last check, April 30, 2009].

the evolution of DL. P2P systems have often been described as the counterpart of client/server networks.²⁵ In client/server systems, centralized servers manage and control the network, provide services and resources whereas the clients consume these resources. Several client/server networks can hardly meet the demand for resources because of an increasing number of users, higher bandwidth traffic and the arrival of a variety of applications. The major drawbacks of client/server systems in comparison with P2P is that the client/server models suffer from inefficient allocation of resources and limited scalability which can result in bottlenecks and eventually in single points of failure. Furthermore, additional users stand for additional costs as they consume more bandwidth of the system. Nodes in P2P networks do not only act as clients, but they exhibit server functions as well.²⁶ In addition, client/server networks are not scalable and are susceptible to bottlenecks and single points of failure whereas P2P networks are characterized by scalability, decentralization, transient connectivity, cost efficiency, fault tolerance, self organization, sharing of resources and autonomy.²⁷ In theory, P2P systems exhibit positive network externalities in a way that additional users²⁸ add value to P2P networks by introducing extra resources in the system. In this way, users preserve the system and influence the functioning, performance and control of the network by making their resources available. Therefore, it is critical for DL the deployment of a P2P system that is able to cope with the transient presence of nodes, network/computer failures, and be capable to self-organize itself in the absence, more often than not, of centralized coordinating components.

This kind of needs and technological requirements for DL seems that P2P architectures satisfy, thus P2P plays a crucial role for DL. For example, in the Federated Digital Library (FDL) model, which is one of the classic solutions for sharing information among libraries in relevant topics, there is a group of organizations, working together formally or informally, that agree to support a set of common services

25. See Pourebrahimi, B. Bertels, K., Vassiliadis, S. A Survey of Peer-to-Peer Networks, available at URL: http://ce.et.tudelft.nl/publicationfiles/1075_526_prorisc05.pdf [last check, April 30, 2009].

26. This is why nodes or peers have been described as servants (SERVER + cliENTS).

27. See *Androutsellis-Theotokis, St. & Spinellis, D.* (2004). A Survey of Peer-to-Peer Content Distribution Technologies, *ACM Computing Surveys*, 36(4), 335–371. available at URL: <http://www.spinellis.gr/pubs/jml/2004-ACMCS-p2p/html/AS04.pdf> [last check, April 30, 2009], Eberspächer, J. & Schollmeier, R. First and Second Generation of Peer-to-Peer Systems, In *R. Steinmetz and Wehrle K.* (Eds.). *Peer-to-Peer Systems and Applications*, Berlin Heidelberg: Springer-Verlag, 2005, p. 35-56.

28. Such as DL, content providers, end-users, third-party DL-supporting and affiliated organizations, e.tc.

and standards, thus providing interoperability among their members.²⁹ Conventionally, most of FDL were built in the client/server fashion. However, with the sharp growth of an enormous number of DL, especially those moderate-sized ones, there becomes a performance bottleneck problem in FDL based on the client/server model. The solution to this bottleneck problem in FDL appears to be the P2P architecture. Additionally, P2P architecture in DL seems to provide solutions to problems arising not only from scalability, but also from the granularity and meaning in metadata elements used widely in DL. Due to the semantic heterogeneity resulting from the different metadata schemas employed by various DL, the client/server models cannot furnish DL users with a searching application that does not end up as a difficult, complex, and ineffective task across distributed and heterogeneous digital repositories. DL may, also, collaborate with one another to provide content preservation by storing each other's material. Systems such as OceanStore³⁰ and Intermemory³¹ employ this idea.

By attempting to deprive DL from P2P technological architecture and applications on the grounds of considerations for P2P legality, the result will be to impose higher costs than those projected on the deployment DL projects by depriving them of a method of distributing their output efficiently without incurring high costs. File-sharing software, including the P2P applications is capable of cheaply and quickly distributing thousands of public domain literary works such as those made available through Project Gutenberg³² as well as those historic public do-

29. Gonçalves, M. A., France, R.K., Fox, E. A. & Doszkocs, T. E, MARIAN Searching and Querying across Heterogeneous Federated Digital Libraries, available at URL: http://www.ercim.org/publication/ws-proceedings/DelNoe01/11_Fox.pdf [last check, April 30, 2009]. An interesting example of a FDL requiring interoperability is the Networked Digital Library of Theses and Dissertations (NDLTD), at URL: <http://www.ndltd.org>, which is an international federation of universities, libraries, and other supporting institutions interested in worldwide access to electronic theses and dissertations (ETDs).

30. Chen, Y., Katz, R. & Kubiawicz, J. (2000) SCAN: A dynamic, scalable and efficient content distribution network, Computer Science Division, University of California at Berkeley, USA, available via URL: <http://www.springerlink.com/content/wmxcyp86urbmpx> [last check, April 30, 2009].

31. Chen, Y., Edler, J., Goldberg, A., Gottlieb, A., Sobti, S. & Yianilos, P. - N. (1999) A prototype implementation of archival intermemory, available at URL: <http://pnylab.com/pny/papers/improto/improto.pdf> [last check, April 30, 2009].

32. See Project Gutenberg at URL: http://www.gutenberg.org/wiki/Main_Page [last check, April 30, 2009]. Project Gutenberg is the first and largest single collection of free electronic books, or eBooks. Michael Hart, founder of Project Gutenberg, invented eBooks in 1971 and continues to inspire the creation of eBooks and related technologies today.

main films released by the Prelinger Archive.³³ Distributing works such as books, music, and movies over the Internet can be prohibitively expensive for DL as well as non-profit entities such as Project Gutenberg or the Internet Archive, which must divert scarce resources to purchasing bandwidth and data storage, if they are not allowed to leverage upon the P2P architecture and applications, instead of digitizing more books. File-sharing software permits these entities to shift storage and bandwidth costs onto readers and Internet users more generally, and preserve limited budgets for core mission tasks. File-sharing programs let Internet users do much more than substitute MP3 downloads for CD purchases, including locate public domain music, listen to recordings of live performances in which musicians do not claim copyright, rediscover out-of-print or hard-to-find books or music, and sample albums before buying.³⁴ If it works so for Internet users, it does work the same for DL. P2P represents a great technological advantage in information and communication core technological infrastructure for the evolution of DL. None of the great advances in information and communications technology, from the photocopier to the videocassette recorder, personal computer, and Internet, would have been viable had all copyright infringements by their users been imputed to their manufacturers.³⁵ The legal assault on P2P technologies and the “zero tolerance policy” articulated in the Napster and Aimster cases in the U.S., and which has been used henceforth from content providers and intellectual

33. See Prelinger Archive at URL: <http://www.archive.org/details/prelinger> [last check, April 30, 2009]. The Prelinger Archive was founded in 1983 by Rick Prelinger in New York City. Over the next twenty years, it grew into a collection of over 60,000 “ephemeral” (advertising, educational, industrial, and amateur) films. In 2002, the film collection was acquired by the Library of Congress, Motion Picture, Broadcasting and Recorded Sound Division. Prelinger Archive remains in existence, holding approximately 4,000 titles on videotape and a smaller collection of film materials acquired subsequent to the Library of Congress transaction. Its goal remains to collect, preserve, and facilitate access to films of historic significance that haven’t been collected elsewhere. Included are films produced by and for many hundreds of important US corporations, non-profit organizations, trade associations, community and interest groups, and educational institutions. Users of the Prelinger Archive are warmly encouraged to download, use and reproduce these films in whole or in part, in any medium or market throughout the world. They are also warmly encouraged to share, exchange, redistribute, transfer and copy these films, and especially encouraged to do so for free. Any derivative works that they might produce using these films are theirs to perform, publish, reproduce, sell, or distribute in any way they wish without any limitations. Their right to use these films is granted by the Creative Commons Public Domain license.

34. Travis, H., *ibid.*, (2005), p. 824.

35. Travis, H., *ibid.*, (2005), p. 826, and note 425 attributing this argument to Justice David H. Souter and Justice Antonin Scalia, with respect to Xerox photocopier and Apple iPod MP3 player.

property right-holders with the aim to attack every possible P2P application that come across, represents a radical departure from legal principles of civil law, and will unnecessarily deprive Internet users of a variety of noncommercial content and many of the benefits of P2P technology and DL leveraging upon it.³⁶

DL & coryright

DL as repositories of works include in their contents copyrighted works as well as works under no copyright such as works in the public domain. DL could include in their contents works delivered in any medium and format. Frequently, DL are the licensed creators of derivative works, i.e. a variety of digital media and formats of the same work furnished to it. A DL, also, could contain many different copyrights. For example a textual article or e-book are protected as literary works, photographs as artistic works, a music DVD or CD as a musical work. The DL's website could also subject to different copyrights. Components could be protected by the right of communication to the public or qualify for the database right protection.³⁷ The Copyright issues that pertain to the legitimate operation of a DL are complex and usually difficult to resolve once a dispute arose. Intellectual Property has always been in the crossroads of conflicting interests among the creators, right-holders, and the general public, and it has always been a vexing problem to balance among conflicting copyright-related interests.³⁸ This two-sided reality of Copyright is explicit in article 27 of the Universal Declaration of Human Rights. The Constitution of Greece caters for the protection of intellectual property in its provisions, and specifically the provisions of article 2§1, article 5§1, article 14§1, article 16§1, article 17, article 18§5, and article 28§1. Copyright law in Greece, which is outlined by Law 2121/1993³⁹ as amended since 1997 and onwards, provides definitions and protections for intellectual property rights. The copyright has an exclusive and absolute character, but is subject to limits that are determined by the concept of the work or are explicitly prescribed by law as to the term and extent of the right. The scope of copyright

36. Travis, H., *ibid.*, (2005), p. 826, and note 427 regarding Lawrence Lessig's arguments presented in his book *Free Culture—How Big Media Use Technology and the Law to Lock Down Culture and Control Creativity*, The Penguin Press, 2004, available at URL: <http://www.free-culture.cc/freeculture.pdf> [last check, April 30, 2009].

37. Pedley, P. *ibid.*, (2005).

38. Kallinikou, D. (2008). Archives, Libraries and Copyright. Proceedings of the Conference, Archives, Libraries and the Law in the era of Information Society, Athens, February 2-3, 2006,

39. See Hellenic Copyright Organization, Law 2121/1993 available at URL: <http://web.opi.gr/portal/page/portal/opi/info.html/law2121.html> [last check, April 10, 2009].

comprises works as intangible goods, irrespective of the material on which the work is incorporated. The main features of the work are form and originality.⁴⁰ The idea is not protected by copyright, unless it takes up a specific form. In this regard, processes, operating methods and mathematical concepts per se are not protected. Despite occasional controversies and concerns, the distinction between form and idea is a substantial rule for determining the extent of copyright protection. The idea is free and accessible by anyone, constitutes common property and cannot become subject to copyright, unless it has taken up some form. In this respect, it has been held that scientific discoveries or theories are not per se protected by copyright. A scientist cannot acquire the copyright of a theory or discovery because that would bring scientific and technological progress to a halt. The idea belongs to everybody, but the way that an artist's inspiration and emotion are expressed belongs exclusively to his/her personality and these are the features that are protected. Copyright can only protect the form and, in this regard, scientific legal works are subject to protection, both by national legislation and international treaties such as the Berne Convention⁴¹ article 2§1, as well as the TRIPS Agreement on trade-related aspects of intellectual property rights,⁴² and the WIPO Copyright Treaty⁴³ and the WIPO Performances and Phonograms Treaty.⁴⁴ According to the article 5§2 of the Berne Convention the employment and the exercise of the copyright shall not be subject to any formality.

On the basis of the above points, any scientist or researcher may write a book on civil or administrative law, interpret a law of any content or comment on a court order. It should also be noted that protection does not extend to official texts that express the exercise of state power, especially legislative, administrative or judicial texts, not to the expressions of folk tradition, news and simple events or facts, unless any of them can be included in the category of compilations or derivative works. Specifically, protection does not extend to legislative, administrative or judicial texts because their purpose is to become broadly known for the sake of public interest. However, compilations of laws, decrees, court orders

40. For the concept of work, see (in Greek) Koumantos, G. (2002), Copyright, 8th edition, Ant. Sakkoulas pp. 105 et seq., Kotsiris, L., (2005) Copyright Law, 4th edition, Thessaloniki: Sakkoulas, pp. 53 et seq., Kallinikou, D. (2008) Copyright and Related Rights, 3rd edition, P. Sakkoulas, pp. 29 et seq., Marinos, M.-T. (2005) Copyright, 2nd edition, Ant. Sakkoulas, pp. 71 et seq.

41. Universal Copyright Convention as it was revised in Paris on July 24, 1971, adopted in Greece through Law 100/1975.

42. Adopted through Law 2290/1995.

43. Adopted through Law 3184/2003.

44. Adopted in Geneva on December 20, 1996; adopted through Law 3183/2003.

e.tc. are protected under article 2§2 of Law 2121/1993. Simple events or facts, such as legal literature information, are not protected per se for lack of the element of originality. If, however, they take up the form of a compilation, they may be protected as collective works, provided that the selection or layout of their content is original (e.g. the literature list of a specific field of the legal science), without ruling out the possibility of extending protection to database creators by special rights. Apart from the conceptual limitations of copyright, the law imposes restrictions on the length of the right. According to the community *acquis*, the length of protection under national law is determined on the basis of the lifetime of the author and seventy years after his/her death, calculated from 1st January of the year after the author's death.⁴⁵ Special provisions of Law 2121/1993 cater for the duration of copyright in the case of works of joint authorship,⁴⁶ anonymous or pseudonymous works,⁴⁷ works published in volumes, parts, installments, issues or episodes,⁴⁸ audiovisual works,⁴⁹ and previously unpublished works.⁵⁰ This length applies to both the moral and the property right. On expiry of this period, the work falls into the public domain and may be freely exploited, subject to the exercise by the State represented by the Minister of Culture of the moral right, and specifically the powers to recognize the paternity and safeguard the integrity of the work.⁵¹ On the basis of the copyright law, public domain is relevant only to works whose period of protection has expired. Such works become common property and may be freely exploited.

This time restriction is mainly justified for reasons of protecting society. The recognition of copyright ensures that the author may take financial advantage of his/her work, thus boosting the cultural output of each country and the humanity as a whole. At some point of time, however, the work should be freely disseminated, so that it may become the property of all. It should be underlined that, for works legitimately published or presented to the public for the first time after the expiry of copyright protection, a related right is prescribed in Law,⁵² similar to the author's property right, whose validity expires twenty-five years after the first publication or presentation to the public. The related right of publishers,

45. Article 29§1 of Law 2121/1993.

46. Article 30 of Law 2121/1993.

47. Article 31§1 of Law 2121/1993.

48. Article 31§2 of Law 2121/1993.

49. Article 31§3 of Law 2121/1993.

50. Article 51A of Law 2121/1993.

51. Article 29§2 of Law 2121/1993.

52. See article 51A of Law 2121/1993, see note 58.

mainly as regards typesetting and pagination of published works, is valid for fifty years after the last edition of the work.⁵³ Therefore, it is necessary to examine the origin of the work and the rules applicable each time, since some times the length of the moral right is unlimited, such as in French law, while some works enjoy “perpetual” protection. Works with expired protection may become the subject of any use in the framework of operation of DL without the permission of the right-holder, unless other provisions apply, such as laws on the protection of cultural heritage. Material that has fallen into the public domain may be freely digitized and preserved in the framework of DL and repositories.

The definition of copyright in Law 2121/1993 is similar to the definition posed by international conventions and mandates that one cannot produce, copy, communicate, or transmit to the public copyrighted material such as literary, dramatic, musical, and artistic works, films, and sound recordings without the permission of the copyright owner and/or right-holder. There are exceptions to this broad restriction which are mainly described in Section IV of Law 2121/1993 (articles 18 et al. titled “Limitations on the Economic Right”) including reproduction for private use,⁵⁴ reproduction for textbooks and anthologies,⁵⁵ reproduction for teaching purposes,⁵⁶ reproduction for libraries and archives,⁵⁷ among other provisions of section IV of Law 2121/1993, as well as clauses for exception from the reproduction right⁵⁸ and the three-step test.⁵⁹ Said limitations are set in law for reasons of social policy, aim at the protection of the interests of the public, are close in number, are applied as an exception to the rule of copyright protection, and should not have a broad interpretation.⁶⁰ The clause for the three-step-test in Law 2121/1993 which is in sync with article 9§2 of the Berne Convention for the protection of literary and artistic works as well as article 13 of the TRIPS Agreement on trade-related aspects of intellectual property rights, and articles 10 of the WIPO Copyright Treaty and 16 of the WIPO Performances and Phonograms Treaty imposes limits in the meaning of private use which are subject to revision

53. Article 51 of Law 2121/1993.

54. Article 18 of Law 2121/1993.

55. Article 20 of Law 2121/1993.

56. Article 21 of Law 2121/1993.

57. Article 22 of Law 2121/1993.

58. Article 28B of Law 2121/1993.

59. Article 28C of Law 2121/1993.

60. Marinos, M-T. (1994) The violation of Intellectual Property Right and of Related Rights, Hellenic Justice Magazine (EllDik) 35, p. 1441.

in consideration of technological, societal as well as legal developments in society.⁶¹

The three-step-test in the Greek Copyright Law helps to ensure authors that their works do not get corrupted either accidentally or maliciously. It also allows publishers to develop products without fear that their markets will be destroyed by copies from other sources.⁶² However, the three-step-test should not be interpreted in a manner that jeopardizes an adequate solution which balances the usually conflicting interests of the author, the subsequent right-holders and people's interest in making use of a work in Cyberspace. The public interest is not well served if Copyright Law neglects the more general interests of individuals and groups in society whereas it solely caters for the interests—financial interests most commonly—of the right-holders. The three-step-test is a means to balance the conflicting interests of the author, subsequent right-holders and the general public. It is a means to achieve equilibrium rather than imbalance in favour of either one of said involved parties. The three-step-test should be interpreted in a manner that respects the legitimate interests of third parties including interests deriving from human rights and fundamental freedoms, interests in competition notably on secondary markets, and other public interests notably in scientific progress and cultural, social or economic development.⁶³

61. See article 28C of Law 2121/1993, see note 67. See also article 5§5 of Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonization of certain aspects of copyright and related rights in the information society: The exceptions and limitations provided for in paragraphs 1, 2, 3 and 4 [of article 5 of said Directive] shall only be applied in certain special cases which do not conflict with a normal exploitation of the work or other subject-matter and do not unreasonably prejudice the legitimate interests of the right-holder. Also, see article 9§2 of the Berne Convention for the protection of literary and artistic works article 13 of the TRIPS Agreement on trade-related aspects of intellectual property rights, and articles 10 the WIPO Copyright Treaty and 16 of the WIPO Performances and Phonograms Treaty.

62. Arms, W. (2001), *Digital Libraries*, The MIT Press, p. 117.

63. See article 7 of Law 2290/1995 which transposed into the Greek legal framework for Copyright protection The TRIPS Agreement, and according to which The protection and enforcement of intellectual property rights should contribute to the promotion of technological innovation and to the transfer and dissemination of technology, to the mutual advantage of producers and users of technological knowledge and in a manner conducive to social and economic welfare, and to a balance of rights and obligations. See also the WIPO Copyright Treaty available at URL: http://www.wipo.int/treaties/en/ip/wct/trtdocs_wo033.html [last check, April 30, 2009] the preamble to which emphasizes the need to maintain a balance between the rights of authors and the larger public interest, particularly education, research and access to information.

Digital (copy)rights

Intellectual Property in our legal system, which is belongs to the Continental System, is addressed by two, distinctive and absolute rights, i.e. the moral right and the economic right. Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonization of certain aspects of copyright and related rights in the information society has introduced the so called “Digital Rights” in Copyright, namely the right for reproduction of copyrighted material, the right for distribution of copyrighted works, the right for communication of the work to the public in addition to the other known economic rights of copyright with which the creator and/or right-holder is empowered to permit or forbid the use of his/her work. These “Digital Rights” of the Copyright are of vital importance for the operation of DL. They are also cause of friction between DL and Collecting Societies.

More specifically:

Reproduction of copyrighted material

A user and/or a DL are not allowed to reproduce or communicate copyrighted material unless permission is granted by the intellectual property right-holder. Unauthorized reproduction of such material leads to civil liability in the form of damages and criminal responsibility remedied through fines or potential imprisonment. Permission for exercising the right for reproduction of copyrighted material must be granted in writing otherwise it is null and void.⁶⁴ In case of granted permission which is not in print, nullity may be invoked only by the author.⁶⁵

One of the most important legal issues related to copyright is the need of DL to make copies for preservation purposes and/or for future use. It is more than common in the non-print environment, where either the fragility of the infrastructure per se or the obsolescence of the equipment to direct a depository institution to reproduce the material. Given that the copyright law gives the author the exclusive right to authorize the reproduction and dissemination of his work, such activities can be performed legally only by the copyright right-holder.⁶⁶ Although the copyright legislation in many countries, Greece among others, provides the right of libraries to reproduce protected works, that have been lost or damaged and are no longer available in the market, no legislation provides permission to make backup copies of all kinds of material deposited in a library. Article 22 of

64. Article 14 of Law 2121/1993.

65. Kallinikou, D. *ibid.*, (2008), pp. 204-205.

66. Kallinikou, D. (2007), Copyright and Libraries, pp. 66-76, 162-167.

Law 2121/1993 provides an exception to the rule of forbidden reproduction of a work without the written permission granted by the copyright holder,⁶⁷ for the sake of the interests of the general public.⁶⁸ In consideration of the provisions of article 22, reproduction of a copyrighted work is allowed if a) it is made by a non-profit library or archiving organization, b) the work belongs to a copy in the library or archive's permanent collection, c) the reproduction aims at retaining that additional copy or at transferring it to another non-profit library or archive, and d) the reproduction is deemed necessary since it is not possible for the library or archive to obtain an additional copy from the market promptly and on reasonable terms.⁶⁹ Under these exceptional circumstances as they are described in law, reproduction of a copyrighted work is permissible by a DL. The aforementioned circumstances are applicable not only in the case of works of authorship, but also in the case of audiovisual works since the exception to the rule introduced by article 22 applies in the case of related rights *mutatis mutandis*.⁷⁰ In consideration of the provision of article 22 of Law 2121/1993, it goes without saying that said article is a roadblock in the way of the evolution of traditional libraries into DL. Therefore, it is necessary to amend the copyright legislation in order to encompass provisions allowing reproduction of analogue and digital material for preservation and/or for legal deposit purposes.⁷¹

The transposition of Directive 2001/29/EC into the Greek Copyright Law that implemented through Law 3057/2002 which amended Law 2121/1993 did not elaborated upon libraries and archives' right for reproduction of a copyrighted work, but rather it left said issue to be judged through the provisions of article 18 of Law 2121/1993 and the meaning of the three-step-test which is subject to interpretation by the hearing judge of a case submitted to court.⁷² According to many librarians, the legislator's option not to elaborate upon libraries right for reproduction of a copyrighted work in the process of the transposition of Directive 2001/21/EC was a wrong choice and it was severely criticised by the librarian community in Greece which considers that libraries—at least public libraries

67. Article 22 of Law 2121/1993.

68. Kotsiris, L. *ibid.*, 2005, pp. 224-225.

69. Koumantos, G. *ibid.*, 2002, p. 294.

70. For cinematographic works of special artistic value see the provisions of article 23 of Law 2121/1993 in consideration of Law 1567/1986.

71. Strakantouna, V. (2007) Legal deposit of works protected by Copyright, addressed at international conference titled "Rethinking the boundaries of copyright," Istanbul 15-16 November, 2007.

72. It was deemed that the provisions of article 5§2(c) of Directive 2001/29/EC are covered in Law 2121/1993 by article 18.

and non-profit archiving organizations—should have been vested with the exceptional right to reproduce copyrighted works, and thus become easier for them to achieve their statutory goals in respect of Copyright Law.⁷³

And that severe criticism for legislator's option to omit elaborating upon libraries and archives' right for reproduction of the works was an outburst of libraries and archives' clamour against Collecting Organizations' practices and pressure for collecting the arbitrarily charged at high rates equitable remuneration for photocopies of works made through public libraries and in accordance with the provisions of article 18§§3-10 of Law 2121/1993.

Librarians and many others, too, seem to believe that rather than investing money and resources in developing useless and proprietary electronic platforms for the online distribution of works, subsequent right-holders could have opted for the examination and application of the best model for legal reproduction of works through public libraries. There are many European countries which have permitted said reproduction of works. In these European countries wherein public libraries' right for reproduction of works does not conflict with copyright the cost of licensing said reproduction of works is covered either from consumers, e.g. the UK-model for reproduction of works through public libraries, or from the State's Annual Financing Plan, e.g. the Norwegian model for reproduction of works through public libraries. The latter model could have been applied in Greece, too. In addition, subsequent right-holders could have leverage upon the examples of Collecting Societies in French, The Netherlands, or Denmark which have opted for solutions that consider technological evolution, social trends, and the need for legal amendment to adapt to reality, nowadays.

In June 2008, the French music group *Petit Homme* signed a special contract with SACEM, the French collecting society for music composers, agreeing that musician of *Petit Homme* could post their work online by excluding Internet protocol, wireless application protocol, and similar protocols from their contract. This agreement excludes SACEM of the group's Internet rights and allows the group the control of their Internet rights while SACEM would handle the remaining rights related to the work of the group.⁷⁴

73. See Hellenic Association of Librarians and Information Scientists, 4th and 5th Declarations of the Hellenic Association of Librarians and Information Scientists, September 24-26, 2008, available (in Greek only) at URL: <http://library.aul.gr/files/pdf/pshfismata%20teliko.pdf> [last check, April 30, 2009].

74. See Saez, C. (2008) *Improbable Match: Open Licences And Collecting Societies In Europe*, available at URL: <http://www.ip-watch.org/weblog/2008/10/28/french-deal-highlights-open-licensing-and-collecting-societies-in-europe> [last check, April 30, 2009].

In August, 2007, Dutch collecting societies BUMA and STEMRA and Creative Commons Netherlands launched a pilot project that seeks to provide Dutch musicians with more opportunities to promote their own repertoire.⁷⁵ This project enables members of BUMA/STEMRA to use the three non-commercial Creative Commons licenses for non-commercial distribution of their works. It also allows Dutch composers and lyricists who already use the Creative Commons non-commercial license to join BUMA/STEMRA and have them collect their royalties for commercial use of their works. The Netherlands is the first country to bring such collaboration between a music copyright organization and Creative Commons, a move applauded by Lawrence Lessig, the founder and chairman of Creative Commons International, as “the first step towards more freedom of choice in the field of exploiting music works in the digital world.”⁷⁶

In January 2008, Creative Commons Denmark and KODA, the Danish Authors’ Society, reached an agreement in which KODA accepted to offer non-commercial Creative Commons licensing to its members. This agreement allows creators to rely on the strength of collective rights management for commercial uses of their works, while taking non-commercial online distribution into their own hands by using Creative Commons licenses.” KODA’s adoption of Creative Commons licensing marks a breakthrough for Danish composers and lyricists wanting to explore new ways of making their work available online while at the same time collecting commercial royalties through KODA. Members must sign an agreement with the KODA in which they indicate which works they wish to license, and for the purpose of this arrangement, only Creative Commons licenses with the “non-commercial” condition can be used.⁷⁷

The Right to Distribute & the Rental and Lending Right

The right to distribute is a sine-qua-non service of the operation of libraries, moreover of DL. The rental and lending right is also understood as a necessary service in the operation of DL. Yet, despite the fact that in many European coun-

75. See Creative Commons Netherlands, Buma/Stemra and Creative Commons Netherlands launch a pilot—More opportunities for music authors to promote their own music, Press Release, Amsterdam, August 23, 2007, available at URL: http://www.creativecommons.nl/bumapilot/070823persbericht_en_web.pdf [last check, April 30, 2009].

76. See Reeder, M. (2007) Dutch Collecting Societies welcome CC, August 23, 2007, available at URL: <http://creativecommons.org/weblog/entry/7622> [last check, April 30, 2009].

77. See Thorne, M. (2008) Danish Collecting Society KODA teams up with CC Denmark, January 31st, 2008, available at URL: <http://creativecommons.org/weblog/entry/8012> [last check, April 30, 2009].

tries⁷⁸ there have been provisions for the rental and lending right as of the '50s, the Green Paper of 1998 on certain aspects of copyright and related rights in the information society⁷⁹ did not cater for the harmonization of legislation in the EC market of the rental and lending right. It was not until Directives 2001/29/EC and 100/1992/EEC, the latter as it was codified with Directive 2006/115/EC, that the need for harmonization of legislation in the EC market was stressed. The Greek legislation does not make full use of Directive 100/1992/EEC for the public lending right. This option was criticised.⁸⁰ Retrospectively thinking, though, and in consideration of the European Court of Justice (ECJ), the option of the Greek legislator not to make full use of the provisions for the public lending right might not have been so problematic at least in as much as it have been the relevant provisions of other member-States which have tried to do so rather unsuccessfully⁸¹.

78. Denmark was the first country to establish a Public Lending Right system in 1946, followed by Norway in 1947 and Sweden in 1954. For Denmark's Public Lending Right see URL: <http://www.plrinternational.com/established/plradministrators/denmark.htm> [last check, April 30, 2009], for Norway's Public Lending Right see URL: <http://www.plrinternational.com/established/plradministrators/norway.htm> [last check, April 30, 2009], and for Sweden's Public Lending Right see URL: <http://www.plrinternational.com/established/plradministrators/sweden.htm> [last check, April 30, 2009]. See also The Public Lending Right International Network, the list of Public Lending Right by country titled Established PLR Schemes, available at URL: <http://www.plrinternational.com/established/Established%20PLR%20Schemes.pdf> [last check, April 30, 2009].

79. Opinion of the Economic and Social Committee on the "Proposal for a European Parliament and Council Directive on the harmonization of certain aspects of copyright and related rights in the information society," Official Journal C 407, 28/12/1998 p. 0030, available at URL: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:51998AC1122:EN:HTML> [last check, April 30, 2009].

80. Papazoglou, V. (2008) Horizontal Action of Academic Libraries: Legal Issues. Proceedings of the Conference, Archives, Libraries and the Law in the era of Information Society, Athens, February 2-3, 2006, p. 133, and Marinos, M.-T. (1998), Some notes upon the status of traditional libraries and public digital libraries under the system of Law 2121/1993, Hellenic Justice Magazine (EllDik) p.1484.

81. ECJ, Oct. 26, 2006, JUDGMENT Failure of a Member State to fulfill obligations -Directive 92/100/EEC - Copyright - Rental and lending right - Failure to transpose within the prescribed period, Commission v. Spain, C-36/2005, Collection 2006, p.I-10313, ECJ, Oct.26, 2006, JUDGMENT Failure to fulfill obligations - Directive 92/100/EEC - Rights related to copyright in the field of intellectual property - Public lending right - Failure to transpose within the period prescribed, Commission v. Italy, C-198/2005, Collection 2006, p.I-107, ECJ, Jul.6, 2006, JUDGMENT (EC) Failure of a Member State to fulfill obligations - Directive 92/100/EEC - Copyright - Rental and lending right - Failure to transpose within the prescribed period, Commission v. Portugal, C-53/2005, Collection 2006, p.I-06215, ECJ, Oct.16, 2003, JUDGMENT (EC) Directive 92/100/EEC Copyright. Remuneration of

The Right to communicate the work to the public

According to national and international law, the copyright owner has also the exclusive right to communicate his work to the public. To ensure unobtrusive access to collected cultural material, provisions should be enacted in law, which will allow depository institutions to enable unobtrusive access to works by giving them the rights of public lending online,⁸² digital disposition,⁸³ and creation of multiple copies,⁸⁴ in a way that caters for both the author's copyright as well as a user's right to access information and cultural works. Though access to legal deposits of collections should be free, it should not be unconditional and/or vague in the permitted uses. Traditionally, one of the main reasons for the existence of public libraries is the provision of access to their collections to the public without any financial demand on the part of the author and/or publisher. That is to say, to make the works of culture, arts, and sciences hosted in copies in public libraries available to the public in its quest to access cultural, artistic, and scientific resources.

Copyright & technology

Technology was always a big opportunity as well as a danger for copyright. Strengthen the law while holding technology constant and the right is stronger. Proliferate copying technology while holding the law constant and the right is weaker.⁸⁵ While in the analogue world, life sans copyright law is possible, in the digital world life that does not subject to copyright law is not possible. In the digital world, every single act triggers the law of copyright. The emergence of digital technologies has radically increased the domain of copyright law from regulating a small portion of human life to regulating absolutely every bit of life lived through a computer.

authors in the event of public lending of their literary or artistic works, *Commission v. Belgium*, C-433/2002, Collection 2003, p. I-12191.

82. *Kallinikou, D. ibid.*, (2007), pp. 76-88, & the same, *ibid.*, 2008, pp. 157-160.

83. *Kallinikou, D. ibid.*, (2007), pp. 89-93, & the same, *ibid.*, 2008, pp. 161-185.

84. *Kallinikou, D. ibid.*, (2008), pp. 142-154.

85. *Lessig, L.* (1999) *Code and Other Laws of Cyberspace*, Basic Books pp.124-127, *Lessig, L.* (2008) *Remix—Making art and commerce thrive in the Hybrid Economy*, The Penguin Press pp.96-97, 276-277, 289-291, *Litman, J.* (2006) *Digital Copyright*, Prometheus Books, pp.35-69, *Benkler, Y.* (2006) *The Wealth of Networks—How Social Production Transforms Markets and Freedom*, Yale University Press, pp. 273-300, *Paul Goldstein, P.* (2007) *Intellectual Property: The Tough New Realities That Could Make or Break Your Business*, Portfolio pp. 24, 27-29.

In consideration of the conflicting relationship between Copyright and Technology one wonders upon the essence of the threat that Copyright is met with in the era of digital information technologies and digital libraries which we all have been living in during the last two decades, at least. Is Copyright at stake because of the nature of its conflict to Technology? The answer is definitely No! We are not entering a time when copyright is more threatened than it is in real space. We are instead entering a time when copyright is more effectively protected than at any time since Gutenberg.⁸⁶ The power to regulate access to and use of copyrighted material is about to be perfected.

An important thing about Copyright law is that, though designed in part to protect authors, the control it was designed to create was never to be perfect. Copyright protection has never accorded the copyright owner complete control over all possible uses of his work. Almost since the inception of copyright regulation, there have always been limitations to copyright.⁸⁷ Perfect control is not the control that law has given owners of intellectual property. Historically, the law of Copyright has been focused mainly on commercial life, i.e. it has laid down the rules according to which for profit exploitation of intellectual property is permissible and doable. Most exceptions to the rules for commercial exploitation of intellectual property, namely the restrictions to intellectual property regulation are triggered by the idea of Copyright's commercial use. Most of these restrictions make provisions for permissible non-commercial use of intellectual property in the sense that all other non-commercial uses of intellectual property aside from these provided and allowed by law, are not permissible without the prior consent of the intellectual property right-holder.

This stance of intellectual property law which is a pure depiction of the traditional intellectual property regulation conceived to fit the analogue world seems that it does not fit in the digital world. It does not fit the user-generated creativity that digital technologies have empowered through the Internet. To the extent that people's creativity finds its expression on the Net, it is inevitably subject to the regulation of Copyright law. To the extent peoples' creativity is based upon that of others, peoples' creativity is in the need of the prior permission of others. To the extent it builds upon the creativity of others, it needs to be sure that this creativity can be built upon legally.

We have learnt through the many, lawsuits over the distribution of peer-to-peer (P2P) file-sharing software for .mp3 formatted music that while technology can

86. Lessig, L., *ibid.*, (1999) p. 127.

87. Lessig, L. *ibid.*, (1999) pp. 130-135, Kallinikou, D. *ibid.*, (2008) pp. 238-278, Sinodinou, T.-E., (2008) Intellectual Property & New Technologies, Sakkoulas, pp. 138-154.

provide enormous scope for access, unless the law supports such access, it will be unauthorized and could lead to legal liability. The future of the learning process lies in seamless access to educational resources available through (digital) libraries and most of them accessible through Information & Communication Technologies (ICTs) and Internet Protocol (IP) networks. In consideration of the situation nowadays, we understand that while digital libraries deploying ICTs and IP networks inherently produce and communicate copyrighted material in their normal process of operation, they activate the potential for copyright infringement. Therefore, we need to encourage the existing copyright legal framework including laws regulating the operation of digital libraries and legal deposit of works to accommodate the disruptive energies of ICTs & IP networks in a way that promotes openness and open access to educational resources.

Copyright & openness (open access)

The term Openness (Open Access) was coined to typify the open access to information or material resources needed for projects; openness to contributions from a diverse range of users, producers, contributors, flat hierarchies, and a fluid organisational structure. In the context of Budapest Open Access Initiative,⁸⁸ Open Access means the free availability of literature and works of authorship, audiovisual works etc on the public Internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the Internet itself. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited. The Bethesda Statement on Open Access⁸⁹ and the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities⁹⁰ seem to agree that for a work to be considered for Open Access, the copyright holder must consent in advance to let users copy, use, distribute, transmit and display the work publicly and to make and distribute derivative works, in any digital medium for any responsible purpose, subject to proper attribution of authorship. With Open Access individuals can take

88. See The Budapest Open Access Initiative at URL: <http://www.soros.org/openaccess/read.shtml> [last check, April 30, 2009].

89. See the Bethesda Statement on Open Access at URL: <http://www.earlham.edu/~peters/fos/bethesda.htm> [last check, April 30, 2009].

90. See the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities at URL: <http://oa.mpg.de/openaccess-berlin/berlindeclaration.html> [last check, April 30, 2009].

projects in their own direction without necessarily hindering the progress of others. Openness is being put forward to facilitate the growth of the open source and free software programming communities, and may involve the consumption and production of free content.⁹¹ The appeal of Openness has become so great that it is sometimes difficult to recognize that limits on Openness are not only necessary but desirable. The virtues of an open environment are undeniable; what is more difficult is negotiating the proper levels of Openness for a given realm of online life.⁹²

The sense for movement of Openness was first understood according to Professor Yochai Benkler, at a conference at Yale University that Professor James Boyle⁹³ organized in April 1999, which was already planned as a movement-building event. That conference, "Private Censorship/Perfect Choice,"⁹⁴ looked at the threats to free speech on the Web and how the public might resist. It took inspiration from John Perry Barlow's 1996 manifesto "A Declaration of the Independence of Cyberspace."⁹⁵ The stirrings of a movement were evident in May 2000, when Yochai Benkler convened a small conference of influential intellectual property scholars at New York University Law School on "A Free Information Ecology in the Digital Environment."⁹⁶ This was followed in November 2001 by a large gathering at Duke Law School, the "Conference on the Public Domain," the first major conference ever held on the public domain.⁹⁷ It attracted several hundred people and permanently rescued the public domain from the netherworld of

91. See Wikipedia, Openness, available at URL: <http://en.wikipedia.org/wiki/Openness> [last check, April 30, 2009].

92. *Bollier D.*, (2008) *Viral Spiral: How the Commoners Built a Digital Republic of their Own*, The New York Press pp.40, available at URL: <http://www.viralspiral.cc/download-book> [last check, April 30, 2009].

93. See Boyle, J. A Politics of Intellectual Property: Environmentalism For the Net? available at URL: <http://www.law.duke.edu/boylesite/Intprop.htm> [last check, April 30, 2009], was an influential piece that James Boyle wrote in 1997, calling for the equivalent of an environmental movement to protect the openness and freedom of the Internet.

94. See *Yale Bulletin & Calendar*, Private Censorship and Perfect Choice Conference to explore Speech and Regulation on the Net, April 5-12, 1999 Volume 27, Number 27 available at URL: <http://www.yale.edu/opa/arc-ycb/v27.n27/story3.html> [last check, April 30, 2009].

95. See Barlow, J-P. A Declaration of the Independence of Cyberspace, available at URL: <http://homes.eff.org/~barlow/Declaration-Final.html> [last check, April 30, 2009].

96. See The Information Law Institute at New York University School of Law supported by Arthur S. & Marilyn Penn Foundation, Conference A Free Information Ecology in the Digital Environment, available at URL: <http://www1.law.nyu.edu/ili/conferences/freeinfo2000/aboutconf/index.html> [last check, April 30, 2009].

97. See Duke Law School supported by the Center for the Public Domain, Conference on the Public Domain, November 9—11, 2001, available at URL: <http://www.law.duke.edu/pd> [last check, April 30, 2009].

“non-property.” People from diverse corners of legal scholarship, activism, journalism, and philanthropy found each other and began to re-envision their work in a larger, shared framework.⁹⁸

The Open Access movement cropped up as a reaction of academia in the increasingly rising pricing of scientific publications and subscriptions controlled by publishers and distributors that intervene in the process of scientific knowledge dissemination and stifle competition in scientific publishing and distribution.⁹⁹ By the time¹⁰⁰ Open Access started to be a central point of discussion in the agenda of academic institutions, prices had risen many times faster than inflation since 1986.¹⁰¹ Fortuitously, just as journal prices were becoming unbearable, the Internet emerged to offer an alternative. The Internet has played a catalytic role in the evolution of the Open Access movement because of the radical changes it has imposed in the process of authoring, publishing, distributing, and pricing content via the Internet networked public sphere. The evolution of the Web into Web 2.0 has enabled more interaction and participation among users and empowered them to undertake action both as readers and authors, publishers and distributors, in the process of production and consumption of knowledge. Since the beginning of the Internet era, openness of scientific knowledge, art, and culture has been fostered and cultivated in way that indicates that openness is somewhat intrinsically connected to the hierarchical anarchy of the Net. While Open Access was born because of the need to remove price barriers (subscriptions, licensing fees, pay-per-view fees), it was soon realized that its survivability was subject to the need to remove permission barriers as well (most copyright and licensing restrictions).

98. *Bollier D.*, *ibid.*, (2008), p. 67.

99. See *Lessig, L.* Answers to Written Questions. The Senate Judiciary Committee, “The Microsoft Settlement: A Look to the Future”, available at URL: <http://www.lessig.org/content/testimony/answers.doc> [last check, April 30, 2009].

100. *Suber P.*, Timeline of the Open Access Movement, revised February 9, 2009, available at URL: <http://www.earlham.edu/~peters/fos/timeline.htm> [last check, April 30, 2009].

101. See *Peter Suber*, Open Access Overview, Focusing on open access to peer-reviewed research articles and their preprints, revised June 19, 2007, available at URL: <http://www.earlham.edu/~peters/fos/overview.htm> [last check, April 30, 2009]. See also Kyrillidou, M. and Young, M, ARL Statistics 2001-2002, Association of Research Libraries, available at URL: <http://www.arl.org/bm~doc/arlstat02.pdf> [last check, April 30, 2009], the same, ARL Statistics 2002-03, Association of Research Libraries, available at URL: <http://www.arl.org/bm~doc/arlstat03.pdf> [last check, April 30, 2009], the same, ARL Statistics 2004-05, Association of Research Libraries, available at URL: <http://www.arl.org/bm~doc/arlstat05.pdf> [last check, April 30, 2009].

The Internet and the Open Access movement has inevitably altered the way in which librarians see their own profession, and their role in saving, archiving, and distributing knowledge, art, and culture to the general public. Because of the nature of their profession, librarians, even those working for for-profit organizations and/or private libraries and archives, are prone to adopt solutions that remove both price and permission barriers in order to keep the knowledge commons open to the general public. After all, the general public is their most important clientele. The general public is the most important stakeholder of any public and/or private library. Librarians, and especially the younger generations of them who have had the opportunity through their undergraduate and graduate studies in academic institutions to get a grasp of what a DL might be and/or could evolve to become, usually strive to develop strategies favouring the outcomes best corresponding to the deepest values of their profession, in particular the desire to overcome barriers in the open access of knowledge, art, and culture saved, archived, and distributed through libraries. From that perspective, it is clear and reasonable that librarians throw all of their weight to the Open Access movement for the following reasons:¹⁰²

1. It is the only alternative to present publishing that has a chance to develop without the economic penalties associated with present, digital publications peddled in the form of site licenses.
2. It is the only alternative that, although relying on some external, public support, has a chance to withstand the competition of the large publishers over the middle and long term, unlike most learned societies and similar, generally irreproachable, institutions.
3. It is the only way for librarians to recover responsibility over traditional concerns such as classification and conservation. In this manner, they can also get involved with the elaboration of various tools that add values to any collection of scientific articles (what Professor Jean-Claude Guéron calls “epistemological engineering”).
4. It is the only way to ensure that powerful panoptic effects, either already identified or to be discovered, do not remain the exclusive preserve of private, unaccountable, profit-driven companies, many of them operating offshore.

102. See Guéron, J-C. (2001), In Oldenburg's Long Shadow: Librarians, Research Scientists, Publishers, and the Control of Scientific Publishing, Association of Research Libraries 2001, available at URL: <http://www.arl.org/resources/pubs/mmproceedings/138gueron.shtml> [last check, April 30, 2009].

5. Open libraries and archives provide a very good way to develop new and positive relationships with scientists, particularly gatekeepers, and administrators to review in depth the processes of scientists' evaluation now that these questions can be treated independently of print-related constraints.

A closer look at Openness in Greece reveals that though there is a diverse set of sources of openness, yet very few of these are legal. Mostly, they are based on technological and social practices, including resistance to legal and regulatory drives toward enclosure. For the most part, the drive for openness is based on individual and voluntary cooperative action, not law. The social practices of openness take on a quasi-normative face when practiced in standard-setting bodies.¹⁰³ While very few of them exist, currently in Greece, some of them are very characteristic and a source of hope for the sincere consideration and support of Openness in Greece in the foresighted future. The National Hellenic Research Foundation¹⁰⁴ in Greece, though not the first scientific institution to consider Openness, yet it's probably the most pronounced proponent of Openness and its meaning in academic development and innovation in the country. There are also a few ex-

103. *Benkler, Y. ibid.*, (2006) p. 394 reports a couple of pronounced examples in the US, such as Internet Engineering Task Force (IETF) or the World Wide Web Consortium (W3C).

104. See the National Hellenic Research Foundation (NHRF) in Greece at URL: <http://www.eie.gr/index-en.html> [last check, April 30, 2009]. Currently NHRF issues a number of academic peer-reviewed open access journals such as BYZANTINA SYMMEIKTA through NHRF's Institute for Byzantine Research; see BYZANTINA SYMMEIKTA at URL: <http://www.byzsym.org/index.php/bz> [last check, April 30, 2009], and the HISTORICAL REVIEW, a peer-reviewed open access journal issued by NHRF's Institute for Neohellenic Research; see HISTORICAL REVIEW at URL: <http://www.historicalreview.org/index.php/historicalReview> [last check, April 30, 2009]. Also, NHRF's Open Access—Knowledge for All site at URL: http://www.openaccess.gr/?language_id=1 [last check, April 30, 2009] which is an online platform providing updated and comprehensive information on Open access issues and latest trends in Greece and elsewhere, Open access infrastructure currently being developed by the National Documentation Centre (EKT) which is part of NHRF. It also links to NHRF's Open Access libraries and repositories such as HELIOS repository at URL: <http://helios-eie.ekt.gr/EIE> [last check, April 30, 2009] still operating in a Beta version, PANDEKTIS digital thesaurus of primary sources for Greek history and culture at URL: <http://pandektis.ekt.gr/dspace> [last check, April 30, 2009], the HELLENIC PH.D. DISSERTATION THESIS database at URL: <http://argo.ekt.gr/opac2/zConnectENU.html> [last check, April 30, 2009]. This database consists of the national archive of the Hellenic Dissertations. It includes elements for dissertations that have been written in Greek universities as well as for dissertations written abroad by Greek Doctors. The database consists of the 80% of the total dissertation production in Greece; it includes the dissertations since 1985 until today, as well as a proportion of 5% of the dissertations from 1932 until 1985.

amples of academic repositories and DL in Greece.¹⁰⁵ These include NEMERTES which is the institutional repository of the University of Patras that aims to accommodate all the intellectual work produced by the academic and research staff,¹⁰⁶ PSEPHEDA which is the academic digital repository of the University of Macedonia,¹⁰⁷ ANEMI which is a DL founded by the University of Crete's Library with the aim to provide simple and quick access to a rich collection of digitized material related to modern Greek studies,¹⁰⁸ PERGAMOS, an integrated DL system that offers a platform for the perseverance, documentation, cataloguing, management and prominence of various and heterogeneous digital collections of the National and Kapodistrian University of Athens,¹⁰⁹ DSPACE, which is the DL of the University of Piraeus,¹¹⁰ PANDEMOS, the DL of the Panteion University,¹¹¹ EUREKA! which is the open access institutional repository of AT-EI of Thessaloniki,¹¹² E-LOCUS, the institutional repository of the University of Crete's library,¹¹³ @NAKTISIS, the DL of the TEI of Western Macedonia,¹¹⁴ ESTIA which is the digital repository of the Harokopio University,¹¹⁵ and THEOFRASTOS, which is the DL of the School of Geology of Aristotle University of Thessaloniki.¹¹⁶ However, none of these few examples of institutional repositories and academic DL have the force of law. Last, but not least, one cannot miss

105. Ten out of twenty three Greek academic institutions and four out of sixteen technological educational institutions have developed digital collections and repositories aiming at Greek digital Grey Literature. See Nikolaidou, M. Digital Repositories: The Greek Reality, International Conference Open Access Infrastructures: The Future of Scientific Communication, Hellenic National Research Foundation & National Documentation Centre, 2009, available at URL: <http://www.openaccess.gr/dotAsset/13728.pdf> [last check, April 10, 2009].

106. See Nemertes at URL: <http://nemertes.lis.upatras.gr/dspace/?locale=en> [last check, April 30, 2009].

107. See Psepheda at URL: <http://dspace.lib.uom.gr> [last check, April 30, 2009].

108. See Anemi at URL: <http://anemi.lib.uoc.gr> [last check, April 30, 2009].

109. See Pergamos at URL: <http://pergamos.lib.uoa.gr/dl/index> [last check, April 30, 2009].

110. See DSpace (in Greek only) at URL: <http://digilib.lib.unipi.gr/dspace> [last check, April 30, 2009].

111. See Pandemos (in Greek only) at URL: <http://library.panteion.gr:8080/dspace> [last check, April 30, 2009].

112. See Eureka! at URL: <http://eureka.lib.teithe.gr:8080/dspace> [last check, April 30, 2009].

113. See E-Locus at URL: <http://elocus.lib.uoc.gr> [last check, April 30, 2009].

114. See @naktisis at URL: <http://eprints.teikoz.gr> [last check, April 30, 2009].

115. See Estia (in Greek only) at URL: <http://195.251.30.202:8080/dspace> [last check, April 30, 2009].

116. See Theophrastos at URL: <http://geolib.geo.auth.gr/digeo> [last check, April 30, 2009].

to report regarding Openness in Greece that fact that there is one political party, namely the main opposition Panhellenic Socialist Movement (PASOK) party in Greece which has set all sites and communication elements of it under the Creative Commons licensing.¹¹⁷ The fact that PASOK has opted for Creative Commons licensing for its political communications is indicative of its orientation towards Openness, and is encouraging for further development both in political and legal frameworks in Greece that could enhance and multiply the Open Access initiatives in the country.¹¹⁸ Most legal devices that support Openness in Greece are used voluntarily like the GRL and LGPL free software licensing and the Creative Commons licensing.

While “open” means “without cost”, it does not follow that it also means “without conditions.” This conditional use of educational resources available in an information commons is a distinctive characteristic of Open Educational Resources and could best be served through the Creative Commons licenses.¹¹⁹ The term Open Educational Resources which first came into use at a conference hosted by UNESCO in 2002, is used to describe the open provision of educational resources in the form of digitized materials, offered freely and openly for educators, students, and self-learners to use and reuse for teaching, learning, and research, which are accessible mainly through ICTs and IP networks and are available for consultation, use and adaptation by a community of users for non-commercial purposes.

Thus, the meaning of Open Educational Resources includes:¹²⁰

1. Learning content such as full courses, courseware packs, content modules, collections, journals e.tc.

117. See the Panhellenic Socialist Movement (PASOK) portal at URL: <http://www.pasok.gr/portal> [last check, April 30, 2009].

118. PASOK's option to leverage upon the openness momentum of Creative Commons licensing resembles President Elect Barack Obama's option to set his political communication during the U.S. Presidency campaign under the Creative Commons licensing indicating his favorable stance towards open access; see more on Barack Obama's CHANGE.GOV copyright policy at URL: http://change.gov/about/copyright_policy [last check, April 30, 2009]. It remains to be seen whether pre-election option will elevate into formal policy and/or legal framework favouring Openness.

119. Organisation for Economic Cooperation & Development, *Giving Knowledge for Free: The Emergence of Open Educational Resources*, 2007, p.34, available at URL: http://www.oecd.org/document/41/0,3343,en_2649_35845581_38659497_1_1_1_1,00.html [last check, April 30, 2009].

120. Organisation for Economic Cooperation & Development, *ibid.*, 2007, pp. 35-36.

2. Tools such as software to support the development, use and reuse, delivery of content, searching mechanisms for locating content, learning management systems, content development tools, communities for content aggregation e.tc.
3. Implementation resources such as the Creative Commons licenses which promote open publishing of materials, design principles and policies that mandate the provisions upon which content is accessible and available for use, reuse, adaptation e.tc.

Openness is about the right and the ability to modify, repackage, and add value to a resource.¹²¹ This kind of openness blurs the traditional distinction between the consumer and the producer of resources. The term “user-producer” is sometimes used to highlight this blurring of roles. In that sense, Openness should leverage upon Open Educational tools that make possible the following three freedoms:

1. The freedom to study a work and apply knowledge offered from it.
2. The freedom to redistribute copies, in whole or in part, of a work.
3. The freedom to make improvements or other changes, i.e. to make adaptations, to the content of a work, and the freedom to release modified copies of it.

Open Educational Resources do not occur just because of the fact that the implementation resources such as the Creative Commons licenses are at hand. Targeting mainly academic institutions in Greece and in consideration of the potential by leveraging upon implementation resources such as the Creative Commons licenses, our recommendations for the development of Open Educational Resources in conjunction with the Creative Commons licenses include the following:

1. Each institution and especially public libraries and depositories funded by the Government should develop and publish its policy on open access, clearly declaring its objectives and interests in providing materials by this means. Template guidelines and model documents should be developed to assist institutions practically in the establishment and management of open access systems, and should include:
 - a. Guidance on the development of institutional open access policies, outlining different models of open access and providing means for determining and reviewing the categories of materials which are to be made available by open access and the scope of open access which is to be afforded, in terms of classes of persons who are to be allowed access and the external rights granted to access and reuse of the materials.

121. Organisation for Economic Cooperation & Development, *ibid.*, 2007, pp. 32-36.

- b. Examples of model institutional open access policies accompanied by explanatory statements of each open access policy.
 - c. Guidance on matters to be considered by formally allocating responsibility to an appropriate office within the institution's governance structure, in order to ensure appropriate ongoing administration of the open access policy.
 - d. Guidance on the operation of copyright and contract in structuring an open access system.
2. In order to ascertain who is permitted to use academic materials deposited in a repository and the extent of the permitted use of such materials, it is necessary to identify the various stakeholders and their respective roles, describe the legal relationships among them, and understand how copyright interests are allocated among them and how the Creative Commons licenses can serve such an allocation.
3. Each institution must address conjointly and make decisions about the following factors for the sustainability of Open Educational Resources projects:
- a. The size, structure, and degree of centralization of the organization which will implement an Open Educational Resources project.
 - b. The types of resources it will offer and the media formats in which these resources will be shared.
 - c. The types of the end user reuse that is most likely to help the project meet its goals.
 - d. Incentives for engaging as many participants as possible.
 - e. Ways to reduce costs while still meeting the Open Educational Resources goals.
 - f. Choose among the many available funding models the one which is most likely to result in levels of funding sufficient to allow the Open Educational Resources project to survive.

Strategic planning and implementation for the creation of Open Access libraries and depositories requires thorough studying of all the aforementioned issues—probably more than these mentioned hereto—while at the same time emphasis should be placed on understanding the perceptions upon Openness in the public. Keeping an inquisitive eye locally, should not distract attention from the environmentalism of Openness and the examples of best practices for depositories and libraries worldwide, and especially notable examples of DL and/or Open Access Journals that leverage upon openness-enhancing legal tools such as the Creative

Commons licences in their strategically important choice to Open Access. In the field of Law studies, there are already twenty-one law reviews which have adopted the Open Access Principles, or have policies that are consistent with them. Leading journals such as *Animal Law*, *Harvard Journal of Law & Technology*, *Indiana LawJournal*, *Lewis & Clark Law Review*, *Michigan Law Review*, *Michigan State Law Review*, *New York Law School Law Review*, *Texas Law Review*, *Vanderbilt Law Review*, and *Wayne Law Review* have signed on, as have all of the journals published by Duke Law School and Villanova Law School.¹²²

In Europe, there are also notable examples of Open Access Law Journals such as *Ancilla Juris* in Switzerland,¹²³ IDP of the Universitat Oberta de Catalunya in Spain,¹²⁴ the *Journal of International Commercial Law and Technology (JICLT)* of the International Association of IT Lawyers in Denmark,¹²⁵ the *Utrecht Law Review (ULR)* of the Universiteit Utrecht in the Netherlands,¹²⁶ which all leverage upon the Creative Commons licensing and in most cases the BY-NC-ND license (Attribution+NonCommerical+NoDerivatives). There are also other notable examples of European Open Access Law Journals such as the *Electronic Journal of Comparative Law (EJCL)* of Tilburg University Schoordijk Institute in the Netherlands,¹²⁷ the *Erasmus Law and Economics Review (ELER)* in Italy,¹²⁸ the *InDret Review on the Analysis of Law* of the Universitat Pompeu Fabra in Spain,¹²⁹ the *International Journal of Communications Law and Policy* of the Centre for Socio-Legal Studies (IJCLP) at Oxford University

122. Raul (2005) Creative Commons and Science Commons Announce Open Access Law Program, Creative Commons, Press Release June 6, available at URL: <http://creativecommons.org/press-releases/entry/5464> [last check, April 30, 2009]. About the Open Access Law Program see Science Commons, The Open Access Law Program, a part of the Science Commons publishing project, supports "open access" to legal scholarship, available at URL: <http://sciencecommons.org/projects/publishing/oalaw> [last check, April 30, 2009].

123. *Ancilla Juris* leverages upon Creative Commons licensing. See *Ancilla Juris* at URL: <http://www.anci.ch> [last check, April 30, 2009].

124. IDP leverages upon Creative Commons licensing. See IDP at URL: <http://www.uoc.edu/idp/7/cat/index.html> [last check, April 30, 2009].

125. JICLT leverages upon Creative Commons licensing. See JICLT at URL: <http://www.jiclt.com/index.php/JICLT> [last check, April 30, 2009].

126. ULR leverages upon Creative Commons licensing. See ULR at URL: <http://www.utrechtlawreview.org/index.html> [last check, April 30, 2009].

127. See EJCL at URL: <http://www.ejcl.org> [last check, April 30, 2009].

128. See ELER at URL: <http://www.eler.org/index.php> [last check, April 30, 2009].

129. See InDret at URL: <http://www.indret.com> [last check, April 30, 2009].

in the U.K.,¹³⁰ the Journal of Academic Legal Studies (JOALS) of the University of Hannover in Germany,¹³¹ the Juridica International of the University of Tartu in Estonia,¹³² the Lex et Scientia of the Universitatea Nicolae Titulescu in Romania,¹³³ the Papers Lextra of the Institut Joan Lluís Vives in Spain,¹³⁴ the Review of Economic Research on Copyright Issues (SERCI) in Spain,¹³⁵ the Rivista di Criminologia, Vittimologia e Sicurezza in Italy,¹³⁶ the SCRIPT-ed of the Research Centre for Studies in Intellectual Property and Technology Law in the U.K.,¹³⁷ to name a few. Unfortunately, there is no analogous example of an Open Access Law journal in Greece, currently.¹³⁸ To the point that we are aware of, though, there is only a digitized collection of materials referring to Civil Law, namely all the dissertations submitted by graduate students in Athens Law School's Graduate Studies in Civil Law program, that is available through the Athens Law School Library of the National and Kapodistrian University of Athens.¹³⁹

REFERENCES

Androutsellis-Theotokis, S. & Spinellis, D. (2004). A Survey of Peer-to-Peer Content Distribution Technologies, *ACM Computing Surveys*, 36(4), pp.335–371, available at URL: <http://www.spinellis.gr/pubs/jrnl/2004-ACMCS-p2p/html/AS04.pdf> [last check, April 30, 2009].

Arms, W. Y. (2001). *Digital Libraries*, The MIT Press.

Atkins, D. (1997). Report of the Santa Fe Planning Workshop on Distributed Knowledge Work Environments: Digital Libraries, Report Version September 20,

130. See IJCLP at URL: <http://www.ijclp.net> [last check, April 30, 2009].

131. See JOALS at URL: <http://www.joals.org> [last check, April 30, 2009].

132. See Juridica International at URL: http://www.juridica.ee/international_en.php?submit_year=1&selected_year=default [last check, April 30, 2009].

133. See Lex et Scientia at URL: <http://lexetscientia.univnt.ro/?lang=en> [last check, April 30, 2009].

134. See Papers Lextra at URL: <http://www.lextra.uji.es/papers> [last check, April 30, 2009].

135. See SERCI at URL: <http://www.serci.org/default.asp> [last check, April 30, 2009].

136. See Rivista di Criminologia, Vittimologia e Sicurezza at URL: <http://www.vittimologia.it/rivista> [last check, April 30, 2009].

137. See SCRIPT-ed at URL: <http://www.law.ed.ac.uk/ahrc/script-ed> [last check, April 30, 2009].

138. See The Directory of Open Access Journals at URL: <http://www.doaj.org> [last check, April 30, 2009], in which there is no entry for Greece.

139. See Athens Law School Library of the National and Kapodistrian University of Athens (in Greek only) at URL: <http://www.lib.uoa.gr> [last check, April 30, 2009].

1997, available at URL: <http://www.si.umich.edu/SantaFe> [last check, April 30, 2009].

Barlow, J.-P., A Declaration of the Independence of Cyberspace, available at URL: <http://homes.eff.org/~barlow/Declaration-Final.html> [last check, April 30, 2009].

Benkler, Y. (2006). *The Wealth of Networks—How Social Production Transforms Markets and Freedom*, Yale University Press.

Bollier, D. (2008). *Viral Spiral: How the Commoners Built a Digital Republic of their Own*, The New York Press, available at URL: <http://www.viralspiral.cc/download-book> [last check, April 30, 2009].

Bokos, G. (2001). *Introduction to Information Science*, Papassotiropoulos.

Boyle, J. (1999). *A Politics of Intellectual Property: Environmentalism For the Net?* available at URL: <http://www.law.duke.edu/boylesite/Intprop.htm> [last check, April 30, 2009].

Boyle, J. (2008). *The Public Domain: Enclosing the Commons of the Mind*, Yale University Press, available at URL: <http://www.thepublicdomain.org/download> [last check, April 30, 2009].

Benkler, Y. (2006). *The Wealth of Networks—How Social Production Transforms Markets & Freedom*, Yale University Press, available at URL: http://cyber.law.harvard.edu/wealth_of_networks/Main_Page [last check, April 30, 2009].

Candela, L., Castelli, D., Ferro, N., Ioannidis, Y., Koutrika, G., Meghini, C., Pagano, P., Ross, S., Soergel, D., Agosti, M., Dobрева, M., Katifori & V., Schuldt, H. (2007). *The DELOS Digital Library Reference Model*, available at URL: http://www.delos.info/files/pdf/ReferenceModel/DELOS_DLReferenceModel_0.98.pdf [last check, April 30, 2009].

Chen, Y., Katz, R. & Kubiawicz, J. (2000). *SCAN: A dynamic, scalable and efficient content distribution network*, Computer Science Division, University of California at Berkeley, available at URL: <http://www.springerlink.com/content/wmxcyyp86urbrnpx> [last check, April 30, 2009].

Chen Y., Edler J., Goldberg A., Gottlieb A., Sobti, S. & Yianilos, P. (1999). *A prototype implementation of archival intermemory*, available at URL: <http://pnylab.com/pny/papers/improto/improto.pdf> [last check, April 30, 2009].

Coyle, K. (2004). *The rights in the Digital Rights Management*, D-Lib magazine, 10 (9) available at URL: <http://www.dlib.org/dlib/september04/coyle/09coyle.html> [last check, April 30, 2009].

Christodoulou, K. (2007). Notes upon the general theory of intangible goods, DiMEE magazine, vol.2007, p.180-197,

De Boever, J. (2007). Peer-to-Peer Networks as Distribution and Publishing Model, available at URL: http://elpub.scix.net/data/works/att/128_elpub2007.content.pdf [last check, April 30, 2009].

Dole, W., Hurych, J. M. & Koehler, W. C. (2000). Values for Librarians in the information age, Library Management, 2000, 21(6) pp. 285-286.

Eberspächer J. & Schollmeier R., First and Second Generation of Peer-to-Peer Systems, In R. Steinmetz and Wehrle K. (Eds.) (2005). Peer-to-Peer Systems and Applications, Berlin Heidelberg: Springer- pp. 35-56.

Iglezakis, I. (2007). Issues related to open content availability in the framework of Information Society, Review of Hellenic Justice magazine, 2007, pp.1065-1082.

Ioannidis, Y., Schek, H.-J. & Weikum, G. (2005). Future Digital Libraries Management Systems: System Architecture and Information Access, 8th DELOS Thematic Workshop, Schloss Dagstuhl, Germany available at URL: http://dbis.cs.unibas.ch/delos_website/D1.1.2%20-%20Workshop%20II%20on%20DL%20Access%20and%20Architecture%20jointly%20with%20WP2%20FUTURE%20Digital%20Library%20Management%20Systems%20System%20Architecture%20and%20Information%20Access.pdf [last check, April 30, 2009].

Gonçalves, M.-A., France, R.-K., Fox, E.- A. & Doszkocs, T.- E., MARIAN Searching and Querying across Heterogeneous Federated Digital Libraries, available at URL: http://www.ercim.org/publication/ws-proceedings/DelNoe01/11_Fox.pdf [last check, April 30, 2009].

Goldstein, P. (2007). Intellectual Property: The Tough New Realities That Could Make or Break Your Business, Portfolio 2007.

Griffin, S. (1998). NSF/DARPA/NASA Digital Libraries Initiative, A Program Manager's Perspective, available at URL: <http://www.dlib.org/dlib/july98/07griffin.html> [last check, April 30, 2009].

Guédon, J.-C. (2001). In Oldenburg's Long Shadow: Librarians, Research Scientists, Publishers, and the Control of Scientific Publishing, Association of Research Libraries, available at URL: <http://www.arl.org/resources/pubs/mmproceedings/138guedon.shtml> [last check, April 30, 2009].

Hoffman, K. (2005). Professional Ethics and Librarianship, Texas library journal.

Kallinikou, D., (2007). Copyright and Libraries, Sakkoulas.

Kallinikou, D., Papadopoulos, M., Kaponi, A. & Strakantouna, V. (2009). Alternative system for non-commercial use of intellectual property in consideration of free P2P file-sharing, available at URL: <http://www.marinos.com.gr/bbpdf/pdfs/msg77.pdf> [last check, April 30, 2009],

Kallinikou, D. (2008). Proceedings of Conference “Archives, Libraries and the Law in the era of Information Society, Athens, February 2-3, 2006, Athens: National Library of Greece.

Kallinikou, D., Papadopoulos, M. & Kaponi, A. (2008). Re-examining the limits of regulation for intellectual property protection on the eve of Openness, December 15-16, 2008, speeches at the international conference titled “Open Access Infrastructures: The Future of Scientific Communication” organized by the National Hellenic Research Foundation & the National Documentation Centre at NHRF’s Leonidas Zervas Auditorium in Athens, Greece, Kallinikou’s part one of speech at URL: http://www.marinos.com.gr/bbpdf/pdfs/EIE_15.Dec.08_Dionysia.pdf [last check, April 30, 2009], and Papadopoulos’ part two of speech at URL: http://www.marinos.com.gr/bbpdf/pdfs/EIE_15.Dec.08_Marinos.pdf [last check, April 30, 2009].

Kallinikou, D., Papadopoulos, M. & Strakantouna, V. (2008). The Creative Commons v.3.0 licenses and Academic Libraries, June 10, 2008, Presentation at Workshop organized by the Ionian University, TAB at Nomiki Bibliothiki EUROPE Conference Hall, available at URL: <http://www.marinos.com.gr/bbpdf/pdfs/msg75.pdf> [last check, April 30, 2009].

Kallinikou, D., Papadopoulos, M. & Kaponi, A. (2008). The Creative Commons v.3.0. GREECE licenses and digital repositories of works of music, May 29, 2008, lecture for an audience of graduate students and professors at the Department of Communication & Mass Media of the National & Kapodistrian University of Athens, available at URL: <http://www.marinos.com.gr/bbpdf/pdfs/msg74.pdf> [last check, April 30, 2009].

Kallinikou, D., Papadopoulos, M. & Karounos, T. (2008). The Creative Commons v.3.0. GREECE licenses as Free Culture applications for the promotion of Open Educational Resources, May 27, 2008, speech at the Conference titled 3rd FREE / LIBRE / OPEN SOURCE SOFTWARE Conference organized by the Greek Research & Technology Network s.a., the e-Business Forum, and the Greek Open Source community, available at URL: <http://www.marinos.com.gr/bbpdf/pdfs/msg73.pdf> [last check, April 30, 2009].

Kallinikou, D., Karounos T. & Papadopoulos, M. (2007). The Greek version of Creative Commons licenses, DiMEE magazine, vol.2007, p.377-386.

Kallinikou, D., Karounos T. & Papadopoulos, M. (2007). The Greek version of Creative Commons licenses, Presented at 16th Pan-Hellenic Librarians Conference 2007, available at URL: <http://www.marinos.com.gr/bbpdf/pdfs/msg63.pdf> [last check, April 30, 2009].

Kotsiris, L. (2005). Copyright Law, 4th edition, Thessaloniki: Sakkoulas.

Kallinikou, D. (2008). Copyright and Related Rights, 3rd edition, P. Sakkoulas.

Koumantos, G. (2002). Copyright, 8th edition, Ant. Sakkoulas.

Krishnan, R., Smith, M., Tang, Z. & Telang, R. (2006). Digital Business Models for Peer-to-Peer Networks: Analysis and Economic Issues, available at URL: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=917899 [last check, April 30, 2009].

Kyrillidou, M. & Young, M. (2002). ARL Statistics 2001-2002, Association of Research Libraries, available at URL: <http://www.arl.org/bm~doc/arlstat02.pdf> [last check, April 30, 2009].

Kyrillidou, M. & Young, M. (2003). ARL Statistics 2002-03, Association of Research Libraries, available at URL: <http://www.arl.org/bm~doc/arlstat03.pdf> [last check, April 30, 2009].

Kyrillidou, M. & Young, M. (2005). ARL Statistics 2004-05, Association of Research Libraries, available at URL: <http://www.arl.org/bm~doc/arlstat05.pdf> [last check, April 30, 2009].

Lemley, Mark A. (2005). Property, Intellectual Property, and Free Riding, Texas Law Review, Vol. 83, p.1031, available at URL: <http://ssrn.com/abstract=582602> [last check, April 30, 2009].

Lessig, L. (2004). Free Culture—How Big Media Use Technology and the Law to Lock Down Culture and Control Creativity, The Penguin Press, available at URL: <http://www.free-culture.cc/freeculture.pdf> [last check, April 10, 2009].

Lessig, L. (2006). Code 2, Basic Books, available at URL: <http://pdf.codev2.cc/Lessig-Codev2.pdf> [last check, April 30, 2009].

Lessig, L. (2002). The Future of Ideas, The Fate of the Commons in a Connected World, Random House, available at URL: <http://www.the-future-of-ideas.com> [last check, April 30, 2009].

Lessig, L. (1999). Code and Other Laws of Cyberspace, Basic Books 1999.

Lessig, L. (2008). Remix—Making art and commerce thrive in the Hybrid Economy, The Penguin Press 2008.

Lessig, L., Answers to Written Questions. The Senate Judiciary Committee, "The Microsoft Settlement: A Look to the Future", available at URL: <http://www.les-sig.org/content/testimony/answers.doc> [last check, April 30, 2009].

Litman, J. (2006). Digital Copyright, Prometheus Books.

Lohmann, F. (2006). IAAL: What Peer-to-Peer Developers Need to Know about Copyright Law, Electronic Frontier Foundation, available at URL: <http://www.eff.org/wp/iaal-what-peer-peer-developers-need-know-about-copyright-law> [last check, April 30, 2009].

Mason, M. K. (2009). The ethics of librarianship, available at URL: <http://www.moyak.com/papers/ethics-librarianship.html> [last check, April 10, 2009].

Marinos, M.-T. (1998). Some Notes upon the status of traditional libraries and public digital libraries under the system of Law 2121/1993, Hellenic Justice Magazine (Elldik) p.1484.

Marinos, M.-T. (1994). The Violation of Intellectual Property Right and of Related Rights, Hellenic Justice Magazine (Elldik) 1(35).

Marinos, M.-T. (2005). Copyright, 2nd edition, Sakkoulas.

Mitrou, L. (2002). The Law in the Information Age, Sakkoulas.

Mitrou, L. (2005). Self-Regulation in Cyberspace, Sakkoulas.

Montagnani, M.-L. & Borghi, M. (2008). Positive Copyright and Open Content Licenses: How to make a marriage work by empowering authors to disseminate their creations, International Journal of Communications Law & Policy, 12, available through SSRN at URL: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1020997 [last check, April 30, 2009].

Papadopoulos, M. (2008). Creative Commons licenses v.3.0, Presentation to the National & Kapodistrian University of Greece, Athens Law School, available at URL: <http://www.marinos.com.gr/bbpdf/pdfs/msg70.pdf> [last check, April 30, 2009].

Papazoglou, V. (2008). Horizontal Action of Academic Libraries: Legal Issues, Proceedings of the Conference Archives, Libraries and the Law in the era of Information Society, Athens, February 2-3, 2006, National Library of Greece, 2008.

Pedley, P. (2007). Digital Copyright, Facet.

Pourebrahimi, B., K. Bertels, K. & Vassiliadis, S., A Survey of Peer-to-Peer Networks, available at URL: http://ce.et.tudelft.nl/publicationfiles/1075_526_prorisc05.pdf [last check, April 30, 2009].

Reeder, M. (2007). Dutch Collecting Societies welcome CC, August 23, 2007, available at URL: <http://creativecommons.org/weblog/entry/7622> [last check, April 30, 2009].

Rubin, R. (2000). *Foundations of Library and Information Science*, New York, Neal-Schuman.

Saez, C. (2008). Improbable Match: Open Licences And Collecting Societies In Europe, available at URL: <http://www.ip-watch.org/weblog/2008/10/28/french-deal-highlights-open-licensing-and-collecting-societies-in-europe> [last check, April 30, 2009].

Schachaf P. & Rubenstein, E. (2007). A Comparative Analysis of Libraries' approaches to Copyright: Israel, Russia and the U.S., available at URL: <http://dlist.sir.arizona.edu/2117/01/approachesToCopyright.pdf> [last check, April 30, 2009].

Schaumann, N. B. (2005). Direct Infringement on Peer-to-Peer Networks, William Mitchell Legal Studies Research Paper No. 9, available at URL: <http://ssrn.com/abstract=703882> [last check, April 30, 2009].

Severson, R. (1995). The recovery of ethics in librarianship, *Journal of information ethics*, 2(2).

Sinodinou, T.-E. (2008). *Intellectual Property & New Technologies*, Sakkoulas.

Strakantouna, V. (2007). Legal deposit of works protected by Copyright, addressed at international conference titled "Rethinking the boundaries of copyright," Istanbul 15-16 November, 2007.

Strakantouna, V., Piskopani, A.-M. & Mitrou L. (2007). Personal Data and Libraries, *Private Law Chronicle (Xronika Idiotikou Dikaiou)*, Z.

Suber, P. (2009). Timeline of the Open Access Movement, revised February 9, 2009, available at URL: <http://www.earlham.edu/~peters/fos/timeline.htm> [last check, April 30, 2009].

Suber, P. (2007). Open Access Overview, Focusing on open access to peer-reviewed research articles and their preprints, revised June 19, 2007, available at URL: <http://www.earlham.edu/~peters/fos/overview.htm> [last check, April 30, 2009].

Thorne, M. (2008). Danish Collecting Society KODA teams up with CC Denmark, January 31st, 2008, available at URL: <http://creativecommons.org/weblog/entry/8012> [last check, April 30, 2009].

Travis, H. (2005). Building Universal Digital Libraries: An Agenda for Copyright Reform, *Pepperdine Law Review*, vol. 33, pp. 761-829, available at URL: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=860784 [last check, April 30, 2009].

Zittrain, J. (2008). *The Future of the Internet and How to Stop it*, Yale University Press 2008, available at URL: <http://futureoftheinternet.org/static/ZittrainTheFutureoftheInternet.pdf> [last check, April 30, 2009].

The Mirror and the Spell: Order and Organization in the Age of the Digital Reserve

Hannah Knox^{*},
Centre for Research
on Socio-Cultural Change (CRESC).

Damian O'Doherty^{}**,
Manchester Business School,

Theo Vurdubakis^{*}**
& **Chris Westrup^{****}**
Centre for the Study of Technology
and Organization (CSTO)

The paper takes as its starting point the diffusion of ICT applications associated with so-called 'customer relationship management' (CRM). CRM encourages organizations to shift their understanding of customers from an episodic and

^{*} Hannah Knox is a Research Associate at the ESRC Centre for Research on Socio-Cultural Change (CRESC) in the University of Manchester. Her research focuses on the anthropological study of technological projects of social change. She is currently working on an ESRC project which is looking at road building in Peru. Address: Centre for Research on Socio-Cultural Change, 178 Waterloo Place, The University of Manchester, Oxford Road, Manchester, M13 9PL, UK. [email: hannah.knox@manchester.ac.uk]

^{**} Damian O'Doherty is a Senior Lecturer in Organization Analysis at the Manchester Business School in the University of Manchester. His research examines organization in terms of order and disorder and has published his work in a number of journals including Culture and Organization, Sociology, The Sociological Review, The Journal of Management Studies, International Studies in Management & Organization and Ephemera. Address: Manchester Business School, The University of Manchester, Booth Street West, Manchester M15 6PB, UK. [email: damian.odoherty@mbs.ac.uk]

^{***} Theo Vurdubakis is Professor of Organization and Technology at the Department of Organization, Work and Technology, Lancaster University UK. He is also the director of the Centre for the Study of Technology and Organization (CSTO). His research focuses on the role of technologies in social organization. He is currently working on an Economic and Social Research Council project investigating knowledge management and enterprise systems. Address: Centre for the Study of Technology and Organization, Lancaster University, Lancaster, LA1 4XY, UK. [email: t.vurdubakis@lancaster.ac.uk]

^{****} Chris Westrup is a Senior Lecturer at the Manchester Business School in the University of Manchester. His interests include information systems and changing forms of organization and the use of information systems in development. He has recently been a member of an ESRC funded study with his fellow co-authors. Address: Manchester Business School, The University of Manchester, Booth Street West, Manchester M15 6PB, UK. [email: chris.westrup@mbs.ac.uk]

transaction-based perspective to one that emphasises continuous 'relationship management'. CRM applications thus promise to deliver more, and increasingly accurate, information about consumer habits and behaviours therefore allowing organizations to maximise their extraction of business value. In this paper we explore the ways in which such inscriptive technologies are not merely referential but also performative of contemporary re-presentations and ideals of the consuming subject. Focusing on what we might call the 'digital subjects' of customer relationship management we also explore how such inscriptive apparatuses also appear endemically prone to instability and representational excess. Through an investigation of managerial imagery of the computer enabled CRM, the paper explores the ways in which ambiguity and ambivalence continue to haunt advances in corporate technologies of surveillance and tele-control.

Keywords: Customer Relationship Management; Digital Subjects; Information and Communication Technologies; Surveillance; Performativity.

Introduction

'One of the most significant products of the culture of capitalism', argues Dan Rose (1990:158) 'is reality', as it is 'continually formed and reformed within large companies'. Inquiry into reality, he suggests, 'must thus be inquiry into reality formation and reality consumption'. Clearly, any such inquiry cannot but also be an inquiry into the information and communication technologies (ICTs) through which the twin labours of reality formation and reality consumption are increasingly performed. To paraphrase Austin (1962), the still unfolding 'computational rendition of reality' (Kallinikos, 2006) imparts social agents the ability 'to do things with symbols', the power so to speak, to manipulate the world through the manipulation of signs. At the same time, this technological mediation of the world, and of the human subject, remains a socially, culturally and historically situated set of practices and procedures. As such, the computational machineries of reality fabrication and consumption are, we shall argue, also ways of acting out longstanding cultural preoccupations.

Our focus in this paper is the consuming subject¹ as an object of the corporate and institutional apparatuses of reality 'fabrication' and 'consumption'. In line

1. The argument presented in this paper draws upon an ongoing research engagement with the role played by ICT applications in the production and consumption of business knowledge. The most recent round of empirical work was conducted between 2004-6 in 4 organisations: One major systems vendor and three business users. All authors were involved in the empirical research, which combined the use of a range of qualitative research methods. These included a series of open-ended and semi-structured interviews, non-participant observation and the investigation of documents. The majority of interviews were carried out with two or more

with much of social theory (e.g. Hayek, 1949; Giddens, 1994a;b), management texts typically represent the consuming subject as an active, autonomous and increasingly knowledgeable agent. Collectively, consumers comprise an often nebulous but nevertheless powerful entity: 'the market'; and as everybody knows, the market cannot be 'bucked'. Consumer choice, exercised through the market, thus functions as a sort of 'reality principle' for business enterprises. In one of a series of turn-of-the-century IBM² television ads on 'business nightmares', a businessman is describing to his psychoanalyst a disturbing dream in which he is being chased by an anonymous crowd whom he cannot outrun.

Analyst: "From whom?"

Businessman: "I am not sure. They are wearing masks – They are tough customers".

Analyst: "Customers? What do they want?"

Businessman: "Everything! Now!"

The meaning of the dream, the psychoanalyst opines, is that he is too slow, he can't adapt - and he is 'in denial'. Such a diagnosis is hardly surprising. In the 'digital era', it has become an article of faith that traditional business models are in the process of breaking down under pressure from increasing competition and consumer empowerment and sophistication. This new reality requires new 'individual [customer] centric' models of doing business. There is an urgent need, it is claimed, to re-configure the organization in ways that ever more closely reflect the shifting consuming subject. Staying for a moment with the psychoanalytical conceit of the IBM ad, it is worth noting the premise that the anonymity of the customer (underscored by the wearing of blank masks – see figure1) is the stuff of which business nightmares are made.



Figure 1: The chase - from an IBM 'business nightmare'.

members of the research team present. A key feature of the methodology was our efforts to see and hear the interviews not simply as an exercise in information retrieval, but rather to engage with the interview as an occasion of 'practical reasoning' in which organisation (as a verb) was being variously represented and 'accomplished' by its members and users. See for instance Authors 2007a;b; 2008 and forthcoming for accounts of this work. We gratefully acknowledge the support of the (UK) Economic and Social Research Council's Evolution of Business Knowledge research programme (Grant no RES-334-25-0012).

2. IBM e-Business Solutions. Uploaded 22 May 2003.

Typically, the 'problem' of the customer (the object of knowledge) is posed as one of fragmentation. While electronically mediated transactions produce more and more recorded trails, much of this material remains, it is claimed, post hoc and unsystematic (Abbott, 2001). The meaningful re-integration of this data is thus the Holy Grail of 'forensic marketing'. Customer relationship management (CRM) is an ICT mediated approach that aims to '[align] business processes with customer strategies in order to build customer loyalty and increase profits overtime' (Rigby et al, 2002:102). In recent years, CRM standard packages have tended to become 'obligatory passage points' (Latour 1987) in corporate attempts to effect such an alignment. CRM systems promise user organizations the ability to, inter alia, identify customers by attribute and behaviour; distinguish between them by profit contribution; facilitate better decisions on product design and promotion; target customers as individuals and as segments; as well as measuring promotional effectiveness and return on investment. CRM packages are thus promoted as fully-fledged Latourian (op.cit) 'centres of calculation' where a comprehensive (and it is claimed predictive) knowledge of the consuming subject can be assembled out of partial traces (such as transaction data, surveys, questionnaires, loyalty schemes, e.tc.). CRM thus claims to do what 'all the king's horses and all the king's men' could not, and put Humpty Dumpty together again.

The Family from Hell and Other Tales of Twenty-first Century Consumption

CRM is said to represent the culmination of a decade long shift away from an emphasis on the management of transactions to the management of relationships. In the course of a presentation at FreeWorldSys³, (a vendor of ICT 'business solutions'⁴) this was emphasized by 'Anna', one of our interlocutors, when – for dramatic effect – she pointed in quick succession at her wedding ring and then at an image of a supermarket loyalty card arguing that both those artifacts meant that a relationship was in existence:

"A different sort of relationship sure, but one that also needs work if it is to continue...Your customers have signed on the dotted line, they have taken your shilling⁵... They have taken you into their confidence, they are [en]trusting you with their habits and preferences."

In this account, the loyalty card represents the willingness to be identified, the willingness to, as it were, stand out from the faceless multitudes of the 'mar-

3. A pseudonym.

4. 'We take care of the technology so you can take care of your business'.

5. An allusion to the 'Queen's shilling' given by recruiters to enlistees in the British armed forces.

ket'. CRM, she went on to argue, enables business to build on this recognition. It is about the realisation that 'your customer is a resource, your main resource'. Until relatively recently however the knowledge necessary to fully render the customer into a productive resource was said to be lacking, to have been out of date or to have remained entrapped in various informational 'silos'. Members of the audience were encouraged to reinforce this point by airing their own tragicomic tales of corporate flat-footedness, ignorance and confusion. For example a manager in a major financial services organization reminisced about the bad 'old days' (1970s-80s) when information about bank customers was kept by account number only and the organization had but the vaguest knowledge about who banked with them or how many accounts a given customer had. Now, however, things are different. According to an FreeWorldSys publication,

'FreeWorldSys' CYCLOPSTM [software platform]⁶ ...by leveraging customer and transaction information already in your systems, we provide individually targeted loyalty programmes that maximise the lifetime value of every customer'.

The imagery of lack followed by technological fulfillment tended to underpin the accounts of our interlocutors, such as the tale of 'The Family from Hell' that was related to us during a visit to FreeWorldSys. Different members of the family in question, so the story goes, would lurk near the checkouts of various branches of 'GoodSense'⁷, a large British supermarket chain. The family would collect discarded receipts looking for the ones where the shopper had not collected loyalty points. They would then approach the checkout staff and, claiming the receipt as their own, ask that the missing points be registered retrospectively to their account. By this means, the family is said to have been amassing (and cashing out) loyalty points in astronomical quantities. Their *modus operandi* however is, in the tale, also the cause of their undoing as the customer loyalty account failed to show anything approaching a stable shopping 'pattern'. This in turn enabled GoodSense to discover the scam, identify the perpetrators, and ensure a conviction: a triumph then for corporate knowing over fraudulent behaviour. However the tale as told to us in FreeWorldSys hinted at a different moral. According to Anna the story told not only of a victory for the corporate gaze but also of its blind spots. For her it was instructive that all the action in the tale takes place around supermarket checkouts.

'What happens up to now? Say the customer walks in. And they have your loyalty card. You do not know they are in here. You only find out they've

6. A pseudonym.

7. A pseudonym.

been when it's too late [to sell them anything], when they are at the check-out. Well, that's what we are now changing'.

In this type of narrative, technological innovations such as the introduction of smart cards and RFID (radio frequency identification) technologies have the potential to transform what used to be post hoc into 'real time' knowledge. The chip in the loyalty card will be able to signal the presence of the customer as soon s/he comes within read range. If so desired, closed circuit television cameras can track the bearer's movements throughout her/his visit. To quote again: 'You will be able to see what they looked at and did not buy but put it (sic) back on the shelf'. It is of course easy to make too much of the 'ideal logic' of a technological system. For instance, in the course of illustrating the above point, Anna had to excuse herself: she was too short she explained for the camera to be able to focus on her. She quickly pulled out a crate apparently kept there for that purpose, climbed on top and, what we might call, the 'eye of the CYCLOPSTM' obligingly swerved to focus on her.

CRM could be understood as located at the nexus of a number of technologies of visibility, inspection and inscription (e.g. Lace, 2005). According to the late Michel Foucault (1977), contemporary techniques and practices of surveillance and documentation should be viewed as attempted solutions to the political-administrative problems posed by the 'mass'. In the modern era, he claimed, rational administration can no longer tolerate the opacity and sheer unknowability of the mass. In Western modernity, it is no longer the 'anonymous' masses that observe the rituals and ceremonies of the elites. Instead it is now the many who are subjected to the gaze of the few. The various administrative apparatuses of surveillance, Foucault (1977:191) argued, have "lowered the threshold of describable individuality and made of this description a means of control". For Foucault, this individuality is fabricated by means of a range of administrative technologies and devices such as identity documents, dossiers, filing cabinets, and ultimately, calculating machines. These work to individualise the formerly 'nameless' mass(es) into knowable and calculable subjects. There is by now an extensive literature that identifies and explores how these processes operate in contemporary forms of organizing (e.g. Miller and O'Leary, 1987; Knights and Morgan, 1993). Increasingly, it is not only the employees of an organization who are subject to the corporate gaze. Computer-assisted CRM offers the means through which an organization can expand its knowledge and control of the customer and hence extend its operations beyond currently taken-for-granted organizational boundaries. The technologies upon which CRM depends provide the means for carrying out the ongoing corporate labour of assembling and putting into circulation a veritable 'digital double' of the consumer. In principle this is nothing new: marketing has always sought to construct ideal-type classifications. Nevertheless, if as

Walter Benjamin (1979) wrote, 'every day the urge grows stronger to get hold of an object at a very close range by way of its likeness, its reproduction', then the promise of contemporary information technologies is the ultimate gratification of that urge.

It should be clear from what has been said so far that role of such digital doubles is not merely referential but also performative. They could be viewed as variants of what Charles Sanders Pierce (1960) has called 'indexical icons' or rather, as a special creative kind of indexical icons. They constitute 'a self-reflexive use of reference that in creating a representation of an ongoing act, also enacts it' (Lee and LiPuma, 2002:195). For instance, in FreeWorldSys and elsewhere we were proudly shown various technological devices designed to utilize the information collected and archived in order to predict and shape the behaviour of the consuming subject. Among those devices, the one that is to-date most widely deployed in actual retail settings is the 'smart trolley' (e.g. Ody, 2003). A trolley is 'smartened-up' by being equipped with a small computer which identifies the customer by 'reading' the information on a loyalty card. The computer is able to sense the proximity of the individual shopper to a product and drawing on stored knowledge of that individual's purchasing history adjust the prices shown on its miniature screen: "For you Harry, this Argentinean Red, 20% off!"⁸. A 'smart trolley' then, can give voice to a siren song of consumer commodities. The mission of CRM technologies then is as summed-up in Hamel and Prahalad's (1996:Ch.4), exhortation, "to amaze customers by anticipating and fulfilling their unarticulated needs".

As we have seen, the promise of CRM is often expressed as the vision of the customer as a 'resource'. Consumer behaviour will, it is hoped, no longer be a source of uncertainty – the stuff business nightmares are made of – but will become instead the object of reliable, even predictive, knowledge. The various information technologies employed in the delivery of CRM (including RFID tagging, data mining, and so on) thus collect, compile and combine continually updated data (e.g. Kallinikos, 2006) -including age, income, education, consumption habits, credit history and the like, in order to fabricate and place under corporate control ever more efficacious electronic re-presentations of the consuming subject. Perhaps such labours are, as some argue, invested in a fantasy and the consumer is indeed ultimately 'unmanageable' (Gabriel and Lang, 2006). Nonetheless, simulation, as Bogard (1996) argues (drawing on Baudrillard), is precisely the process that feigns what it does not (or cannot) possess (Nichols, 2004). The figure of the digital double thus appears to enjoy a peculiar, even ambivalent, form of exist-

8. Anna in the course of a 'smart trolley' presentation.

ence. An existence located in the intersection of reality and fantasy⁹, of the actual and virtual worlds (Bogard, *ibid*: 27).

The post-modern predicament of yakov petrovitch golyadkin

It is perhaps more than a coincidence that the seemingly inexorable spread of the administrative machineries of representation described by Foucault (1977), was, almost from the beginning, haunted by spectres of the subject in crisis. From E.T.A. Hoffman's stories of doppelgängerism, to Robert Louis Stevenson's Dr Jekyll and Mr Hyde, to Ira Levin's *Stepford Wives*, the tale is told of the Occidental subject prone to fragmentation and dissolution (e.g. Bainbridge, 2006). Against this backdrop the ambivalent figure of the 'double' functions as a potent vehicle for such anxieties (Herdman, 1990). In Dostoevsky's 1845/1866 story *The Double*, for instance, the anti-hero, Titular Councillor Yakov Petrovitch Golyadkin, is haunted by another Golyadkin, one who is

'completely different but at the same time absolutely identical....so that if one were to ... place them side by side, no onewould have taken upon himself to determine just who is the real Golyadkin and who the counterfeit, who the old and who the new, who the original and who the copy.'

Golyadkin finds himself unable to shake off this 'other' Golyadkin, who keeps impersonating him at work and at home and interfering in his life, gradually driving him mad.

'Either you or I, but not the two of us! And therefore I am declaring to you that your strange, ludicrous and at the same time impossible desire to seem to be my twin and to pass yourself off as such will only lead to your utter dishonour and defeat'.

Alas, as Golyadkin comes to realise, he is fighting a losing battle. Wherever he goes, more 'exact likenesses' of his' keep springing up.

'And all of these exact likenesses, immediately upon making their appearance, began running after the other and stretched out like a string of geese waddling after Mr Golyadkin Senior, so that there was nowhere to escape from these exact likenesses.'

In Dostoevsky's tale, the origin of these doubles remains obscure, leaving open the possibility that they might be products of Golyadkin's increasingly disturbed mind. Oscar Wilde's (1891) *Picture of Dorian Gray*, on the other hand, locates the genesis of the double in the realm of representation. Dorian Gray never seems

9. For the role fantasy plays in ICT research and design see Boland (1986).

to age, despite living a life of debauchery and dissolution. Instead the symptoms only show up on his portrait kept locked away in the attic. There has been in other words an illicit substitution of the represented for the real, of the sign for the referent: the double sin of witchcraft and idolatry¹⁰. The Faustian nature of these substitutions is in turn further spelled out in the 1913 German film, *The Student of Prague* (see Baudrillard, 1988). In the film, the Devil offers an impoverished university student a great deal of money for something seemingly insignificant: his reflection in the mirror. The student, Balduin, readily agrees and the Devil summons the image from the mirror and walks away with it. The newly wealthy Balduin does not miss his mirror image, not, that is, until he suddenly comes across himself. The Devil has endowed Balduin's reflection with an independent life and has put it into circulation. Now there are two of them. Like Golyadkin, Balduin engages in a futile struggle to reclaim his identity as the double starts to shadow him, interfere in his life, and even commit crimes for which he is blamed.

If information age folklore is to be believed, Golyadkin's and Balduin's predicament is now becoming endemic (e.g. Johnson, 2008; see Figure 2). In one well-known tale of contemporary doppelgängerism, former salesman Bronti Wayne Kelly spent years unable to obtain employment despite extensive experience and good qualifications. During his time in the wilderness he accumulated hundreds of rejections from prospective employers. In those rare cases where he was offered a job, he would be fired within days and without an explanation. Kelly's savings soon ran out, he had to sell his house and file for bankruptcy. He became homeless, sleeping rough in the streets and washing in public toilets. It was only after four desperate years that he was eventually given an explanation. Kelly's wallet had been stolen back in 1990 while Kelly was serving as a USAF reservist. The wallet had contained his driver's license, Social Security card, and military ID. It turned out that during all those years the (still unknown) thief had been giving Kelly's identity to the authorities every time he was arrested for his numerous crimes. Thus, Kelly's 'data profile', circulating and proliferating in the network of computer databases used by employers in their background checks, included an ever lengthening record of arrests and convictions for crimes ranging from shoplifting to arson. Eventually the police provided Kelly with a 'Certificate of Clearance,' which he has to carry with him, and which officially states that the authorities had determined that Bronti Kelly was not the criminal with whom he co-habits that identity. Even this however seems to have failed to rid Kelly of his malevolent 'digital double'. He has since been rejected from another 50 jobs – a sign that the erroneous data is continuing to

10. That is performative signification, and objectified subjectivity respectively (see Hawkes, 2004).

haunt him. He is said to be considering changing his name, the ultimate surrender so to speak, to the 'Other Kelly'.

To a computer this could be you



Figure 2: Bank mail-shot addressed "To the Real Mr Vurdubakis".

For many contemporary commentators, corporate concerns with protecting the public from identity theft merely diverts attention from the real villain, being none other than those same corporations whose data collection and manipulation activities increasingly imperil the autonomy (to say nothing of the privacy) of the subject. The so-called 'sovereign consumer' thus appears as a kind of 'corporate dope' with the invisible hand of the market seemingly increasingly replaced by the concealed hand of management (e.g. Bakan, 2006; Packard, 1970). Corporate digital doubles and their associated computer mediated, forensic, actuarial and diagnostic procedures, are even viewed as symptoms of an ongoing transformation of Euro-American societies into 'control societies' where:

'Marketing is now the instrument of social control and produces the arrogant breed who are our masters. Control is short term and rapidly shifting, but at the same time continuous and unbounded, whereas discipline was long term, infinite and discontinuous' (Deleuze, 1995:181).

According to The Economist (1999) the Faustian deal has already been struck. Privacy, the editor argues, has been irreversibly eroded. This erosion does not come about as the result of the actions of some Orwellian central authority. Rather, it is the logical outcome of consumers' willingness to give away piece-by-piece seemingly insignificant but ever-increasing amounts of personal data in return for various benefits such as supermarket loyalty points. (To paraphrase Arendt (1965) we might speak of the banality of surveillance.) While it is no longer possible to restore even the levels of privacy enjoyed in the 1970s, most people, the editorial claims, do not actually care. They do not miss what has been given

away. Further erosion is therefore inevitable, so '[t]he best advice is: get used to it'(ibid: 12). If that is indeed the case, then Euro-American consumers appear to have gotten a particularly bad deal, having sold their doubles for a fraction of the price obtained by Balduin....¹¹

For those whose faith in the science(s) and technologies of management survives unshaken, it may be tempting to interpret the various tales of disruption and dop-pelgängerism recounted in this and the previous sections as merely describing sites and instances where the technological-administrative labour of reference was not properly performed. (Recall for instance the moral of the tale of the 'family from hell' as related to us in FreeWorldSys). In this view, such phenomena point to the need for further, and more creative, technological and administrative solutions to the problem of the (faceless) mass, solutions that would admit less noise and outside interference.

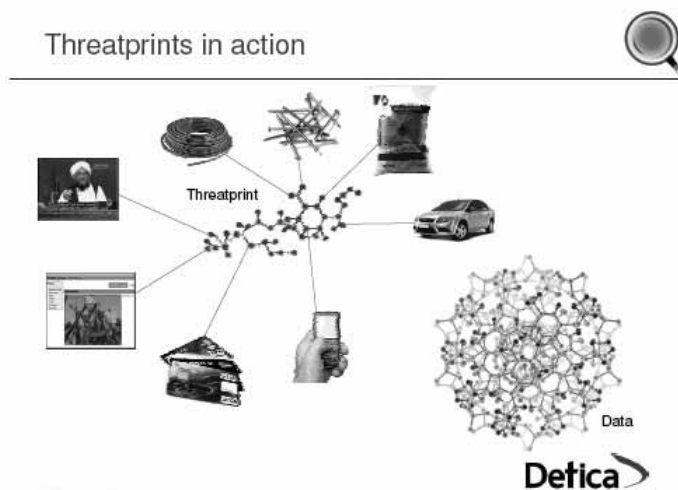


Figure 3: 'Threatprint' (Source: Sutherland/Detica, 2008).

In this vein the IT consultancy Detica recently proposed a technique for predicting criminal activities (especially terrorism) that, it claims, will enable law enforcement agencies, 'to avoid being overwhelmed by data'; (ii) 'to avoid a systemic invasion of Privacy'; (iii) 'to predict, and ultimately influence' [criminal and] terrorist behaviour' (Rutherford/Detica, 2008:9). This technique is the 'threatprint' a virtual object assembled via the mining and profiling of electronic data generated by eve-

11. The Devil pays Balduin 100.000 gold pieces in the 1913 film. By the time of the 1926 remake (dir. Henrik Galeen), the price had risen to 600.000 gold pieces, possibly a reflection of Weimar era inflation.

ryday consumption activities, movements and social interactions (see Figure 3). By focusing on linkages between the 'digital footprints' generated by possible criminal activities (rather than persons) it is possible to trace out and pre-dict the contours of events that are yet to happen. The electronic traces of a criminal/terrorist action, it is argued, pre-figure the event itself and a thorough study of such traces in the light of particular threat scenarios (the 'threatprint') will reveal the preventative actions to be taken in the present.

Steven Spielberg's (2002) *Minority Report* can be seen as a dramatisation of the cultural logic that underwrites both CRM and notions like 'Threatprints'. Based on a story by Philip K. Dick (2000), the film depicts the society in the year 2054 as one preoccupied with the quest for predictive knowledge of the human subject. One way in which this preoccupation is manifested is in the operations of "Pre-crime". The aim of the "Pre-crime" is the identification and apprehension of criminals before their crimes have been committed. In the film, the hero, John Anderton (played by Tom Cruise), is on the run having been officially identified as a pre-criminal. Anderton attempts to evade detection in the anonymity of urban crowds. In 2054, however, this is no longer possible. Foucault's (1977) opaque anonymous mass which had so offended the modern will to knowledge and control is no more. Public spaces are saturated with commercial iris scanners designed to identify, monitor and target consumers. Wherever Anderton goes he is hailed by name by the 'smart' advertising billboards attempting to entice him: "John Anderton, you look like you could use a Guinness!" By 2054, it appears, Hamel and Prahalad's (1996) vision has indeed come to pass. In desperation Anderton resorts to an eye transplant which is carried out by an underground doctor. After the operation he is shown entering a GAP store whereupon he is immediately greeted by the apparatus as "Mr Yakamoto" - presumably the late owner of his new black-market eyes - and asked whether he enjoyed the tank tops he acquired in his previous visit there.

Minority Report clearly belongs to the popular cinematic genre of 'humanist' critiques of what we might call 'administrative reason' and its technological (over) writing of the world (Czarniawska and Joerges, 1999). At the same time, however, the film partakes of the phenomena it critiques. Not only is it a commercial product in itself, but it is also a vehicle for other commercial products. The extensive usage of brands such as GAP, Guinness, or American Express, is both a deliberate narrative device and part of a lucrative product placement strategy through which Spielberg was reportedly able to defray 20% of the production costs. The film therefore partakes of the practices and techniques it critiques.

In the wave of publicity that accompanied the film's release much was made of the fact that the technological devices portrayed were designed by a panel of

technical experts and futurists convened by Spielberg with the aim of producing an accurate picture of the future. In fact *Minority Report* was often referenced by our interlocutors, usually in order to note how something that had appeared amazingly futuristic only two or three years earlier, could be achieved “right now” with RFID technology much more cost effectively¹². According to ‘Jerry’¹³, a management consultant specializing in Enterprise System implementations including CRM, RFID should be understood as performing in the ‘real world’ a role not dissimilar to that of cookies. The shopping experience as portrayed in *Minority Report*, he argued, is one already familiar to online shoppers who, as a matter of course, expect to be recognized, to have shopping suggestions thrust at them, and so on¹⁴. Widespread use of RFID would therefore allow this mode of consumption to seep out of the ‘virtual’ into the ‘real’ world. There are some signs that such “seepage” is taking place. UK retailers such as the House of Fraser, Tesco and Marks & Spencer have recently (and controversially) been experimenting with RFID tags. In what is probably the best-known case, Phillips announced that it was providing clothing retailer Benetton with washable RFID tags, which would be woven into the labels of 15 million items (e.g. RFID Journal, 2003a). Benetton sales staff, it was argued, “could easily identify repeat visitors on arrival and give priority service to these more loyal - and hence more valuable - customers” (Banham, 2003). Even though the ensuing outcry seems to have resulted in Benetton’s plan being put on hold, such cases were nevertheless often mentioned to us as indicating the building up of a possibly unstoppable technological and business momentum (see RFID Journal, 2003b; 2003c) (see Figure 4). As Evans (2005:111) notes, RFID tags in clothing offer a cost-effective “alternative [to *Minority Report*’s biometrics] way of identifying customers - as long as they didn’t swap clothes!”

12. In the film, it will be recalled, it is biometrics that is used to identify and monitor customers.

13. A pseudonym.

14. Jerry was critical of the way the future of consumption was portrayed in *Minority Report*. For him it illustrated a confusion of the medium with the message, in that essentially mass media such as billboards were being used in one-to-one marketing.

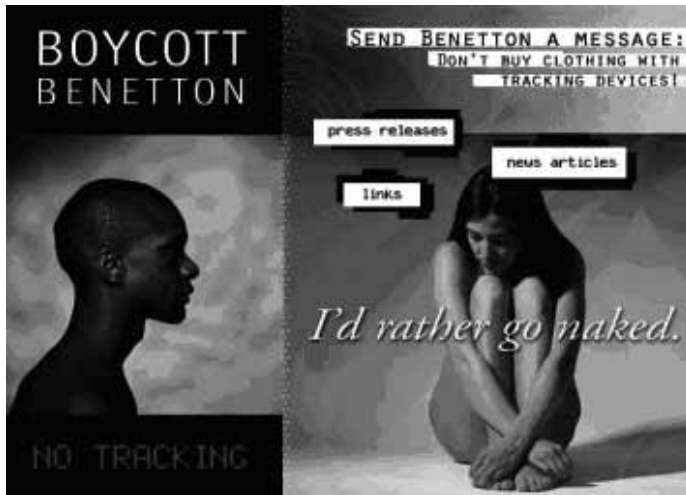


Figure 4: Anti-RFID protest (Source: www.boycottbenetton.com).

Much of the technological gadgetry discussed here, from CYCLOPSTM, to the biometric technologies portrayed by Spielberg (see also Evans, 2005; Knights et al, 2000), to the weaving of RFID tags in clothing, can be viewed as enrolled in the performance of the complex referential work deemed necessary to keep representations and represented, signs and referents, properly aligned with one another. If the vision of 'Customer Relationship Management' in *Minority Report* (the 'future of CRM') is compared with that of IBM's 'business nightmare' described earlier (presumably now the past of CRM) what is particularly striking is the reversal in the imagery of the chase. In the IBM advertisement it is the businessman that is chased by a faceless crowd of consumers. In the future conjectured by *Minority Report* on the other hand, it is the (no-longer anonymous) consumer that is relentlessly pursued by the electronic apparatuses of customer relationship management.

The theme of all-seeing 'hidden persuaders' (Packard, 1970) is of course a common trope in conspiracy theories (see Parish and Parker, 2001). Indeed, the RFID debate has often shaded into conspiracy theories of varying degrees of (im)plausibility. For instance, Albrecht and McIntyre's (2005) well-received *Spychips: How Major Corporations and Government Plan to Track Your Every Move with RFID*, has also been published (in the US) under the title *The Spychips Threat: Why Christians Should Resist RFID and Electronic Surveillance* (2006) including additional material which describes RFID tags as analogous to the 'Mark of the Beast' prophesied in the Book of Revelation (e.g. op.cit:xii).

Information technology and the performativity of the sign

Dorian Gray's picture, Balduin's mirror image, and Minority Report's Pre-crime (or for that matter 'pre-consumption') are all testaments to an enduring fascination with, what we might call, the double life of the sign: representations are portrayed as not merely referential, but as performative. Commenting on the view of sympathetic magic propounded in James Frazer's (1890) *Golden Bough*, Taussig (1993:47-8) dwells upon Frazer's

"notion of the copy, in magical practice, affecting the original to such a degree that the representation shares in or acquires the properties of the represented. To me this is a disturbing notion foreign and fascinating not because it so flagrantly contradicts the world around me but rather, that once posited, I suspect if not its presence, then intimations thereof in the.... habits of representation in the world about me".

Intimations -or perhaps something more? Current Euro-American habits and practices of representation are crucially dependent upon the various 'information' technologies which mediate social organization. 'Modern' information technology and ('pre-modern') magic could be said to share a particularly close kinship in that both are conceived as loci for the ability to effect transformations of the material world through the manipulation of 'mere' symbols (Davis, 1999). This claimed power is best encapsulated in the concept of a 'virtual reality', as popularised by the 'cyberpunk' literary genre (e.g. Gibson, 1984; 1997) and as depicted in films like the Matrix trilogy (dir. Wachowski & Wachowski, 1999: 2003). In this sense, the different applications and associated imageries that make up CRM could be viewed as reflecting a broader ('post-modern') cultural preoccupation with the efficaciousness of the electronic sign.

Of course, in 'pre-modern' Christian Europe, magic, the esoteric art of the efficacious sign was widely assumed to presuppose and require a deal with the Devil. For instance, Hawkes (2004) notes that for Thomas Aquinas the agency that executes a magician's spells

'cannot be the magician himself, for human beings cannot achieve objective effects by subjective force alone. It cannot be God, for God does not submit His will to human command, nor can He be invoked with spells or images. It cannot be the signs themselves, for signs naturally possess no performative power. The agent who performs the magical action can only be a spirit who does not serve God. Despite what the magicians claim to believe, all magic is in fact performed by Satan or his subsidiary demons'.

Such Christian anxieties-anxieties that culminated in the convulsions of iconoclasm and were enacted in Protestant prohibitions against images of the divine-were underpinned, at least in part, by the suspicion that if nature and reality are the realms of God, then artifice and simulation may be the realm of the Devil (e.g. Leyerle, 2001) -a notion echoed in Balduin's sale of his image to Satan.

Commentators from Sigmund Freud (1933)¹⁵ to Arthur C. Clarke¹⁶ (1976) have often alluded to the tendency of advanced technologies to adopt the themes - and one might add replicate the mystique - of magic. If magic is to be understood as the promise of power over things and over others, this is also what information technologies currently promise (e.g. Davis, 1999). Whatever we might mean by 'post-modernity' it is in many ways coterminous with an increased awareness of, and preoccupation with, the performative rather than merely denotative aspects of representation (Hawkes, 2004, Anderson, 1990). It is therefore highly appropriate that the computer has been elevated to the emblematic technology of the present age (Poster, 1990).

How then, to return to Aquinas' question, do we account for this purported ability? What exactly are the processes through which representations, as Taussig (1993) puts it, come to share in or acquire the properties of the represented? In their account of the rise of Tesco to pre-eminence among British supermarkets, Humby et al (2006:139) describe such a process:

"Take each product, and attach to it a series of appropriate attributes, describing what that product implicitly represented to Tesco customers. Then by scoring those attributes for each customer based on their consistent shopping behaviour, [known from their loyalty card accounts] and building those scores into an aggregate measurement per individual Measuring customers on these criteria should start to create distinct profiles ... in shopping, so that it might be possible to identify the busy urban couple who shopped for ready meals but loved to be adventurous and spend a little extra. Or the health-conscious shopper who buys fresh fruit and vegetables, and avoids red meat but sometimes eats chicken. Or indeed any one of the distinct shopping characters we all see in our local supermarket, and think 'I know who

15. Freud (1933) - presumably drawing on Frazer-conjectures 'early men' as attempting to exert control over the world (and their fellows) by means of sympathetic magic: "[Magic] depended for success upon the performance of an action which would cause Nature to imitate it. If he wanted it to rain, he himself poured out water; if he wanted to stimulate the soil to fertility, he offered it a performance of sexual intercourse in the fields".

16. According to Clarke's (1976) 'Third Law' for instance, advanced technologies are 'indistinguishable from magic'.

you are. We recognize a mirror image of ourselves when we look at someone else's basket".

When each profile was isolated, it was "tested against other factors – for example, how often did they typically shop and when? What magazines did they read?" (op.cit:144). The tools employed in the diagnostic process include Osgood's (1964) "Semantic Differential Procedures" which are used for measuring "abstract or judgmental concepts" (Humby et al, 2006:140-2), such as "fast", "fresh", or "adventurous", by means of 20 seven-point Likert scales. "Using the scores for all these attributes, an Osgood profile becomes a map of the connotations for the product or concept" (ibid). "With 20 scales agreed as a way of grading every product" on Tesco's shelves, 45,000 Osgood profiles were constructed, "one for every product from anchovies to asparagus, whisky to washing powder", yielding approximately 1.2 million individual ratings (ibid). Combining these profiles with information such as income, age, marital status e.tc., gained from other sources (such as geo-demographics, loyalty card data, e.tc.), Tesco created "Shopping Habits", "a picture of what attitudes and beliefs drive our [consumer] behaviour" (ibid). The range of "data layers" that might be involved in the construction of pre-dictive knowledge(s) of the consuming subject are summarised in fig. 5 (from Evans, 2005:104). These are the basis for identifying, or perhaps constructing, Hamel and Prahalad's (1996) "unarticulated needs": Customers can be given incentives to purchase the same products that others who share their "Shopping Habits" (say "adventurous", or "finest") have already purchased (an experience familiar from the online world: "We've noticed that customers who have expressed interest in X have also ordered Y"). To paraphrase Taussig (1993), what we have in this account, is the notion of the copy, (the digital double), affecting the original to such a degree that the represented begin to share in or acquire the properties of the representation.

As described above, computer mediated-processes of consumer re-representation are redolent of the echoes of 'magical' practice that Taussig detects in our contemporary "habits of representation". One quick way of summing-up such echoes would be through the concept of "totemism" (e.g. Lien, 1997). Totemism, it will be recalled, describes the establishment of classificatory correspondences between the natural and the social orders. Totem animals and plants, for instance, provide an ostensibly natural basis for the classification of social groups (e.g. Needham, 1978)¹⁷. In Mary Douglas' (1966) analysis of the dietary prohibitions in Leviticus for instance, the separation of clean from unclean (and hence forbidden) animals, reflects a preoccupation with affirming a clear boundary between

17. As Levi-Strauss (1962) put it, the distinction between the classes of "man" and "animal" provides the conceptual basis for social differences.

Jews and Gentiles. It is worth noting that in contemporary anthropological accounts the concept of totemism does not comfortably map on, but instead tends to undermine, any conventional distinction between 'modern' and 'pre-modern' cultures. Indeed, as outlined above, 'totemism' is a fairly accurate description of 'customer relationship management's' preoccupation with re-configuring the world of goods and the world of the consuming subject into mirror images of one another (Lien, 1997).

At the same time, for many commentators CRM and its associated techniques epitomise a quintessentially Foucauldian (post?) 'modern' exercise of power (e.g. Zwick and Dholakia, 2004), a panoptic power which shapes identity and works through the channelling of desire. In their discussion of the Tesco case, Humby et al (2006:146) note that the success of the programme is shadowed by a certain cultural ambivalence. Consumers, they say, "are enthusiastic about participation in the process" but when made aware of the sheer "volume of data that the company running the scheme has collected, they are less comfortable". They quote Simon Davies, founder of the privacy rights group Privacy International:

"to get discounts you have to accept the imposition of a loyalty card and consent to give up personal data. That consent is a fraud, it's a bit like slavery to your supermarket."

To borrow from Anna's wedding ring metaphor, the 'relationship' part in the customer relationship management is according to Davies, more akin to a 'Stepford' marriage. The "supermarket slave" is an oxymoronic creature, the monstrous progeny of the Friedmans' (1976) free-market choice and Deleuze's (1995) control society. Unsurprisingly, Humby et al, having spent most of the book explaining how, with sufficient data, consumer agency can indeed be harnessed in the pursuit of corporate objectives, attempt to banish associations with the 'Stepford' consumer, arguing that,

"[a]ny supermarket that attempted to 'enslave' its customers, rather than reward them, won't be able to run a scheme in the long term because participation ... will dwindle" (ibid).

And yet, given that for many commentators the supermarket represents the paradigmatic institution of post-modernity-like the medieval castle or the industrial age factory, a metaphor for the zeitgeist of an age (Anderson, 1990) - the spectre of the Stepford consumer cannot be so easily exorcised. A caricature though it might be, it is at the same time a reminder of the cultural ambivalence that surround the status of the human subject in informational capitalism.

Concluding remarks

Advocates and critics of informational capitalism tend to share the assumption that digital record keeping is assembling an ever more detailed picture of social life. The electronic record has become, it is implied, a kind of mirror where social events and individual actions are routinely reflected. There is by now an extensive body of work, within sociology, cultural theory, the new (sub)discipline of surveillance studies and beyond, which endeavours to highlight and explain this process and its social, moral and organizational consequences. We have in this paper sought to trace out some of the new twists that the continuing diffusion of CRM-related technologies add to these ongoing debates. We have seen, for instance, how such technologies involve more than just the accurate recording of the activities of organizational actors (including consumers) – the standard concern of surveillance studies. Rather, systems of this kind appear increasingly involved in the making of ourselves into the kind of people we are, in fabricating the kinds of identities we might assume.

Heidegger (1977), famously, argued that the relationship between knower and known in occidental modernity—whether the object of knowledge is the natural or the social world—increasingly becomes a relationship of appropriation. He therefore describes the distinctiveness of modern technology in terms of ‘enframent’, as the process through which the world and everything in it, human beings included, is harnessed as a repository of resources, a ‘standing reserve’ or ‘stock’ (Bestand) (Heidegger, 1977). The mighty Rhine, for example, is ‘enframed’ by the power plant and ‘put to work’ as a reliable source of hydroelectric power. Against this backdrop, supermarket shelves in their transcendence of seasons and geography exemplify no less than the Rhine dam, a world remade as a standing reserve. Correspondingly, the consuming subject itself is, as we have seen, electronically re-presented in terms of the goods on those shelves. As Humby et al (2006:122) put it, ‘You Are What You Eat’.

Insofar as contemporary computer-mediated forms of surveillance and digital inscription also involve the attempted technological conquest of *tyche*, they can be said to create their own versions of a digital ‘standing reserve’ (see Knox et al., 2008). They indulge the desire to tame the flux of social life, to ‘enframe’, so to speak, the Heraclitian river (Chia and Tsoukas, 2002). They thus instantiate what MacKenzie (1998) calls ‘the historically recent and still growing technological channelling of events into networks of inscriptions or marks mobilised through algorithmic or programmed devices’ Events are, as it were dis-assembled and re-assembled “by passing through the bottleneck of coding”, and thus subjected to the rules, practices and procedures on which coding techniques depend (Kallinikos, 2009:189). The creation of ‘digital reserve’ is thus the condition of exist-

ence for the various wonders of 'information age' social organization (see Figure 5). Yet we find that these same processes render the orderings they produce increasingly vulnerable to disruptions and translation errors as the 'real world' gets processed through the medium of its digital mirror. For Heidegger it will be recalled, the processes of ordering the natural and social worlds are processes punctuated by technological breakdown, followed by technological fix, further breakdown and further fix. Tyche is therefore not eliminated, but is reinvented in ways that continue to haunt proceedings. We argue that the 'digital reserve' is generative of its own forms of instability, its own forms of flux and disorder. The 'digital reserve' could be said to go hand-in-hand with a 'digital uncanny' (Waldby, 1997).

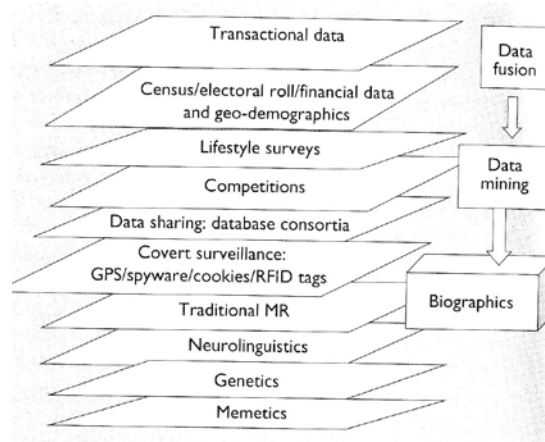


Figure 5: A digital (standing) reserve? Data layers (Source: Evans, 2005).

In the traffic between the digital and real, the thing and its reflection, all sorts of strange creatures seem to emerge. For instance, on March 6th 2001 there was alarm at Derby Magistrates Court when Brendan Michael Forrester of Kingsway Place Swadlincote failed to appear to face charges which inter alia included obstructing a police officer in the course of his duties, various drug offences, living off immoral earnings, causing death by dangerous driving and being in possession of a rocket launcher. Panic ensued at the thought that such a dangerous individual was at large in Derbyshire. When police inquiries failed to produce Forrester or even locate 'Kingsway Place' where he supposedly resided, the conclusion was reached that he did not actually exist '[e]xcept on the computer', that 'Forrester' might in fact be 'a police training exercise that went too far' (Seddon, 2001:312-3) or the creation of a bored court clerk with too much time on his hands.

More seriously, ever since May 23 1993 German police had been hunting the (woman) murderer of churchwarden Lieselotte Schlenger. When the killer's DNA was fed into the Interpol database she was immediately linked to a bizarre variety of crimes, from a heroin-filled syringe discovered in a German wood, to the strangulation of a pensioner, to the gangland style execution of three Georgian car dealers found dumped in the Rhine, and most notorious of all, the 2007 shooting of policewoman Michele Kiesewetter. In all, her DNA turned up at over thirty crime scenes in Germany, Austria and France making her one of Europe's most prolific criminals, a veritable female Moriarty. At the time of writing, and after a hunt lasting 16 years, the police came to suspect that she did not actually exist and that the DNA found at the various crime scenes and circulated in the DNA databases actually belongs to the woman who packed the cotton swabs used to collect it. In terms of the concerns of this paper such incidents provide a reminder of the (Baudrilardian) status of the 'digital double' as a presence in its own right and not necessarily a reflection cast by an 'original'.

The inscriptive machinery in *Minority Report*, which might be taken to portray the ultimate in attempts to overcome the unpredictability of human agency ('enframing' taken to its logical conclusion) still appear endemically prone to instability and representational excess. Anderton, whose job is to enforce the system's conclusion, and whose faith in it is – to start with – unquestioned, soon comes to realise that the facts which the system records are not merely re-presentations but performances of reality. Thus the crime he is projected to commit is being orchestrated by person or persons unknown while the crime scene itself, which the system has ostensibly accurately recorded, is in fact 'a set-up purposely designed to enrage him' into committing the pre-dicted crime (see Friedman, 2006).

The still in-formation 'digital reserve' is, we propose, not unlike the mirror described in Jorge Luis Borges' (1990) *Book of Imaginary Beings*. In it, Borges tells the tale of the fish, an ever-shifting 'shining creature that nobody had ever caught' but which the people of Canton say can sometimes be fleetingly glimpsed swimming in the depths of mirrors. Once upon a time, the tale goes, the "world of mirrors and the world of men were not, as they are now, cut off from each other". Rather mirrors were, so to speak, portals between the specular world and our own. It was only the magic arts of the Yellow Emperor who in order to avert an invasion by the mirror creatures, stripped the specular world of its autonomy and made it into a slavish reflection of ourselves. The day however is bound to come when the Emperor's spell will finally fail and the mirror creatures will begin to stir: "Little by little they will differ from us; little by little they will not imitate us". The first sign that the spell has failed, the first creature to be sighted, will be the fish moving in the mirror depths, presaging the onslaught of the mirror creatures whom the flimsy glass barriers will no longer be able to contain. Perhaps

our age, the age of the digital reserve is indeed the age when the spell has began to fail.

'Doubles' and 'originals', we have intimated, both mimic and destabilize one another. After all, even Deleuze (1995) did not claim that a 'control society' would actually function as intended. Once more, it is hardly surprising to discover that the corporate systems and techniques that seek to, as it were, 'iron out' the unpredictability of human agency, are themselves not immune to subversion, inversion and drift (e.g. Ciborra, 2002). It is because of, rather than in spite of this, that 'Customer Relationship Management' provides us with a useful vantage point from which to observe how our (post-modern?) preoccupations regarding the performativity of the electronic sign are enacted in contemporary institutional settings.

REFERENCES

Abbott, Julie (2001) 'Data data everywhere – and not a byte of use?', *Qualitative Market Research*, 4(3): 182-92.

Albrecht, Katherine & McIntyre, Liz (2005) *Spychips: how major corporations and government plan to track your every move with RFID*, Nelson Current, Nashville, TN

_____ (2006) *The Sychips Threat: Why Christians Should Resist RFID and Electronic Surveillance*, Nelson Current, Nashville, TN

Anderson, Walter T. (1990) *Reality is Not What It Used to Be*, HarperSanFrancisco.

Arendt, Hannah (1965) *Eichmann in Jerusalem*, New York: Viking.

Aunger, Robert (2001) *Darwinizing Culture: The Status of Memetics as a Science*, Oxford: Oxford University Press.

Bainbridge, Simon (2006) 'Lord Ruthven's Power', *The Byron Journal*, 34/1:21-34.

Ball, Kirstie (2005) 'Organization, Surveillance and the Body: Towards a Politics of Resistance', *Organization*, 12(1):89-108.

_____ and **Wilson, D.** (2000) 'Power, Control and Computer-based Performance Monitoring: Repertoires, Resistance and Subjectivities', *Organization Studies* 21(3):539-65.

Banham, Russ (2003) 'CRM, As You Like It', CFO.com. Available at <http://www.cfo.com/article.cfm/3009214?f>

Bauman, Zygmunt (1991). *Modernity and Ambivalence*. Cambridge: Polity Press.

Benjamin, Walter (1979) 'The Work of Art in the Age of Mechanical Reproduction' in *Illuminations*, London: Fontana.

Blackmore, Susan (1999) *The Meme Machine*, Oxford: Oxford University Press.

Bogard, William (1996) *The Simulation of Surveillance: Hypercontrol in Telematic Societies*, Cambridge University Press.

Boland, Richard (1986) 'Fantasies of Information', in *Advances in Public Interest Accounting*, Vol.1, London: JAI Press: 49-64.

Borges, Jorge L. (1990) 'The Fauna of Mirrors' in *The Book of Imaginary Beings*, (trans.) M. Guerrero, and N. Di Giovanni, Harmondsworth: Penguin.

Chia, Robert and Tsoukas, Haridimos (2002) 'Everything Flows and Nothing Abides: Towards a Rhizomic Model of Organizational Change, Transformation and Action', *Process Studies*, 32(2):196-224.

Ciborra, Claudio (2002) *The Labyrinths of Information*, Oxford, Oxford University Press.

Clarke, Arthur C. (1976) *Profiles of the Future*, Athens: Cactus (Greek translation.).

Czarniawska Barbara & Joerges, Berward (1998) 'The Question of Technology, or How Organizations Inscribe the World', *Organization Studies*, 19(3): 363-85.

Davis, Eric (1999) *TechGnosis: Myth, Magic, and Mysticism in the Age of Information*, London: Serpent's Tail.

Dawkins, Richard (1989) *The Selfish Gene*, Oxford: Oxford Paperbacks.

Deleuze, Gilles (1995) 'Postscript on Control Societies.' in *Negotiations 1972-1990*, Martin Joughin (trans.), New York: Columbia University Press.

Dick, Philip K. (2000) *Minority Report; The Collected Short Stories of Philip K. Dick*, London: Gollancz.

Dostoevsky, Fyodor (1985 [orig. 1845/1866]) *The Double; Two Versions*, (trans. Evelyn Harden), Ann Arbor: Ardis.

Douglas, Mary (1966) *Purity and Danger*, London: Routledge.

The Economist (1999) 'The End of Privacy: The Surveillance Society', May 1st: 19-23

Elmer, Greg (2000) 'The Politics of Computer Profiling', *EASST Review*, Vol. 19(1), March.

Evans, Martin (2005) 'The Data-Informed Marketing model and Its Social Responsibility' in *Lace S, (ed.) (2005) The Glass Consumer: Life in a Surveillance Society*, Bristol: NCC/The Policy Press: 99-132.

Fickle, L. (1999) 'Know your customer', *CIO Magazine*, Vol.12(21):62-72.

Frazer, James G. (1994 [orig. 1890]) *The Golden Bough: A Study in Magic and Religion*, London: Touchstone.

Freud, Sigmund (1933 orig. 1932) 'Lecture XXXV', in *New Introductory Lectures on Psychoanalysis*, London: Hogarth Press.

Friedman, Lester (2006) *Citizen Spielberg*, Chicago: University of Illinois Press.

Friedman, Milton & Friedman Rose (1990 orig. 1976) *Free to Choose*, London: Harvester.

Foucault, Michel (1977) *Discipline and Punish: The Birth of the Prison*, (trans.) A. Sheridan, Harmondsworth: Allen Lane.

Gabriel, Yiannis & Lang, Tim (2006) *The Unmanageable Consumer*, London: Sage.

Giddens, Anthony (1994a) 'Living in a Post-Traditional Society', in U. Beck, A. and S. Lash, (eds.) *Reflexive Modernization*, Cambridge: Polity:55-109.

_____ (1994b) *Beyond Left and Right*, Cambridge: Polity.

Hamel, Garry and Prahalad, C.K. (1996) *Competing for the Future*, Boston MASS: Harvard Business School Press.

Hawkes, David (2004) 'Faust Among the Witches', *Early Modern Culture*, 4.

Hayek, Friedrich (1949) *Individualism and Economic Order*, London: RKP.

Heidegger, Martin, (1977) *The Question Concerning Technology and Other Essays* New York: Harper and Row.

Herdman, John (1990) *The Double in Nineteenth Century Fiction*, London: Macmillan.

Holstein, J.A. & J. Gubrium (1997) 'Active Interviewing', in D. Silverman (ed.) *Qualitative Research; Theory, Method and Practice*. London: Sage.

Humby, Clive, Hunt, Terry & Phillips, Tim (2006) *Scoring Points: How Tesco Continues to Win Customer Loyalty*, London: Kogan Page.

Johnson, Simon (2008) 'Eight million at risk of ID fraud after hackers hit hotel chain', *The Daily Telegraph*, August 25th:1.

Kallinikos, Jannis (2006) *The Consequences of Information*, Cheltenham: Edward Elgar.

_____ (2009) 'On the Computational Rendition of Reality: Artefacts and Human Agency', *Organization*, 16(2):183-202

Knights, David & Morgan, Glenn (1993) 'Organization Theory and Consumption in a Post-Modern Era', *Organization Studies* 14(2):211-34.

_____ **Noble, Faith, Vurdubakis, Theo, & Willmott Hugh** (2001) 'Chasing Shadows: Control, Virtuality and the Production of Trust', *Organization Studies* 22(2):311-36.

Lace Suzanne (ed.) (2005) *The Glass Consumer: Life in a Surveillance Society*, Bristol: NCC/The Policy Press.

Latour, Bruno (1987) *Science in Action*. Milton Keynes: Open University Press.

Lee, Benjamin & LiPuma, Edward (2002) 'Cultures of Circulation', *Public Culture*, 14(1): 191-213.

Levi-Strauss, Claude (1962). *Totemism*, Chicago: University of Chicago Press.

Leyerle, Blake (2001) *Theatrical Shows and Ascetic Lives*, University of California Press.

Lien, Marianne (1997) *Marketing and Modernity*, Oxford: Berg.

MacKenzie, Adrian (1998) 'The Future Left to Its Own Devices', *Culture Machine*.

Miller, Peter and O'Leary, Ted (1987) 'Accounting and the Construction of the Governable Person Accounting Organizations and Society', 12(3) 235-265.

Needham. Rodney (ed.) (1978) *Primordial Characters*, Charlottesville: Virginia University Press.

Nichols, Joshua (2004) 'Data Doubles: Surveillance of Subjects Without Substance' *CTheory*. Available at www.ctheory.net/articles.aspx?id=410

Ody, Penelope (2003) 'Channel hoppers pick' n' mix in the aisles', *Financial Times IT Review*, October 1:1-2.

Osgood, Charles E. (1964) 'Semantic Differential Technique in the Comparative Study of Cultures', *American Anthropologist* 66/3:171-200.

Packard, Vance (1970) *The Hidden Persuaders*, Harmondsworth: Penguin.

Pagels, Elaine (1996) *The Origin of Satan*, London: Allen Lane.

Parish, Jane & Parker, Martin (2001) *The Age of Anxiety: Conspiracy Theory and the Human Sciences*, London: Blackwell.

Peirce, Charles S. (1931-58) *Collected Papers I-VIII*, Hartshorne, C, Weiss, P, and Burks, A, (eds.). Cambridge, MA: Harvard University Press.

Poster, Mark (1990) *The Mode of Information: Poststructuralism and Social Context*, Chicago: University Of Chicago Press.

Privacy Rights Clearinghouse (2000) 'Oral Testimony of Michelle Brown' to the U.S. Senate Committee Hearing on the Judiciary Subcommittee on Technology, Terrorism and Government Information, 'Identity Theft: How to Protect and Restore Your Good Name': July 12. Available at <http://www.privacyrights.org/cases/victim9.htm>

RFID Journal (2003a) 'Benetton to tag 15 Million Items': March 12. Available at <http://www.rfidjournal.com/article/articleview/344/1/1/>

_____ (2003b) 'A Setback for RFID? Some press reports suggest that the Benetton situation could hurt the RFID industry. Nothing could be further from the truth': April 14. Available at <http://www.rfidjournal.com/article/articleview/382/1/1/>

_____ (2003c) 'Behind the Benetton Brouhaha: Despite reports to the contrary, the Italian retailer continues to explore the supply chain benefits of RFID': April 14. Available at <http://www.rfidjournal.com/article/view/381/1/1/>

Rigby, Darrell, Reichheld, Frederick, & Schefter, Phil (2002) 'Avoid the four perils of CRM', *Harvard Business Review*, 80 (2): 101-9..

Rose, D. (1990) 'Quixote's Library and Pragmatic Discourse: Towards Understanding the Culture of Capitalism', *Anthropology Quarterly*, 63(4): 155-68.

Seddon, Peter (2001) *The Law's Strangest Cases*, London: Robson.

Serres, Michel (1995) *Angels: A Modern Myth*. Paris: Flammarion.

Silverman, David (1993) *Interpreting Qualitative Data: Methods for Analysing Talk, Text and Interaction*. London: Sage.

Taussig, Michael (1993) *Mimesis and Alterity*, London: Routledge.

Thompson, James. D. (1967) *Organizations in Action*, New York: McGraw Hill.

Waldby, Catherine (1997) 'Revenants: The Visible Human Project and the Digital Uncanny', *Body and Society* 3(1):1 - 16.

Westerman, Paul (2004) Data Warehousing: Using the Wal-Mart Model. Morgan Kaufmann.

Wilde, Oscar (1985) The Picture of Dorian Gray, Harmondsworth: Penguin.

Zwick, Detlev, & Dholakia, Nikhilesh (2004) 'Whose Identity Is It Anyway? Consumer Representation in the Age of Database Management', Journal of Macromarketing, 24/1:31-43.

A brave new world for libraries: the sui generis right

Rania Konsta*

Ionian University, Corfu, Greece

Michalis Gerolimos**

Ionian University, Corfu, Greece

The Sui Generis Right

The ways in which databases¹ affect, within a modern information environment, scientific communication, the exchange of ideas among peers, as well as the search and retrieval of information by any potential user through a commercial and financial framework that is still being shaped, has led the European Union the creation of the sui generis right (special right), through which it is the first time ever to establish a copyright on databases (EU Directive 96/9/EC). What is exceptional about this development, which is a worldwide legal breakthrough, is the fact that this new special right protects databases regardless of the legal framework governing intellectual property. In essence, it constitutes a protection effort through an organised ensemble aiming at safeguarding private investment in databases. Thus, this framework rather than being related to the selection and arrangement of database content, it aims at protecting investors from the parasitic use of this content by competitors and simple users alike (Colston, 2001).

Therefore, Directive 96/9/EC raises new issues concerning the integration of database protection both in the traditional framework of intellectual property, as

* *Rania Konsta* is head Librarian at Ionian University Library. She has a MSc in Information Science and she is candidate Phd.in Library and Information Science.

** *Michalis Gerolimos* holds a PhD in Library and Information Science and he is a graduate of Department of Archives and Library Science, Ionian University, Greece. He has a professional experience in Library Science for over 8 years as an Academic Librarian. He has, also, been a Teaching Assistant and Research Assistant in Department of Archives and Library Science for several semesters.

1. The term 'database' includes both electronic and non-electronic ones. Therefore, a database is defined as the collection of independent works, data or other materials which are systematically or methodically arranged and can be individually accessed using electronic or any other means (Directive 96/9/EC).

well as in the overall law of immaterial goods. It is the first time in the history of immaterial goods that “bare” data and facts (information) are protected from being retrieved and exploited, by a regulation that is characterised by ingenuity, originality, individuality, discreet force, e.tc. This protection is essentially provided by avoiding the risk of third parties –other than its author- of acquiring and exploiting database material. A key requirement in order to recognise the sui generis right for a database author are both the investment’s form and nature.

According to the Directive, the right in question pertains to each database in which a “substantial” investment exists², regardless of its originality, of whether or not it constitutes an actual intellectual creation of its author, e.tc. Moreover, it should be noted that the sui generis right does not impose any restrictions as to the aim for which the database was created, i.e. it is not required to correspond to an investment’s scope, like –for instance- being profitable. As a consequence, the sui generis right does not fall under unfair competition stipulations, as it is considered an intellectual property right and it thus protects the databases from the moment the latter is created, regardless of its function (e.g. commercial) or the aim (e.g. profit) for which it has been created.

Sui generis right protection covers databases for which content obtaining, verification or presentation signify a substantial or quantitative investment (Article 7). However, the quantitative and qualitative criteria that could in fact, even typically, define a “substantial investment”, are not mentioned in the Directive, resulting in significant interpretation and application problems, given that “substantiality” is a relative term, that depends on subjective factors. Moreover, investments may have to do with the financial capital, the time allocated by the investor or the database creation effort as a whole. This definitional vagueness in the section in question is subject to interpretation and may, ultimately, be considered similar to the one pertaining to “sweat of the brow” right for the beneficiary of a database.

The Directive also provides for certain restrictions of the sui generis right. The lawful user of a database (a CD-ROM buyer, or a subscriber in the case of an on-line connection) may, without prior permission by or payment to the author, extract or reutilise for any reason whatsoever “insubstantial” parts of the database content, while any contrary contractual provision between the user and the owner is considered null and void. It is a minimum right of the lawful user to be able to use in their personal computer insubstantial parts of the database content, as well as the accompanying software. It is, therefore, reasonable to assume that

2. The investment is a key point of protection: it constitutes (theoretically) a condition for recognising a database author’s sui generis right. The right in question protects those databases, the obtaining, verification or presentation of the contents of which, require a substantial investment.

even those who do not have any legal rights on the database, for instance as a result of a use concession contract or purchase, to be able to extract and reutilise insubstantial parts of the database content. However, banning the systematic and repeated extraction of insubstantial parts of the database content, which damages the author's rights or conflicts with the proper exploitation of the database, would mean that users such as scientists, librarians or journalists –to name but a few- whose work makes it imperative that they access database information, may have problems legalising the extraction and reutilisation of database content.

Extraction and/or reutilisation of "insubstantial" parts of the database content is prohibited, when such actions are being repeated and assume a systematic character, resulting in conflict with normal exploitation of the database or unreasonable prejudice of the legitimate interests of the maker of the database (Article 8, paragraph 2). The issue with the provision in question is that the reference to temporary extraction also means that even temporary digital copies constitute an unauthorised extraction and, therefore, simply reading on a personal computer monitor a substantial part of the database constitutes temporary extraction and can be hence forbidden.

As in the case of "substantial investment", where the lack of term definition creates interpretation problems, there are no guidelines explaining the notion of the "substantial" part of a database. Moreover, the substantial or insubstantial character of the extraction/reutilisation³ may be judged either quantitatively or qualitatively, an option that, in all likelihood, protects the database owner rather than the user. According to Article 7, paragraph 2, public lending does not constitute an act of extraction or reutilisation and, as a consequence, a library allowing the public to access the database, over a limited period of time and provided that no financial or commercial gain is pursued by said use licence, is free to do so. Article 9 of the same European Commission Directive does not include some of the users' rights on authors' works which are, by now, recognised and established by other provisions, like –for instance- the possibility to cite passages, the compulsory exception in cases of teaching and scientific research, as well as reproduction by archives and libraries. It is possible that the authors of the Directive in question considered that all relevant rights can be well-protected by the provision regarding the free extraction and reproduction of part of the database which is considered insubstantial.

3. 'Reutilisation' is defined as any form of making available to the public all or a substantial part of the contents of a database by the distribution of copies, by renting, by on-line or other forms of transmission. The database author's right to reutilisation corresponds to the creator's right to distribution, according to copyright law.

To the above data, one should add that Article 9 paragraph 1 of the Directive stipulates that it is prohibited to reproduce a digital database for private use, with the exception of the right of Member States to limit the sui generis right of the author of a digital or other database, provided that the extraction is carried out within the framework of an administrative or judicial procedure or for the purposes of illustration for teaching or scientific research. It is only in the case of this exception, within the framework of an administrative or judicial procedure that the extraction and reutilisation of data is allowed. On the contrary, should the exception to be established by the Member State be related to education or research, only the extraction of data is allowed and not its reutilisation, while Member States should ensure that the source of the data in question is appropriately cited. A further limitation in the use of a database is the fact that the aim of said extraction must be non-commercial and that said extraction does not exceed a certain degree that justifies meeting its non-commercial aim. However, since it is of vital importance to scientists not only to extract data and information from a database, but also to be able to process and reutilise them (i.e. to be entitled to publish and share these data with the scientific community), this particular exception is –in essence– an additional restriction imposed on the use of a database.

The sui generis right of a database author expires fifteen years from the first of January of the year following the date of completion, while this deadline is to be renewed following any amendment of, addition to, deletion from the database, e.tc. This is totally contrary to what was in force until today, as even the intellectual property rights of original works have an expiration date. Although it is commercially, scientifically and even socially effective and necessary to renew the content of a database, this, however, gives its author the right to claim that he/she is constantly renewing the database content in question, and, as a result, to constantly maintain this content under protection. As a final point, the result of an incessant protection is that the database content will never come to the public's full and unrestricted possession, even if the data and the information contained therein remain available for over fifteen years. In essence, this problem is created because there was never any provision for compulsory database use licenses, resulting in running the risk of allowing the creation of monopolies in information production, retrieval and utilisation.

Reactions by the academic community and libraries

A typical research requires availability and use of a relatively large amount of information, while in certain cases of specific and specialised research, it is imperative to use systematically numerous databases (like, for instance, in the case of a research focusing on global warming). Should databases that are currently freely available to the public, be placed under the protection of the *sui generis* right, then, inevitably, the cost of such a typical research would definitely rise. Moreover, the very culture governing Scientific Communication is possibly also going to change, given that the process of ideas exchange and common use of data among institutions will be modified, as the institutions themselves will start considering their databases as commercially exploitable sources of income. Of course, such a development would not leave the business world and state services unaffected, since it is highly likely that the ensuing developments –i.e. the increase in research cost- will affect both the current and future potential of carrying out research that will include the greatest amount possible of available information, while at the same time making it available to a wide group of interested parties. Aiming at preventing and dealing with such an eventuality, many associations of scientists, professionals and teachers, like, for example the National Research Council, the National Academy of Sciences, the National Academy of Medicine and the National Academy of Engineering, are strongly opposing the *sui generis* right as well as the protection that it enforces in favour of database authors-owners. Similar reactions have also been noted on behalf of many library associations against HR 3531 and its pertinent provisions regarding databases (Band and Gowdy, 1997).

The establishment of the *sui generis* right has caused intense reactions both in the American academic community and the database authors themselves. More precisely, the academic community on the one hand opposes the *sui generis* right, claiming that it is bound to have catastrophic repercussions on research and scientific development, as a result of the expected long-term information monopoly; database authors, on the other hand, are trying to achieve their fullest possible protection (through an absolute and exclusive right) so as to safeguard their investment. Although the European *sui generis* right has led to the genesis of at least two bills in the America on the legal protection of databases, it is worth mentioning that such a right has not as yet been established by law in the United States, a country in which intellectual creation investments are protected par excellence and where the database market is flourishing.

IFLA and sui generis right

Particularly as pertains to libraries and their position regarding the sui generis right, the International Federation of Library Associations and Institutions (IFLA) is attempting –by constantly intervening– to contribute to the creation of a protection framework which, however, will not limit the utilisation conditions of data and information contained in databases, aiming –of course– at a more substantial and effective use of individual library policies. What seems to interest the IFLA the most, as well as other associations that are generally aiming at safeguarding the public's access capability to information through databases, are the serious repercussions that science, education, research and innovation are likely to suffer, should provisions such as those stipulated by the sui generis right ultimately prevail, constituting the legal form of database protection at a global level.

It is equally possible that any debate regarding the establishment of new provisions having to do with database legal protection, will be coming from those who have the most to gain from a readjustment of the existing legislation. Consequently, one may actually go as far as to assume that there is an effort to create an artificial need, the world over, to renegotiate the regime governing database protection, with the aim –of course– for database authors/manufacturers to make an even bigger profit.

IFLA, through committees created ad hoc (such as the Standing Committee on Copyright and Related Rights), raises additional issues that need to be taken into serious consideration, in case there is indeed a change in the provisions regarding the access and utilisation of databases. So, they consider most significant the language in which any such provision will be expressed, as well as the fact that any terms used therein, must be carefully selected so as not to over-protect databases through these new provision. In that sense, terms such as “substantial part”, “insubstantial part” and “substantial investment”, should acquire a very carefully chosen meaning so as to ensure that they will not give rise to different interpretations when the pertinent provisions are applied on a nation-wide level.

In this effort to create a unified front that will deal with issues of copyright of the works contained in databases, it is of the utmost importance not to require a contract between the two transacting parties on the basis of the sui generis right, when there is already legislation in force that provides a corresponding protection to databases. All this, of course, must be in relation to the new standards and models, whenever these may be defined by international organisations and pertinent meetings of experts.

Quite recently, the IFLA responded to the European Commission's Green Paper on Copyright in the Knowledge Economy (COM (2008) 466/3). The main points of its views concerning the rights of authors as well as overall issues concerning database protection may be summarised as follows:

- Authors' rights constitute one of the main pylons for the creation of a regime governing intellectual property rights. However, restrictions and exceptions related to these rights are equally important. The constant changes in the law on copyright are what created the current imbalance between creator rights and user rights, as they have upgraded authors' rights without adequately providing for restrictions and limitations applicable to said rights.
- Authors' rights have been enhanced both in terms of duration and when supported by technical means (as is often the case in digital environments), which means that they are applicable without any significant exceptions. Technical means may not only limit or even eliminate legal exceptions to utilisation but they are in themselves virtually "impervious" to any legal application.
- Contrary to what happened in the case of the Database Directive (Directive 96/9/EC), in the Information Society Directive (Directive 2001/29/EC) there is no provision to amend the contract so as to protect the users. Suppliers' contracts, in case they opt for them to be non-negotiable, and since intellectual property rights constitute –in essence– exclusive rights, may indeed create a monopoly.
- User licence negotiation for compensation in cases of legal exception should not be undertaken by the interested parties, i.e. libraries. It is best avoided, as the "power" of the contracting parties is unequal, as a result of the additional power offered to the copyright owner by the law. A typical example thereof are the various international publishers of scientific journals and books, who –in essence– may indicate to libraries the way in which these documents shall be used in their collections. As a matter of fact, more often than not, such terms in utilisation contracts prevail over any potential legal stipulations providing for utilisation exceptions, e.g. work reproduction for personal use.
- The monopoly power in the hands of those who hold intellectual property rights is a relatively recent development. Back in the days of printed information and document distribution it was –to all intents and purposes– impossible for a publisher to control and/or prevent the creation of copies. However, this has changed in the digital world, as copyright owners have digital document supply contracts that allow them to deny a user's rights to utilisation exceptions. In order to avoid such an exploitation of the monopoly use of intellectual property rights, there must be an overall provision by the legislation governing copyrights, which shall not allow the imposition of terms contrary to the right to exceptions in the utilisation and limitation of intellectual property rights. For instance, Article 15 of the Database Directive, stipulates that it is mainly the legislator's responsibility to provide exceptions in the

utilisation and restrictions in the rights of authors and other owners of copyrights, aiming at ensuring that the needs of society in relation to research, science and education are adequately met.

- IFLA also proposes that the Directive 96/9/EC exception pertaining to the legal protection of databases and their use by people with disabilities, as amended by 2001/29/EC Article 5.3.b, must be mandatory and applicable both for original databases and those protected by the *sui generis* right.
- Finally, as regards the exchange and distribution of works for teaching and research reasons, perhaps it would be advisable to re-examine the decision of the scientific and academic community to participate in schemes aimed at negotiating licences with publishers. Academic institution libraries have been among the key participants of such schemes, aiming at safeguarding their users' access to information. However, research –just like higher education teaching and learning– becomes increasingly international and local negotiation schemes may prove ineffective in this new environment. It is, therefore, significant for intellectual property rights exceptions to include this new environment in question, as well as all new teaching and research methods. In the large majority of cases –with the exception, of course, of the digital documents supply itself– user licenses should be made obsolete. The Information Society Directive Article 5 (3) (a), must be considered adequate, provided it stipulates exceptions “for the exclusive aim of educational or scientific research purposes...”.

IFLA responded to the European Union proposal projecting two essential principles:

- Fundamental rights of expression and information retrieval
- Internal market efficient operation

These two principles are elemental in the effort to acquire the largest possible benefits from the economy of knowledge. A successful intellectual rights regime must, of course, take into consideration authors' rights, but it must also facilitate accordingly other considerable participants in the procedure, like for instance secondary creators, educators, and researchers, all of whom are basing themselves on copyright exceptions in order to create their own intellectual works⁴.

4. International Federation of Library Associations and Institutions: Response from IFLA to the European Commission's Green Paper 'Copyright in the Knowledge Economy' [COM (2008), 466/3] http://209.85.129.132/custom?q=cache:GS1R0XlkdnEJ:www.ifla.org/III/clm/pl/IFLA_Response_Green-paper-copyright.pdf+sui+generis&cd=7&hl=en&ct=clnk.

ALA and sui generis right

In 2006, the American Library Association (ALA) responded to European Union policy on the sui generis right and, more particularly, voiced its concerns regarding the 1996 EU Database Directive (96/9/EC), in an attempt to convince the EU to withdraw all pertinent provisions. So, according to the Resolution in opposition to sui generis database protection (CD #20.6, January 25, 2006) ALA urges the European Commission either to repeal its Database Directive or to withdraw the sui generis right while maintaining copyright protection for “original” databases. In the same resolution it is stated that the European Commission itself, in December 2005 concluded, among other things, that:

- There is no evidence that the Database Directive has achieved its goal of stimulating the production of databases in Europe;
- The sui generis right for database protection has given rise to legal uncertainty and to significant litigation in European courts and the courts of its Member States;
- The sui generis right for database protection may harm legitimate business, research and education activities and threaten the fair use of information, including information in the public domain.

According to the ALA⁵, the sui generis right gave a new, unprecedented opportunity for database protection, even if they are not sufficiently original to be copyrighted. It also stresses that many databases, which consist of individual pieces of information that have been organised in a single collection so that the data are easier to access – are protected under copyright law because of the creative way that the information in them is selected, coordinated and arranged. However, under traditional copyright law, basic factual information is in the public domain and is not entitled to copyright protection. That means that databases that do not have a creative or original element – such as phone book white pages – are not protected under US copyright law.

In the years that passed since the European Commission issued the Database Directive, large database producers and publishing houses have attempted to persuade the US Congress to pass a similar law of database protection. In response to these efforts, the American libraries have been among the first organisations to react and fight against all attempts to change the legal framework concerning database protection. Such protection would reverse the fundamental US information policy that facts are not creative in nature and, therefore, cannot be owned. ALA continues to

5. American Library Association: Resolution in opposition to “sui generis” database protection. Available at: www.ala.org/ala/aboutala/offices/wo/referenceab/colresolutions/012506-CD20.6.pdf.

insist that any database protection bill must allow “fair use” of databases comparable to that under copyright law and permit downstream, transformative use of facts and government-produced data contained in a database⁶.

Conclusions

Immoderate exclusive rights in information constitute an offence to fundamental constitutional freedoms; restrict people’s information environment; impede the function of democratic institutions; and ultimately block the very creation of future databases. If there is indeed a need for a provision of some sort, which would protect authors’ rights (which, incidentally, are not really harmed, as proven by the flourishing database market in the US and the rest of the world), it should under no circumstances grant exclusive rights to authors (Benkler, 2000). Similarly, scientists, researchers and educators should be allowed to use database in the same manner they have been using all collections of works until today. When there is overprotection of the data contained in databases, as in the case of the *sui generis* right that suppresses the three main components of intellectual property (originality, finite duration and exceptions for scientific research and teaching), then the additional protection in question is reduced to a disproportionate restriction of the freedom of expression (Torremans, 2004).

In essence, what the *sui generis* right does is to create obstacles for people’s research, educational and scientific activity. Instead of that –and, naturally, any other similar effort whose sole aim would be to increase database protection and, as a result, increase database owners’ rights to the detriment of users’ rights–, there should be a collective effort aimed at establishing a “balanced copyright”. Creators and intermediaries should benefit from their works, keeping in mind that those who buy and use creations also have rights. Balanced copyright cites the Constitution in granting limited terms for the copyright monopoly – perhaps the 14 or 28 years that sufficed in the United States in the past, may be a longer plausible limit. At some point, works should enter the public domain to encourage the progress of science and the useful arts. Balanced copyright means people and institutions should be able to use their purchased copies of mass-produced works pretty much as they please: copying for personal use or preservation, lending to others, excerpting for use within other works. We should be able to copy text and images from e-journals and books to use in reports and new creations. And libraries should be able to preserve born-digital materials, which frequently mean bypassing copy protection and digital-rights management (Crawford, 2007).

6. American Library Association: Database protection legislation. Available at: <http://www.ala.org/ala/aboutala/offices/wo/woissues/copyrightb/federallegislation/dbprotection/databaseprotection.cfm>.

Even further efforts are in order so as to re-establish a balance in copyrights, like in the case of the Digital Media Consumers' Rights Act, which would allow copy-protection circumvention for fair use or research purposes and the Public Domain Enhancement Act, which would make it easier to find rights holders for older materials. However, what we should always keep in mind is that libraries need intellectual property rights, and by that we mean intellectual property rights that guarantee equal rights for database creators, owners and users; otherwise, their capacity to preserve documents and lend them is gravely jeopardised.

No business or organisation can afford to ignore the issue of *sui generis* database protection. Depending on where a company falls in the data food chain -- and almost every business is somewhere in the food chain -- *sui generis* protection will either add to the bottom line or take away from it. Accordingly, companies would be well advised to study last year's HR 3531, and new legislation if introduced, so that they can determine their position on this controversial issue and act to support or oppose it.

While it is still unclear whether the Courts will revert to granting copyright protection under a "sweat of the brow" standard, it is certain that the frequency of these cases is on the increase. As unusual compilations of mundane information become more valuable to marketing firms and consumers alike in this information age, what was "original" a decade ago has become essential today. Regardless of the direction the courts and legislature choose, until the decision is clearly codified, the information economy is wise to combine innovation with caution, spending as many resources protecting their creations personally as they spend developing them.

REFERENCES

ALA (2006). Database protection legislation, available at: <http://www.ala.org/ala/aboutala/offices/wo/woissues/copyrightb/federallegislation/dbprotection/databaseprotection.cfm>.

ALA (2006). Resolution in opposition to "*sui generis*" database protection, available at: www.ala.org/ala/aboutala/offices/wo/referenceab/colresolutions/012506-CD20.6.pdf.

Band J. and Gowdy J.S. (1997). *Sui generis* Database Protection: Has Its Time Come? D-Lib Magazine, Washington, available at: <http://www.dlib.org/dlib/june97/06band.html>.

Benkler, Y. (2000). "Constitutional Bounds of Database Protection: The Role of Judicial Review in the Creation and Definition of Private Rights in Information", *Berkeley Technology Law Journal* (15).

Bitlaw: a resource on Technology Law, available at: <http://www.bitlaw.com/copyright/database.html>.

Cardinale, P.J. (2007). Sui generis database protection: seconds thoughts in the European Union and what it means for the United States, available at: <http://jip.kentlaw.edu/art/volume%206/6%20Chi-Kent%20J%20Intell%20Prop%20157.pdf>.

Conley J.M. et al. (1999). "Database Protection in a Digital World", The Richmond Journal of Law & Technology, (6), available at: <http://www.richmond.edu/jolt/v6i1/conley.html>.

Colston C. (2001). "Sui Generis Database Right: Ripe for Review?", Journal of Information, Law and Technology, available at: http://www2.warwick.ac.uk/fac/soc/law/elj/jilt/2001_3/colston.

Crawford, W. (2004). A middle ground on copyright, available at: <http://www.ala.org/ala/online/thecrawfordfiles/crawford2004/crawfordSept04.cfm>.

Freno M. (2001). "Database Protection: Resolving the U.S. Database Dilemma with an Eye Toward International Protection", Cornell International Law Journal, 34, p. 165.

Geneva Declaration on the Future of WIPO, available at: <http://www.cptech.org/ip/wipo/genevadeclaration.html>.

IFLA (1999). Report from the second session of the Standing Committee on Copyright and Related Rights, Geneva May 4-11, 1999. Protection of Databases, available at: <http://archive.ifla.org/III/clm/p1/wipo2rpt.htm>.

IFLA (2008). Response from IFLA to the European Commission's Green Paper 'Copyright in the Knowledge Economy' [COM (2008)466/3], available at: http://209.85.129.132/custom?q=cache:GS1R0XlkdnEJ:www.ifla.org/III/clm/p1/IFLA_Response_Green-paper-copyright.pdf+sui+generis&cd=7&hl=en&ct=clnk.

Nimmer M. B. (1999). Cases and Materials on Copyright. Mathiew Bender.

Reichman J.H. (1998). Database Protection at the Crossroads: Recent Developments and the Long term Perspective, available at: <http://library.findlaw.com/1998/Dec/2/131341.html>.

Ross A. (1999). Copyright Law and the Internet: Selected Statutes and Cases, Thelen Reid & Priest LLP, available at: <http://library.findlaw.com/1999/Jan/1/130344.html>.

US Copyright Office (1997). Report on Legal Protection for Databases, Available at: <http://www.copyright.gov/reports/dbase.html>.

Torremans P.L.C. (ed) (2004). Copyright and humans rights: freedom of expression-intellectual property-privacy, Kluwer.

WIPO World Intellectual Property Organisation, <http://www.wipo.int>.

Computerized support for ethical analysis

Mikael Laaksoharju*

Department of Information Technology –
Human Computer Interaction, Uppsala University

Iordanis Kavathatzopoulos**

Department of Information Technology –
Human Computer Interaction, Uppsala University

Abstract

EthXpert is a computer based, interactive tool to help decision makers analyze the underlying preconditions of real-life moral problems. It is based on a social scientific approach focusing on the process of ethical decision making. EthXpert's goals are: 1) to block heteronomy and support autonomy, 2) to organize interrelationships and data in a systematic way and, 3) present the complexity of the issue in a comprehensive way and provide easy access to all data. By using EthXpert a decision maker can take control of his/her own ethical decision process and of the moral problem itself.

Keywords: Ethical tool, Decision making, Problem solving, Ethical competence

Introduction

Today most groups, organizations and institutions have ethical guidelines to help their employees make the right decision. In best case these are a compilation of results from ethical contemplation, which would mean that ethically competent people have deliberated over ethical issues in hopefully relevant topics and cre-

* *Mikael Laaksoharju* is a PhD student in Human Computer Interaction. His research is concentrated on developing computer tools to support and train ethical decision making. The work is based on an assumption that structuring and presenting information in a suitable way can improve the handling of moral problems. The tools have been well received, especially within the field of Operations Research.

** *Iordanis Kavathatzopoulos* is Asc. Professor of Psychology and Human-Computer Interaction at the Department of Information Technology, Uppsala University, Sweden. His main interests are in the areas of Information Technology Ethics and Business Ethics. He has developed education programs for ethical competence of professional decision makers, and made special tests for the assessment of ethical problem-solving and decision-making ability. He has constructed tools, methods and computerized instruments for the design of ethically usable IT systems.

ated codes and principles for how to act ethically. This is the ideal, but it is in many ways still limited. For obvious reasons the nature of the resulting rules often becomes general instead of specific. For a person facing an ethical problem this means that the original problem of deciding what to do is being replaced with the problem of determining which of the dictated rules that applies in the current situation.

The focus of problem solving shifts, something that has several undesirable implications: solving ethical problems becomes rule based and thus limited to previously predicted scenarios; the awareness of unique features in a specific problem decreases and relevant questions that should be raised about a problem may end up disguised by principles and therefore wind up forgotten (see for example Eriksson et al., 2007). These unique features and relevant questions have to be taken into account in a decision-making process just as they have to be considered during the processes of creation of ethical rules. Ethics is about choice and the right choice is dependent on the quality of the process behind the choice (Ἀριστοτέλης [Aristotle], 1975; Πλάτων [Plato], 1981, 1992a, 1992b; Kant, 2002; Piaget, 1932; Kohlberg, 1985; and others).

In this paper we describe a complement, not only to guidelines but generally to any choice on ethical issues – a computer based, interactive tool to help people analyze and understand the underlying preconditions of a specific ethical or moral problem. To give motives for why this tool can help decision makers reach better, more "ethical" decisions, we will start off with defining some features common for moral problem solving.

1) People are generally not well prepared for handling moral problems. Whenever we experience that we have two or more conflicting moral obligations to fulfill we become insecure about our ability and authority to solve the problem. We generally want to perceive a difficult situation as one where we are out of options and therefore without responsibility. Psychologically we are constituted to avoid situations where we risk losing so we try hard to avoid this personal responsibility in difficult situations. This can lead us to the false conclusion that inaction is better than actions where we have to make an uncomfortable decision (Sunstein, 2005).

2) Ethical problems often regard sensitive issues and are therefore often muddled by emotions and taboos. This means that the decision maker will have a hard time to block her own tendency to biased judgment and might therefore avoid exploring the conditions further (Greene et al., 2004). The result is a submission to dogmatic thinking.

3) Ethical problems often seem impossible to solve without violating at least one important value. Some people have protected values that they will not trade off for any cost (Baron and Spranca, 1997). It could be the protection of a rare type of species; it could regard tampering with the human body.

4) In practice, the ethical problem is often created by a complicating factor that requires unreasonable investments to solve satisfactorily. Example: A medical doctor wants to analyze blood from a patient in order to rule out some possible causes for an observed illness. She therefore sends a blood sample to the laboratory along with a list of tests that she wants to have run. In the laboratory, a machine automatically runs a standard set of tests, which are far more numerous than the ones requested by the doctor. The common practice at the hospital is however to only returning the values that were requested by the physician, even if other values would show abnormalities. It would be easy to condemn this practice as unethical, if there was not the complicating factor that the tests give false positive in 5% of the cases. If all the results from the standard test were returned it would mean that a lot of patients would have to undergo unnecessary further investigations. The expense for the hospital and the discomfort for the patient in this practice have been determined to be enough reason to not overrule the doctor's initial judgment.

The most common conception of ethics is that there exists a set of principles – norms – that decide what is the right thing to do. This can be defined by either actions or results, depending on the theory of ethics. When applying this kind of normative ethics it is presupposed that the general principles described by the ethics can be specified to fit any situation at hand, something that comes with the assumption that internalized morality unconsciously guides us towards right decisions. And it is true. In most situations this kind of morality works very well – instantly and without stress. We do not need to think twice whether or not to help old ladies who have slipped on the sidewalk. We do not need to contemplate over whether it is right or wrong to steal your neighbor's garden flowers. Our morality is a force that prevents our society from decaying into chaos. A crude nonsense interpretation would say that ethics is a description of the morality that has evolved in a society and thus is something that we all share without further thought (see for example Aronfreed, 1976 and Peláez-Nogueras & Gewirtz, 1995). Morality is thus the individuals' interpretations of right and wrong and ethics is the description of it (by all means often also the prescription).

In our research we approach ethics from another angle. Instead of trying to decide the pattern and pursuing the fundamental justification for a nice decision, we focus on the mental process of decision making, which means that we can avoid matching decisions against principles of normative ethics. We define eth-

ics; just as philosophers do when they contemplate over a problem; just as it is in reality; as a description of a choice situation. When doing that, we focus on how to make the decision process optimal. An explicit focus on the process of decision making is desirable since the alternative, locking the mind on a set of untouchable principles, will only lead to knee-jerk neglecting of other important aspects. The analysis should not be constrained by preconceptions.

It should be stressed that this approach does not imply that the decisions made in this tradition will automatically become non-normative. Normative ethics is a description of human co-existence and thus a necessary and inevitable foundation for understanding how people choose. It will therefore implicitly be present in any process regarding human interests but, instead of being blinded by the impression of conflicting principles, we can make the normative values of the involved stakeholders explicit and through that let these become subject for investigation and questioning. It might appear like blasphemy to question ethical principles, but it is not really the principles as such that will be scrutinized – it is the application of them in a certain situation including interpretation of existing rules or creation itself of ethical rules. We believe that true ethical competence requires a deliberate approach toward ethical principles; a view that will allow for the state of reflective equilibrium that is required to neutrally weigh ethical considerations.

Everybody probably agree that it would be better if we all by nature were acting ethically in the sense of philosophizing in the right way. In fact we do try hard to do so. Sadly, we are not always able to handle the moral problem at hand in a satisfactory way. In our daily life we get little training in ethical problem solving and in developing our ethical awareness. Therefore we have come to create and accept ethical guidelines as a necessity to support our handling of moral problems. And this is the rule for most of us (see for example Kavathatzopoulos & Rigas, 1998, 2006).

However, ethical guidelines, beside the interpretation and construction problems, can lead to negative moralizing instead of supporting deliberate ethical thinking. Moralizing is an important aspect when considering moral issues. Controversial matters cause a knee-jerk reaction to take the position that we unconsciously judge will give us the least headache. In other matters we have a hard time to even understand why some behavior is ethical and some other is not. We use our moral heuristics to make effortless decisions in situations calling for moral awareness (Sunstein 2005) and in yet other situations we hurry to adopt the attitude that is closest to our personal interests while still accepted in our social context. Foucault (1987) and Gustafsson (1997) address the importance of moralizing. The former uses the notion of power, or control, to describe how moralizing works as the force that keeps societies together and the latter recognizes the interplay of moralizing as the communication that stabilizes societies.

In both interpretations moralizing is the core of a society, which means that we will inevitably give ourselves over to this in any choice situation. It is necessary and purposeful in our daily life but when it comes to situations where we need to make a decision that affects other people, this attitude can become an obstacle. It prevents us from deliberating clearly on the implications from our choices and it adds stress from deviance on top of the stress from being responsible for a decision. To be able to block this heteronomous behavior and stimulate an autonomous process we need to structure and systematize our thinking. Such a process will also help to justify difficult, possibly uncomfortable, decisions.

EthXpert

EthXpert is a tool to aid people in their analysis of an ethical problem situation (Laaksoharju & Kavathatzopoulos, in press; Kavathatzopoulos, Laaksoharju & Rick, 2007). It is intended to complement codes of conduct and guidelines and builds on the assumption that ethical competence is equivalent to a well-functioning problem-solving strategy (Kavathatzopoulos, 2003, 2004; Erlandsson & Kavathatzopoulos, 2005). The main requirement on the system is that it should not be making any decisions and not even supporting any specific solutions. The sole intention with the tool is to help the user to organize and structure the problem at hand. At the same time the problem should not be narrowed down, thus risking oversimplification, but instead be expanded and widened. EthXpert's aim is: 1) to block heteronomy and support autonomy (Piaget, 1932), 2) to organize interrelationships and data in a systematic way and, 3) present the complexity of the issue in a comprehensive way and provide easy access to all data.

At a first thought the widening of the problem scope might appear as an uncalled-for disservice, but in fact it is exactly the invaluable help that a scrupulous decision maker requires – help to get the fullest possible picture of a problem. The tool does not give any directions about the correctness of any conclusion and will therefore force the decision maker to analyze the problem very carefully. It goes without saying that it is very hard to decide in ethical problem situations – it is in the nature of the problem. The conflicting principles and values will all seem too important to trade off and the outcome of any realistic option will appear to have undesired features. This is the common perception of an ethical problem.

However, in real life many ethical problems occur instead from the lack of information or misinterpretation of responsibilities. The intentions might be good but if sensitive information is missing a wrong decision may be taken. Therefore it is important to allow the person facing the problem to freely add information to the analysis whenever there seems to be a reason for it. To make a well-founded decision, it is desirable to collect as much data as possible but the problem with massive amounts data is however apparent; the chance to make use of it decreases with the

amount. Many of the approaches presented in earlier work (Maner, 2002) suggest different strategies to eliminate matters that are not relevant for the problem. The impending risk with elimination is to lose important aspects. A better approach is to let the decision maker be selective when it comes to analyzing the data. With EthXpert we promote configurable representations, where only the data associated to a specific part of the problem is viewed (Fig. 1).

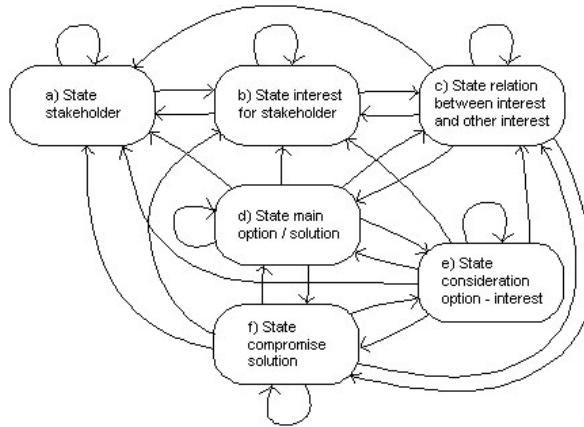


Figure 1: State machine representation of ethical procedure in EthXpert

To illustrate why hesitation toward analyzing a problem situation too much is natural we will make up a simple and perhaps suitable daily-life example: You have to choose whether to present your latest research paper at an important conference or to be present at your six-year old daughter's birthday party. On the one hand you have obligations towards your university to be a diligent researcher, towards yourself to make a decent career and towards the research community to present your research findings. On the other hand you have the obligations to show love towards your daughter and to care for the harmony in your nuclear family. Take a moment of introspection to reflect over how you would approach this dilemma.

- Would you start comparing the weight of the different obligations to find out which one is heavier?
- Would you try to derive the underlying moral principles behind the dilemma and somehow compare them? In this case there are most likely different types of responsibility involved.
- Would you investigate whether there are any compromise solutions to the problem? Maybe you could send someone else to represent you at the conference, maybe you could postpone the celebration of your daughter's birthday or maybe you could even bring your family to the conference site?

- Would you try to imagine the consequences of different solutions? Is there a risk that your daughter will be very disappointed for a long time if you do not stay at home? Are there any complicating factors (maybe your spouse is fed up with you constantly avoiding your responsibility for the family)?
- Would you try to identify the involved stakeholders (i.e. the university, you, your daughter, your spouse e.tc.) and try to find out the opinions of these?
- Would you consult “experts” to help you choose right?
- Would you try to identify what the different stakeholders’ interests really are? Perhaps it is not quite as you thought. Your daughter might be indifferent to whether you are present at the party or not, the university might not care at all about the conferences you choose to attend, the research community do not need your presentation e.tc.; what if you would end up at the conclusion that you are utterly useless and completely uninteresting for anyone around you! Is that what you wanted to find out?

It should be duly noted that this drastic result is not a very likely result from your investigation. If you went through such a process in your mind and then settled for a solution, then you will most likely have reached a good one. The example is too drastic both in the sense that the conflicting obligations are apparent and the conclusions overly brutal. The ethical implications were easily recognized but more often ethical decision making is about identifying problems arising from conflicting values in complex situations where the possible problems are not as apparent.

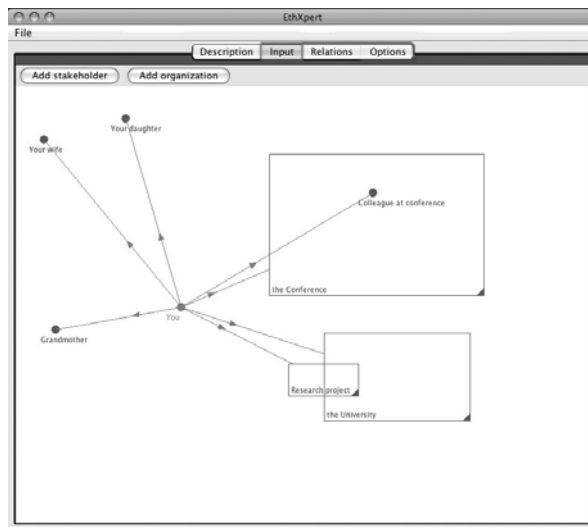


Figure 2: Addition of stakeholders and visualizing relationships

States of the ethical analysis process

Define stakeholders

It can be very difficult to identify all stakeholders that affect or are affected by a decision. Several ethical support systems target this concern in different ways. In Paramedic Ethics (Collins & Miller, 1992) focus is aimed at the obligations and responsibilities of the decision maker. In SoDIS (Gotterbarn, 2002) the decision maker is asked a set of questions about likely reasons for ethical problems. It should be noted that both of these systems are intended for technically skilled IT professionals. With EthXpert we aim wider and do not assume any specific content in the problem to be analyzed. It is therefore impossible to guide the user by asking questions about previously known reasons for ethical problems to occur. This absence we consider as strength when it comes to widening the agenda for the problem situation. The user is thus never lured into the false comfort in believing that the analysis is finished. As with the lack of final state in the state machine representation of the process, ethical deliberation should leave a feeling of insecurity when ending the analysis. Such a setup puts the responsibility for a satisfactory analysis on the analyst. In EthXpert the addition of stakeholders is very easy and therefore supports the problem expansion. For each stakeholder that has a direct relationship to the problem there will most likely be third-party stakeholders that might influence the outcome of the decision. Such stakeholders are naturally identified when the interests of the apparent stakeholders are inspected (Fig. 2).

Define for each stakeholder its interests

In EthXpert it is assumed that the explicit focus on interests of the stakeholders will help the decision maker to identify possible conflicts and also to widen the scope of the problem. All interests that might relate and affect other stakeholders are important to consider and in the process of scrutinizing interests additional stakeholders will naturally become involved in the analysis.

The screenshot shows the EthXpert application window. It features a menu bar with 'File' and a tabbed interface with 'Description', 'Input', 'Relations', and 'Options'. The main area is a table with columns for stakeholder interests. A modal dialog titled 'Responsibility -> Colleague at conference' is open, showing a text area with the text: 'If your presence is invaluable to your colleague, maybe you have a responsibility to attend?'. The dialog has 'Update' and 'Add' buttons.

Stakeholder	Interest	Value	Impact	Priority	Weight	Score	Notes
You	Reputation	A good	Depeding	A University	If you are a	Colleague at conference	Research
Happy	It makes	Keeping	Does your				Grandmoth
Responsibility	You owe to yourself to create opportunities to do the kind of research you if	You naturally have responsibility toward your daughter.	Your responsibility toward you probably based on love and care.	You have a responsibility to fulfill your research commitment towards the	Maybe you want to maintain a good relationship towards the	If your presence is invaluable to your colleague, maybe you have a responsibility to attend?	You have a joint responsibility to promote your research
Your	Happy life	You	If your presence is invaluable to your colleague, maybe you have a responsibility to attend?				
Presents	Prob						
Attention	May						
Your wife	Happy	You need to keep					
Responsibility							
the	Reputation						
Good							
the	interesting						
Reputation							
Economy	Every						
Colleague	Reputation	If you					
Research							
Produce	It does not						
Receive	Having a						
Reputation							
Grandmoth	Stability						
Responsibility							
Gossip							

Figure 3: State how interests affect and are affected by stakeholders

Define how interests relate to other stakeholders

Determining how the interests of the stakeholders relate to other stakeholders is the core of the analysis (Figure 1). This creates a deepened understanding of the dynamics of the ethical problem and will help to track down previously unidentified stakeholders (Fig. 3). Most of the time peripheral stakeholders remain just peripheral, but sometimes these stakeholders prove to have an important influence for the dynamics of the problem. Explicitly stating how the interests affect and are affected by other stakeholders gives a background for the further analysis of implications from different decision alternatives.

Define main options

The most apparent alternatives for handling the ethical problem can be immediately stated. Usually main alternatives are to their character mutually excluding, similar to answering a question with “Yes” or “No”. The procedure in EthXpert helps the user to identify possible ethical implications through

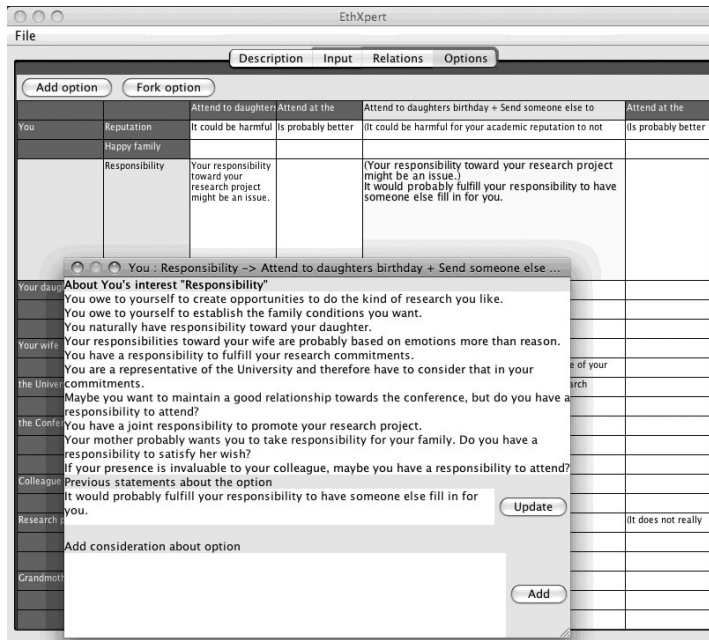


Figure 4: Defining options and translating considerations

Translate considerations

The considerations from the interest-stakeholder matrix will not be automatically copied to the decision matrix. Instead the interest-stakeholder relationships will serve as background and incentive for considering how the different decision alternatives affect the stakeholders (Fig. 4).

Define compromise options

To tackle problems in the main options, i.e. unacceptable negative effects, compromise decision alternatives can be forked off from main alternatives. The compromise option will inherit considerations from the parent, but the analyst should revise these and determine how the compromise addition affects the effect on the stakeholder interest.

So if collecting as much information as possible about a problem helps us to better deal with it, why don't we just let everybody interested contribute to the process? That idea is not bad as such. Living Labs show that there is much to gain from allowing the public to influence also complex decision processes. For the decision maker, just skimming through the mass of information created by multi-

ple sources would most likely help to acquire a wider perspective on the problem. The difficulty with such an approach is how to digest huge amounts of unclassified data; the problem becomes a cognitive or computational one. Tools for automatically narrowing down such information will be based on how the software designer has interpreted the original ethical problem, so there is an apparent risk for selective search for evidence and choice-supportive bias if the information is heterogeneous and allows for interpretation. Also, we should not forget that ethical problems are of the type that makes it unsuitable or even counterproductive to ask for everyone's opinion. Some decisions inevitably become uncomfortable, or even unacceptable, for a lot of people in order to protect other values. People do not share values and clearly showing that they can affect a choice situation would probably render in extreme polarization of opinions (Baron & Spranca, 1997). However, the decision maker could ask the right questions, without revealing the sensitive suggested solutions.

Conclusion

EthXpert is an ethical analysis and decision support system that can be used in handling all kind of moral problems. As we have seen ethical rules and guidelines cannot provide solutions. They demand interpretations and adaptation to actual conditions. EthXpert can be used for such purposes too. It is a tool to handle ethical problems in practice. We believe that, by systematically examining how the stakeholders' interests affect and are affected by other stakeholders it is possible to gain a better understanding of the mechanisms of ethical choice. In a world where there are many core values and an infinite number of interpretations we can not expect to be successful in handling issues in a normative, dogmatic way.

REFERENCES

- Ἀριστοτέλης [Aristotle] (1975) Ἠθικά Νικομάχεια [Nicomachean Ethics]. Πάπυρος [Papyrus], Αθήνα [Athens].
- Aronfreed, J. (1976) Moral development from the standpoint of a general psychological theory. T. Lickona (ed.), Moral development and behavior. Holt, Rinehart & Wilson, New York.
- Baron, J. and Spranca M. (1997) Protected values. *Organizational Behavior & Human Decision Processes*, 70, 1-16.
- Collins, W. R. and Miller, K. W. (1992) Paramedic ethics for computer professionals. *Journal of Systems Software*, 17, 23-38.

Eriksson S., Helgesson G. and Höglund A. T. (2007) Being, doing, and knowing: Developing ethical competence in health care. *Journal of Academic Ethics*, 5, 207–216.

Foucault M. (1987) Panoptismen. In *Övervakning och straff – Fängelsets födelse*, 196–228. Arkiv, Lund.

Gotterbarn, D. W. (2002) Reducing software failures: Addressing the ethical risks of the software development lifecycle. *Australian Journal of Information Systems*, 9(2), 155–165.

Greene J. D., Nystrom L. E., Engell A. D., Darey J. M., and Cohen J. D. (2004) The neural bases of cognitive conflict and control in moral judgment. *Neuron*, 44, 389–400.

Gustafsson, C. (1997) Den lystna maskinen [The greedy machine]. In A. Björns-son and P. Luthersson (eds.) *Medialiseringen av Sverige*, 63–75. Carlsson, Stockholm.

Erlandsson, M. and Kavathatzopoulos, I. (2005) Autonomy method: Acquiring skills for ethical analysis of computerisation in car driving [CD-ROM]. In G. Collste, S. O. Hansson, S. Rogerson & T. W. Bynum (eds.), *ETHICOMP 2005: Looking back to the future*. Linköping: Linköping University

Kant, I. (2002) Prolegomena till varje framtida metafysik som skall kunna uppträda som vetenskap [Prolegomena to Any Future Metaphysics]. Thales, Stockholm.

Kavathatzopoulos, I. (2003) The use of information technology in the training for ethical competence in business. *Journal of Business Ethics*, 48, 43–51.

Kavathatzopoulos, I. (2004) Making ethical decisions in professional life. In H. Montgomery, R. Lipshitz & B. Brehmer (eds.) *How professionals make decisions*, 277–288. Lawrence Erlbaum Associates, Inc, Mahwah, NJ.

Kavathatzopoulos, I., Laaksoharju, M. and Rick, C. (2007) Simulation and support in ethical decision making. In T. W. Bynum, K. Murata & R. Simon (eds.) *Globalisation: Bridging the global nature of Information and Communication Technology and the local nature of human beings*, 278–287. Meiji University, Tokyo.

Kavathatzopoulos, I. and Rigas, G. (1998) A Piagetian scale for the measurement of ethical competence in politics. *Educational and Psychological Measurement*, 58, 791–803.

Kavathatzopoulos, I. and Rigas, G. (2006) A measurement model for ethical competence in business. *Journal of Business Ethics Education*, 3, 55–74.

Kohlberg, L. (1985) The Just Community: Approach to moral education in theory and practice. In M. Berkowitz and F. Oser (eds.), *Moral Education: Theory and Application*. Lawrence Erlbaum Associates, Hillsdale, NJ.

Laaksoharju, M. and Kavathatzopoulos, I. (in press) EthXpert: the basic structure and functionality of a decision support system in Ethics. *International Transactions in Operational Research*.

Maner, W. (2002) Heuristic methods for computer ethics. *Metaphilosophy*, 33, 339-365.

Peláez-Nogueras, M. and Gewirtz, J. L. (1995) The learning of moral behavior. In W. M. Kurtines & J. L. Gewirtz (eds.), *Moral development*, 173-199. Allyn and Bacon, Boston.

Piaget, J. (1932) *The moral judgement of the child*. Routledge and Kegan Paul, London.

Πλάτων [Plato] (1981) Θεαίητος, [Theaitetos]. Ι. Ζαχαρόπουλος [I. Zacharopoulos], Αθήνα [Athens].

Πλάτων [Plato] (1992a) Πολιτεία, [The Republic]. Κάκτος [Kaktos], Αθήνα [Athens].

Πλάτων [Plato] (1992b) Ἀπολογία Σωκράτους, [Apology of Socrates]. Κάκτος [Kaktos], Αθήνα [Athens].

Sunstein, C. R. (2005) Moral heuristics. *Behavioral and Brain Sciences*, 28, 531-573.

Is the spammer evil?

Eleftherios Loukis*

University of Piraeus, Department of Informatics,

Ioannis Papadakis**

Ionian University,
Department of Archives and Library Sciences,

Abstract

The email system has evolved from a great way to communicate, share information and collaborate to an inconvenient tool that progressively loses its reliability and becomes obsolete. This paper defines the problem, describes its magnitude and discusses the various measures that have been taken so far to overcome it. Moreover, it argues that the vast majority of spam could be avoided if so-called spammers are confronted as a community of email users with specific, legitimate motives instead of a minority of unethical outlaws.

Keywords: email, spam, CAN-SPAM Act

Introduction

Internet today is a true information highway. Many aspects of everyday life (social, cultural, financial, etc) depend on it in various ways. Its abundant capacity, its global reach and its “no gatekeeper” infrastructure are among the most significant reasons why Internet has become an integral part of our lives.

* *John Papadakis:* He was born in Athens at 1975 and he studied Computer Science (1997) at the Department of Computer Science at the University of Piraeus. He received his PhD with the title: «Digital Libraries: Architectures, Security and Information Retrieval» at the same Department. Since 2005 he works at the Department of Archives and Library Science at the Ionian University. During the past few years his scientific interests include the areas of the semantic web and the web in general.

** *Loukis Eleftherios:* He was born in Athens at 1974 and he received his bachelor diploma in Computer Science from the University of Piraeus, in 1997. At 2001 he graduated from Athens University of Economics and Business with a Master's Degree in Information Systems. He is currently working as a Computer Analyst and IT specialist at various companies as well as at the IT department of the Library of University of Piraeus. He has also participated to various information technology research projects at the University of Piraeus where he developed his interest on computer and information research.

However, it seems that the email system, one of Internet's most widely appreciated applications, has followed a rather opposite course. What started out as a great way to communicate, to share information, to collaborate, is turning into an inconvenient tool that is progressively losing its reliability and becoming obsolete.

Due to the seriousness of the situation, the Internet society was eventually forced to provide formal definitions to the apparent email problem. In this direction, the term 'spam' was introduced. Surprisingly enough, it has been proven to be very difficult to come up with a definition capable of unambiguously identifying all of the emails that are responsible for the email problem. Consequently, a widely agreed and working spam definition does not seem to exist.

Beyond the controversy of the term 'spam' the email problem is an undisputable reality that needs to be confronted. Based on the previously mentioned formal definitions, a number of antispam approaches emerged. Such approaches derive from a vast number of disciplines such as law, mathematics, statistics, sociology e.tc. Despite their diversity, they share the common objective of eliminating the email problem. Depending on their point of view, current antispam approaches could be classified according to the specific goals they are trying to achieve.

In this paper, it is argued that effective antispam solutions should take a closer look at the so-called 'spammers' community, which is comprised of stakeholders with specific motives for relentlessly sending countless emails to unsuspected recipients. However, such motives are not necessarily unethical. Thus, the members of this community could be classified into three broad classes, depending on the various roles an email sender could be associated with. Thus, a sender could be a merchant aiming at finding potential customers for his/hers products. In this case, there are commercial motives for sending such emails. In another case, an email sender could be a person wishing to promote his/hers image or ideas to the public (candidate for elections, celebrity, non-profit organizations, etc). In this case, there are promotional motives for sending such emails. Also, email senders could be ill motivated individuals aiming at deceiving their recipients (e.g. spoofing, phishing, chain emails, etc). In this case, there are unethical motives for sending such emails. Maybe, if email senders belonging to the spammers community had the chance to more efficiently direct their emails to the appropriate recipients, their motives would be satisfied and at the same time the email problem would shrink. It should be noted that the above argument does not refer to senders with deceptive motives such as hackers, frauders, e.tc.

The rest of this paper is structured as follows: Initially, a definition of the email problem is provided. The next section deals with the term spam as a synonym to the email problem. Various existing antispam approaches are mentioned and ac-

cordingly classified in the following section. Next, in order to better understand the problem, an effort is made to take into consideration the spammer's perspective of spam. Finally, conclusions are drawn and the line of thoughts expressed throughout this paper is summarized.

Defining the problem

Today, the answer to the question "what is your opinion about the email system?" is disappointingly negative as expressed from the vast majority of the Internet society. Alternative forms of Internet communications have emerged (i.e. instant messengers, web 2.0 technologies, social networks etc) that attract the attention of a continuously growing number of Internet users (O'Reilly, 2007). After more than 30 years of prosperity, the email system is currently going through a declining course. Email providers waste time, money and resources in order to route an ever-increasing volume of emails to their clients. Most of the times such emails end up in the trash folder¹. At the same time, email users are daily forced to deal with too many useless emails that clutter their inboxes.

It should be noticed though, that not all useless emails are guilty for the email problem. Useless emails originating from the recipient's social network were, are and always will be present. Useless emails coming from outside a user's social network are those responsible for the email problem.

They can be divided into two main categories: objectively useless emails and subjectively useless emails. The former refers to malicious emails that are always useless for everybody (i.e. frauds, viruses, phishing etc) and the latter refers to emails that are useless to a specific set of receivers and/or at a specific period of time (i.e. commercial, promotional, advertisements, etc). For example, there are people that would actually find an email about losing weight useful while at the same time, other people would consider such an email as annoying. Moreover, people that have already dealt with their weight problem would in all likelihood prefer not to receive emails about losing weight any more. The above statement defines the subjectivity principle for useless emails.

The term "spam"

At present, a widely agreed and workable spam definition does not seem to exist (Schryen, 2007a). This is due to the fact that most of current spam definitions are based on various features of spam (i.e. commercial, bulk, unsolicited), that seem to be only partially valid, resulting in vulnerable 'anti-spam' legislation and inefficient software tools against it.

1. Spam-o-meter Statistics By Percentage: Spam statistics, Available at: <http://www.spam-o-meter.com> accessed 15 September 2008.

More specifically, spam definitions provided both by the US and Europe are mainly directed towards email produced for commercial purposes (Schryen, 2007b), neglecting other aspects of the email problem such as fraud, phishing, promotional, e.tc.

Moreover, current spam definitions heavily rely on the term 'bulk'. Bulkiness refers to the aggregation of identical (or near-identical) emails delivered by an email sender. In contrast, the email problem mostly refers to single emails existing within the cluttered inbox of email recipients. Thus, when trying to decide whether an email is bulk or not we actually have to find out whether this email originates from a single sender that has massively sent identical (or near-identical) emails. Such a task is quite often difficult to accomplish, despite current efforts in collaborative antispam solutions (Ramachandran et al., 2006).

The term 'unsolicited' is also commonly met in current spam definitions. However, the fact that spam is unsolicited should not imply that only spam is unsolicited. The very nature of the email process partly relies on unsolicited communication. Thus, useful emails being sent without prior consent from their recipients could be mistakenly considered as spam. Besides, even if a recipient has agreed to receive an email, how and when such consensus is established, may not be obvious. It is very difficult to know whether a relationship between the sender and the recipient already exists or not. To make things worse, even if such a relationship does not exist, there is a lot to be done before accusing the sender for improperly harvesting the recipient's email address. Ultimately, engaging the term 'unsolicited' in the context of spam definition seems to lead to more problems than solutions, at least as far as legislation is concerned.

It is also important that emails which are considered as part of the email problem from 90% of their recipients, are considered as useful from the remaining 10%. Engaging subjectivity within a formal definition renders such definition as erroneous.

Spam could be correctly associated with the act of repeatedly sending the same email to specific recipients. However, spam has evolved in a way that exactly the same email seldom finds its way to the same inbox. Senders usually make minor changes to an email before sending it to receivers that have already received a similar one. Thus, definitions containing such a term could be considered as controversial.

Two more terms finding their way into many spam definitions (Schryen, 2007a) are 'untargeted' and 'indiscriminate'. Similarly to the arguments about 'unsolicited', these two terms characterize many emails that are not necessarily considered

as spam. Moreover, it is difficult to prove whether an email was indiscriminately sent from its sender.

Quite often, emails that are considered as part of the email problem have certain properties (i.e. anonymous and/or disguised emails with deceptive, fraudulent, illegal or offensive content). However, it is risky for formal definitions of spam to incorporate such properties. Not all emails considered as spam have such properties. Thus, even just one spam email that does not have any of the above properties is enough to cause problems both to legislative approaches and antispam mechanisms.

Moreover, the fact that formal definitions of spam do not take under consideration the previously mentioned subjectivity principle for useless emails, renders such definitions as cumbersome: an email has to be treated as spam for all occasions, at all times for everybody.

Beyond the controversy of the term 'spam' the email problem is an undisputable reality that needs to be solved. Based on the previously mentioned formal definitions, a number of antispam approaches have been proposed. However, it seems that in the name of the crusade against spam, some approaches unintentionally contribute to the overall email decline. For example, when useful emails are mistakenly filtered as spam and thus never make it to the recipients' inbox, unreliability is added as another drawback to the already suffering email system.

Antispam approaches

Given the fact that the email problem is apparent to everyone, it is necessary to take a closer look at current antispam approaches in order to track down their vulnerabilities. By doing so, we will be able to avoid pitfalls and propose truly viable solutions that will eliminate, or at least reduce the overall email problem.

In order to better understand current trends in fighting against spam, antispam approaches should be classified into a consistent taxonomy. However, by taking a closer look at the corresponding literature, it appears that it is not an easy task to create such taxonomy. Indeed, many antispam efforts seem to have ambiguous origin, thus being eligible to be classified in more than one category.

Conversely, many existing taxonomies consist of categories with vaguely defined borders, thus allowing such categories to host more than one approach simultaneously. For example, in (Judge, 2006) the category labeled as "legal prosecution" is a subclass of "spam responses" whereas it could easily be a member of "spam deterrence approaches", since "legal prosecution" is a discouraging factor for anybody thinking of sending junk email.

Having the above thoughts in mind, a simple taxonomy is proposed aiming at unambiguously classifying current antispam approaches into two mutually exclusive categories: The first category contains approaches aiming at preventing the act of spam. Usually, prevention is accomplished by discouraging potential spammers. The second category contains approaches that can only be put into action after the birth of spam. Spam is being born at the time when a spammer actually sends emails. In their majority, the main objective of approaches belonging to this category is to hide the problem from the users instead of solving it: despite the fact that spam never arrives to end users, ISPs suffer from resource abuse due to the amount of spam reaching their servers. Thus, the consequences of spam indirectly reflect to end users (i.e. delays, denial of service, etc).

From another point of view, the proposed taxonomy consists of one category aiming at confronting the true causes of spam (i.e. ‘Prevention’) and another one focusing on eliminating the consequences of spam (i.e. ‘Suppression’).

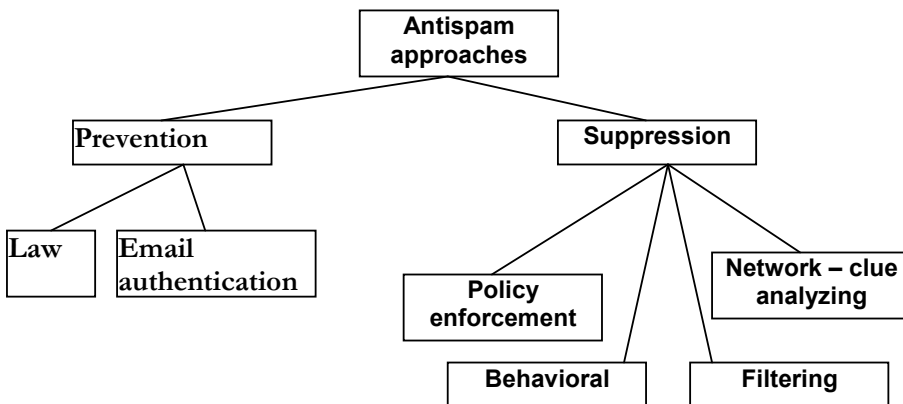


Figure 1. Proposed antispam approaches taxonomy

As illustrated in figure 1, there are two main categories for spam prevention, namely “email authentication” and “law”. On the other hand, “policy enforcement”, “filtering”, “network – clue analyzing” and “behavioral” belong to the spam suppression category. A detailed description of these approaches is presented in the following section.

Email authentication techniques

The lack of authentication is perhaps the most significant vulnerability of the email system. A functional authentication system, incorporated within a frame-

work capable of punishing the ones responsible for today's email problem (spammers, phishers, virus creators, etc), in all likelihood deter malicious email users from sending junk emails.

However, it seems that incorporating authentication in the email system is far from a trivial procedure due to two important reasons: First of all, Internet users are accustomed to a mostly anonymous environment that exercises no control over their actions. Thus, any kind of authentication could be faced with reluctance and distrust, since users would have to comply with additional constraints. In addition, the act of authentication is by itself an additional burden for users that have to be convinced for its necessity. Secondly, email vendors would also be reluctant in deploying email authentication systems. Such deployments usually require fundamental changes to existing email infrastructures (e.g. policies (Lawton, 2006)) for which a great deal of resources have already been invested in terms of money, time and expertise.

To make things worse, even if email authentication was considered as well-established, the email problem could get worse due to the existence of computers that have remotely been taken under the control of spammers, thus operating as "zombies". Indeed, within such an environment, it would be difficult to prove that the authenticated sender of an email identified as spam is not the one that actually sent it. As a result, innocent, technologically naive users would be mistakenly accused as spammers just because their computers have been operating as zombies (Levy, 2003).

Law

In terms of legislation, most countries and regimes that provide legal frameworks against spam are influenced by the rationale of the opt-in or opt-out approaches (Moustakas et al., 2005). In an opt-in system, unsolicited communications – in this case emails sent to recipients without their prior consent – are considered illegal. In an opt-out system, the sender is allowed to send the first email to each recipient, provided that it contains an option to allow the receiver to declare that he/she doesn't wish to get any additional emails from the sender. Whether spam is legal or not depends upon the opt-in or opt-out choice exercised by the recipient of the communications and also on the legal system of reference (Lugaresi, 2004).

It could be argued that the opt-out approach justifies the action of sending spam at least for the first time. From another point of view, in an opt-in solution a spam is prohibited before the spammer has a chance to communicate it, thus eventually restricting freedom of speech (Fingerman, 2004).

European Union's – EU's approach against spam mostly focuses on protecting consumers' privacy and economic interests. Specifically, EU laws and directives do not clearly define the term "spam" and the enforcement of an implied opt-out approach² is restricted to natural persons only. Based on these principles, EU solution should not be regarded as the ultimate answer to spam, but as an attempt to provide for its Member States a rational discipline together with a possible model for a harmonized approach to reducing spam (Lugaresi, 2004).

On the other hand, the United States' Congress enforced the CAN-SPAM Act³, which also envisions fighting spam in legal terms. It is primarily targeted towards unsolicited commercial "spam" email, which is punished by imposing stiff civil penalties and even prison sentences to parties convicted of spamming (Leavitt, 2007). The CAN-SPAM Act is based on the opt-out model. Unfortunately, according to surveys (Soma et al., 2008), the CAN-SPAM Act has had little consequence since going into effect on January 2004. One major drawback of the CAN-SPAM Act is that it drives spammers to hire short-term employees or form shell corporations or otherwise come up with temporary accounts under which to send spam. Moreover, it fails to take under consideration the knowledge of other countries that have tested opt-out legislation and have discovered that it simply does not work (Soma et al., 2008). It is also disappointing that the CAN-SPAM Act is working towards comprehensive "do not email" lists. Indeed, such lists may be dangerous for users' privacy. Apart from security concerns, the risk is to end up with a Big Brother effect (Lugaresi, 2004).

A broader argument against legislation-based solutions to the spam problem (including the US and Europe) is that such legislations are only enforceable within the corresponding regime. Thus, spammers change their tactics or simply move their servers to locations without antispam regulations (Moustakas et al., 2005). Consequently, Internet's universality renders the enforcement of a global anti-spam legislation as a very difficult task to achieve.

Finally, defining yet another problematic antispam law may lead to the opposite effect of legitimizing gigantic amounts of spam (Goodman et al., 2007).

Policy enforcement (Pricing - Hips - Delay)

Policy enforcement is based on reducing the effortlessness of sending email, by changing the action of sending email. This can be achieved either by pricing emails, or by requiring senders to solve some kind of puzzle (Human Interactive Problem Solving - hips (von Ahn et al., 2003) or by simply consuming additional

2. Article 1, Dir. 2000/31/EC, accessed Apr. 7, 2008.

3. CAN-SPAM, www.spamlaws.com/federal/108s877.html, accessed Oct. 1, 2008.

resources from the email sender (e.g. the Penny Black Project⁴ and others ((Cobb, 2008), (Dwork et al., 1993), (Loder et al., 2004))).

Such techniques may be applied in an 'as-is' basis (Hansel, 2004), where all email senders are burdened for each message they send. Alternatively, they could be integrated with spammer identification systems where additional burden is only applied to potential spammers. Consequently, spammers have to think twice before sending spam messages due to the considerable effort required for the sending action.

However, many of the above methods depend on authentication infrastructures that are currently not widely available to the Internet society. Moreover, the additional transactions required for sending email (e.g. challenge-response systems), affect the overall Internet's performance.

And finally, from a social point of view, email users will most likely find the additional burden cumbersome, e.g. by repeatedly being asked to solve a puzzle or being charged for each email they send.

Filtering

This category refers to antispam approaches relying on various kinds of filters in order to decide which email should be considered as spam. Filtering is maybe the most popular method for fighting spam.

There are several kinds of filters available. Depending on the approach method, such filters may be divided into several classes: Thus, one could come across a) individual and collaborative filters (Gray et al., 2004), b) statistical content-based filters, c) static and learning ones, and d) standalone/server-side filters (Garcia et al., 2004).

However, despite the considerable popularity of such approaches, there are a number of issues that inevitably come under discussion.

Initially, the inherent possibility of mistakenly classifying legitimate emails as spam (i.e. false positives) is one of the most serious drawbacks of antispam filtering (Graham, 2004). According to (Yih et al., 2006), users' criteria in selecting antispam filters are not based on their success on blocking as much spam as possible. Instead, they prefer filters that feature low rate of false positives. No matter how sophisticated a filter is, there is always the possibility of blocking legitimate email as spam. Moreover, the evolution of spam techniques over time, leads to the conclusion that spammers are practically unpredictable. Currently, there

4. The Penny Black Project, <http://research.microsoft.com/research/sv/PennyBlack>, Microsoft Research, accessed Oct. 1, 2008.

are spam techniques that manage to bypass filters by employing various methods (images instead of text, poems, obfuscation, etc (Yerazunis, 2004), (Hulten et al., 2004)). It seems realistic to claim that any weaknesses of current or future spam filters will eventually be exploited by spammers (Lowd et al., 2004).

Last but not least, common spam filtering applications eventually hide the spam problem from users instead of eliminating it. They classify and/or delete spam messages after they've been received by systems, but by then much of the bandwidth and processing costs of the spam have already been wasted (Goodman et al., 2007).

To sum up, despite the wide propagation of various filtering models, it seems that the false positives phenomenon will always be a discouraging factor of adopting this kind of antispam approach.

Network - Clue Analyzing

This category contains approaches trying to identify either spam messages or spam servers by monitoring the trails that are being left by the action of sending an email in order to detect suspicious elements.

Approaches belonging to this class focus on examining the context of an email (e.g. recipient's email account, IP address, etc), in contrast to the previous class (i.e. filtering) which contains approaches trying to identify spam messages from their main content. Apart from this difference, both categories aim at identifying spam emails by processing emails that have been arrived at the receiver's ISP. Consequently, this category has many of the drawbacks mentioned in the previous section, such as false positives, false negatives, wasting of resources, e.tc.

Behavioral

Behavioral approaches are trying to fight spam by maintaining registries of already known spam messages and/or spam machines. Such registries are commonly known as lists (e.g. SORBS⁵, Spamhaus⁶, SpamCop⁷, SURBLs⁸, etc). They may be managed either locally by individual users through appropriately designed email clients or published online in so-called collaborative lists. This category comes as a complement to previous categories since messages should already have been created, sent and identified as spam.

5. Sorbs, <http://www.sorbs.net/>, accessed Sept. 1, 2008.

6. Spamhaus, <http://www.spamhaus.org/>, accessed Sept. 1, 2008.

7. SpamCop, www.spamcop.net, accessed Sept. 1, 2008.

8. Surbl, <http://www.surbl.org/>, accessed Sept. 1, 2008.

However, similarly to the previously mentioned classes “clue analyzing” and “filtering”, behavioral approaches are targeted towards the symptoms of the problem (i.e. end-user annoyance) instead of the actual causes. Such approaches not only aim in hiding the consequences of spam from end-users but also they take no measures at all to relieve other parties involved in the email life-cycle (e.g. ISPs). Moreover, the above approaches implicitly raise the issue of correctly identifying the entity to punish. Thus, in the case of a company (i.e. stakeholder) delegating a spammer to send commercial unsolicited bulk emails through a possibly unaware ISP or a compromised machine (i.e. zombie), the question “who should eventually take the blame?” arises. Possible answers include the stakeholder, the spammer, the ISP and the ignorant user of a compromised machine. Finally, blocking spam through the employment of various kinds of blacklists is prone to the fact that there might be a considerable time lag between the creation of a new spam entity and its discovery/inclusion in a list (Pu et al., 2004). Thus, the spammer may already have achieved the anticipated goals and the receiver may already have suffered the corresponding consequences.

Hybrid

Apart from the aforementioned classes for fighting spam, there are also some efforts that try to integrate a number of individual strategies into a single system. Obviously, such hybrid solutions inherit not only the benefits but also the drawbacks of the integrated subsystems. Consequently, their main effort is to compensate the drawbacks of one contained approach with the advantages of another. Ideally, the final system results in an all-in-one robust antispam approach, which performs better as a unified whole than the aggregation of its subparts.

For example, spamassassin⁹ is based on content-matching rules, also supporting DNS-based, checksum-based and statistical filtering, powered by external programs and online collaboration lists. It is considered as a promising solution to the email problem, even though, as stated earlier, it has to deal with a number of issues accumulated from the techniques it encompasses.

Deconstructing spammer

So far, a number of different antispam approaches have been mentioned, based on the common assumption that the spammer is a malicious entity, aiming at circulating large volumes of spam. However, not all spam is useless. Indeed, the previously mentioned subjectivity principle dictates that the same email that has been flagged as “spam” from one recipient could be flagged as “not spam” from

9. SpamAssassin, <http://www.spamassassin.org/>, accessed Sept. 1, 2008.

another one, or, even from the same recipient at another time. Moreover, the fact that 10% of spam constitutes the starting point of some kind of transaction (Weiss, 2003), renders such emails as potentially useful.

As it is mentioned throughout this paper, there are different definitions, approaches and perspectives concerning the email problem. Yet, everybody agrees to one common thing: too much useless email is circulated. So-called “spammers” abuse the email system by sending enormous amounts of email. Consequently, in order to find out the reasons of the email problem, it should be investigated why the email system is abused.

First of all, it is reasonable to assume that the email system wouldn’t be abused without significant motives from the ones that abuse it. Such motives could be organized in a number of distinct categories, depending on the various roles an email sender (or an agent acting on his behalf) can be associated to. Thus, a sender could be a merchant aiming at finding potential customers for his/hers products. In this case, there are commercial motives for sending such emails. In another case, an email sender could be a person wishing to promote his/hers image or ideas to the public (candidate for elections, celebrity, non-profit organizations, etc). In this case, there are promotional motives for sending spam. Also, email senders could be ill motivated individuals aiming at deceiving their recipients (e.g. spoofing, phishing, chain emails, etc). In this case, there are malicious motives for sending such emails. It is the authors’ belief that the above three kinds of motives apply to the vast majority of senders causing the email problem.

Apart from the ill-motivated senders (i.e. hackers, frauders, etc), many of the so-called spammers aim at locating the minority of recipients that will consider their email as useful (e.g. customers, voters). Thus, most spammers have ordinary promotional, and/or commercial motives. However, the existing email infrastructure does not provide a way to properly satisfy such motives without circulating unsolicited emails in bulk rendering the whole system annoying for its users. At the same time, spammers are also unhappy with today’s reality: due to the aforementioned antispam measures, they have to make special efforts to deliver messages (e.g. disguising messages, and/or sending individual email messages rather than batching them together as bulk messages) before sending them to their recipients.

Maybe, if spammers had the chance to more efficiently direct their emails to the appropriate recipients, their goals would be satisfied and at the same time the email problem would be reduced. It should be noted once again that the above argument does not refer to senders with deceptive motives such as hackers, frauders, e.tc.

Towards a motive-sensitive email system

As stated earlier in this paper, there are three main categories of motives for sending spam, namely commercial, promotional and malicious. Efficient solutions to the email problem should take into account such motives that practically represent the spammers perspective to the email problem.

Emails deriving from spammers with malicious motives can only be dealt with through appropriate suppression actions (e.g. laws, punishment, etc). On the other hand, it is argued that promotional and/or commercial motives should not be considered as punishable actions, despite the fact that they are greatly responsible for today's decline of the email system. In this context, it might be possible to satisfy the requirements of spammers with promotional and/or commercial motives without putting additional strain on the email community (i.e. email receivers, ISPs, etc). For example, email senders could be provided with the opportunity to tag the context of their emails according to a well-established taxonomy of email categories (e.g. food, commercial, leisure, etc). At the same time, receivers could be provided with the opportunity to decide a priori what kind of emails to receive. Such decisions could be based on their personal preferences.

Solutions to the email problem should not just focus on fulfilling receivers' needs that consider spam as useless. It is equally important to take into account the minority of receivers that do not consider such emails as useless. Ultimately, such solutions would result in much less spam originating from a minority of ill-motivated spammers.

Finally, it is argued that most of the identified causes of the email problem exist due to the lack of respect exercised by a minority of people (i.e. email abusers) towards the broader email community. To that effect, proper education concerning Internet ethics could prevent the spreading of such beliefs and provide a solid background for the forthcoming solutions.

Conclusions

It is common knowledge that a number of serious problems has brought the email system to its knees. Nowadays, the need to act upon such problems is more pressing than ever. Having these thoughts in mind, this paper attempts a comprehensive review of the email status.

Specifically, an attempt is made to define the problem, in order to accurately determine the main aspects of this phenomenon. In this context, it seems that the widely used term "spam" is only partially capable of providing a semantic equivalent to the email problem. A classification of current approaches aiming at offer-

ing solutions is also provided in an attempt to determine their weak spots, which prevent them from being effective. Further analysis takes under consideration the email abuser's perspective. Thus, in order to deal with the email problem, it is proposed that modern email systems should find a way to regulate promotional and/or commercial emails while at the same time eliminate malicious ones.

REFERENCES

- E. Moustakas, C. Ranganathan, & P. Duquenoy** (2005), Combating spam through legislation: a comparative analysis of US and European approaches, in Proc. Second Conf. on Email and Anti-Spam (CEAS), Stanford, 2005; <http://www.ceas.cc/papers-2005/1>, accessed Sept. 10, 2008.
- A. Weiss** (June 2003), Ending Spam's Free Ride, *netWorker*, 7 (no. 2), 18-24.
- T. O'Reilly** (2007), What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software, <http://www.oreillynet.com/lpt/a/6228>, accessed December 29, 2007.
- G. Schryen G.** (2007), *Anti-Spam Measures: Analysis and Design*, Springer-Verlag New York, Inc. Secaucus, NJ, USA, 7-8.
- G. Schryen** (March, 2007), Anti-spam legislation: An analysis of laws and their effectiveness, *Information & Communications Technology Law*, 16 (no. 1), 17-32.
- A.T. Ramachandran, N. Feamster, & S. Vempala** (2007), Filtering spam with behavioral blacklisting, In Proc. ACM-CCS '07, 342-351.
- P. Judge** (2003), A taxonomy of anti-spam systems, presented at the 56th Internet Engineering task force Conference, Anti-Spam Research Group, San Francisco, 2003; <http://www.ietf.org/Proc./03mar/slides/asrg-6/>, accessed Sept., 2008.
- G. Lawton** (Nov. 2005), E-Mail Authentication Is Here, but Has It Arrived Yet?, *Computer*, 38 no. 11, 17-19.
- E. Levy** (July-August, 2003), The Making of a Spam Zombie Army: Dissecting the Sobig Worms, *IEEE Security and Privacy*, 1 (no. 4), 58-59.
- W. S. Yerazunis** (2004), The Spam Filtering Accuracy Plateau at 99.9 Percent Accuracy and How to Get Past It, in Proc MIT Spam Conference, Cambridge.
- G. Hulten, A. Penta, G. Seshadrinathan, & M. Mishra** (2004), Trends in Spam Products and Methods, in Proc. First Conf. on Email and Anti-Spam (CEAS 2004), Mountain View, CA, 2004, <http://www.ceas.cc/papers-2004/165.pdf>, accessed Sept. 10, 2008.

D. Lowd & C. Meek, (2005), Good word attacks on statistical spam filters, in Proc. Second Conf. on Email and Anti-Spam (CEAS 2005), Stanford, 2005; <http://www.ceas.cc/papers-2005/125.pdf>, accessed Sept. 10, 2008.

N. Lugaresi (2004), European Union vs. Spam: A Legal Response, in Proc. First Conf. on Email and Anti-Spam (CEAS 2004), Mountain View, CA, 2004, <http://www.ceas.cc/papers-2004/145.pdf>, accessed Sept. 10, 2008.

D. Fingerman (2004), Spam canned throughout the land?, Summary of the CAN-SPAM Act, *Journal of Internet Law*, 7 (no. 8), 11-17.

N. Leavitt (March, 2007), Vendors Fight Spam's Sudden Rise, *IEEE Computer*, 40 (no. 3), 16-19.

J. Soma, P. Singer, & J. Hurd (2008), Spam Still Pays: The Failure of the CAN-SPAM Act of 2003 and Proposed Legal Solutions, *Harvard Journal on Legislation*, Harvard College, 45 (no. 1), 165-198.

A. Gray, and M. Haahr (2004), Personalised, Collaborative Spam Filtering, in Proc. First Conf. on Email and Anti-Spam (CEAS 2004), Mountain View, CA, 2004, <http://www.ceas.cc/papers-2004/132.pdf>, accessed Sept. 10, 2008.

J. Goodman, G.V. Cormack, & D. Heckerman (2007), Spam and the ongoing battle for the inbox, *Communications of the ACM*, 50 (no. 2), 25-33.

S. Cobb (2008), The Economics of Spam, http://www.spamhelp.org/articles/economics_of_spam.pdf accessed Oct. 03, 2008.

C. Dwork & M. Naor (1993), Pricing via Processing or Combatting Junk Mail, In Proc. CRYPTO'92, 1993, 137-147.

T. Loder, M. V. Alstyne, and R. Wash (2004), An economic answer to unsolicited communication, in Proc. of the 5th ACM conference on Electronic Commerce (EC'04), 2004, 40-50.

S. Hansell (Feb. 2, 2004), Gates Backs E-Mail Stamp in War on Spam, *The New York Times*.

C. Pu, and S. Webb (2006), Observed trends in spam construction techniques: A case study of spam evolution, in Proc. Second Conf. on Email and Anti-Spam (CEAS 2006), California USA.

F.D. Garcia, J.H. Hoepman, & J. van Nieuwenhuizen (2004), Spam Filter Analysis, Security and Protection in Information Processing Systems, IFIP International Federation for Information Processing, Springer, 147, 395-410.

P. Graham (2007), A Plan for Spam, <http://www.paulgraham.com/spam.html>, accessed Apr. 12, 2007.

W. Yih, J. Goodman, & G. Hulten (2006), Learning at Low False Positive Rates, in Proc. Second Conf. on Email and Anti-Spam (CEAS 2006), California USA.

A Debordian analysis of Facebook

Antonio Marturano*

Faculty of Economics, Sacred Heart Catholic
University of Rome (Italy)

&

Sergio Bellucci**

President, Net Left and CEO, LAIT (Italy)

Abstract

Facebook, the second largest social network on the Web with around 60 million members, is one of the fastest-growing and best-known sites on the Internet today. With the U.S. now accounting for only about a third of all Facebook users, we are starting to see a gradual shift away from its original demographic of college-age users. Very surprisingly, indeed, in the past months Facebook has been literally invaded by Italians, Which is the reason for this huge success of Facebook? One of the reasons is that clearly young Italians' discontent (as it is young people who mainly inhabit Facebook) and frustration with the current political situation and with their political representatives is finding in the Web a channel to let youth voice be heard. Facebook is also a media for channelling Italians' emotions, self representation, and symbolic environment at the same speed of their telefonino (mobile phone): indeed Facebook not only provides multimedia content and a high interactive environment, but it also provides personalised features. In other words, it is my personal content which is available on the web and

* *Antonio Marturano* is adjunct professor at the faculty of Economics, LUISS University and Sacred Heart Catholic University of Rome. He teaches Business Ethics and Leadership. Antonio is an INSEIT founder and chair elected for the Leadership Education ILA MIG. He writes extensively about media and computer-related ethical problems and leadership ethics. Antonio published on Telos, Ethics and Information Technology (ICES), Business Ethics, European Business Forum, Bioethics, and Philosophy of Management. He also published (with Jonathan Gosling, eds.), *Leadership: The Key Themes*, London: Routledge, 2008 and *Etica dei Media*, Milan: Franco Angeli, 2000.

** *Sergio Bellucci* Sergio Bellucci is CEO of LAIT SpA (Lazio Innovation Technology). He was President of CEO of Liberation. He wrote for years on issues of communication and its impact in social and productive system. He published: «Internet, democrazia e politica» Seam (1997), «Dalla meccanica sociale alla termodinamica del collettivo» ManifestoLibri (1998), «E-Work» Derive & Appodi (2005), «L'agire del conflitto. Biopolitica e Mutazione» Mimesis (2008). Being published «The spectrum of the Capital», Codice Edizioni.

it make me feel as if I was in the centre of a virtually worldly networked stage. We will argue Facebook is realizing what Guy Debord calls “the invasive forces of the ‘spectacle’ - “a social relation between people that is mediated by images”: Facebook is seen as an alternative tool able to amplify an individual’s alienation and narcissism, which, are a consequence of the mercantile form of social organization which has reached its climax in capitalism. Under marxist theory Facebook doesn’t appear what Jaron Lanier claims to be collaborative communities. We finally argue that Facebook is not (as Tapscott and Williams claim) a promising example of a new shift from capitalism to a new form of economy based on openness, peering, sharing and global action – which they called Wikinomics; but rather new disguised forms of advanced capitalism aimed at eroding space to more challenging modes of Internet collectivism.

The Facebook Phenomenon

Facebook, the second largest social network on the Web with around 60 million members, is one of the fastest-growing and best-known sites on the Internet today. With the U.S. now accounting for only about a third of all Facebook users, we are starting to see a gradual shift away from its original demographic of college-age users (18-25): 46% of all users are 18-25 years old, down from 51% in late May 2008. The number of users in the 18-25 segment is growing, but at a slower pace than the other age groups. Among the major Facebook age segments, the fastest growing are teens (13-17) and young (26-34) to middle-age (35-44) professionals, with the growth in teens driven by non-U.S. markets. Also it is worth noting the strong growth in the much smaller 45-54 and 55-59 age groups (Ben Lorica, 2008). Looking closely at the top 30 countries, a few European countries have grown more than ten percent over the last four weeks 2008 (France, Spain, Germany, Italy), with France having the most number of users (approx. 2.5 million). Italy, in particular, is a country which still has one of the lowest rates of Internet use in Europe (35.6% according to a 2006 Istat report); but, very surprisingly, in the past months Facebook has been literally invaded by Italians, quickly helping Italy reach first place for the greatest (and fastest) exponential growth in adoption of Facebook by a country (see box below). Italians seem to have a natural affinity with Facebook – they are not only joining in huge numbers (Facebook is now the fifth most popular site in Italy) but they seem to have seamlessly integrated this technology in their everyday life: Facebook is fast becoming the most used accessory in their beloved “telefonino” (mobile phone) (see Di Gennaro, 2008).

Month	FB Subscribers #
January 08	216000
February	238000
March	265000
April	305000
May	355000
June	465000
July	573000
August	622420
September	1294000
October	2215000
November	4152380
December	5587060
January 09	6481280
Number of Italian subscribers to Facebook 2008-2009. Source: http://manyeyes.alphaworks.ibm.com/manyeyes/datasets/facebook-users-in-italy-january-2009/versions/1	

Which is the reason for this huge success of Facebook? One of the reasons is that clearly young¹ Italians' discontent (as it is young people who mainly inhabit Facebook) and frustration with the current political situation and with their political representatives is finding in the Web a channel to let "youth" voice be heard. In fact not only Italian media are not giving enough voice to the younger Italian generations, but also Italy is actually experiencing a strong lack of generational replacement in its leadership (it is worth notice that the average age of Italian leadership is around 70 years old) (See II Rapporto LUISS, 2008). The main channel of information in Italy is still the obsolete TV broadcasting: Italians indeed like to get multimedia information (mainly images) and almost stream chats, rather than word-content information, even though the kind of information supplied is not networked, but broadcasted: few or no interaction is allowed. On the contrary, Facebook not only provides multimedia content and a high interactive environment, but it also provides personalised features.

The most active people on Facebook are those already having a public exposure such as politicians, writers (journalists and novelists), and TV entertainers or actors not fully engaged with TV shows. We have used Facebook from October 2008 to April 2009. We have observed people interacting with us, and have taken notes

1. "Young people" here means, ironically, people less than 50 years old.

about their behaviour by means of ethnographic methods. In particular, We have observed behaviour of the following very active persons in my friends list²:

- a successful writer (RC)³;
- a major of an important Italian city (FZ);
- a journalist (GF);
- a popular TV entertainer and actor (TG).

All of them had a quite huge number of personal photos on display in their Facebook personal webpage and a large cohort of friends (more than 1000). Their behaviour on Facebook is quite different. RC for example exhibited a quite narcissistic behaviour based on continuous search of agreement and acknowledgment of his popularity writing down on his personal page first parts of his novels which “the audience” had to finish according to the actual novel. Usually the audience would not only complete the requesting piece but also added their (favourable, of course!) impressions and feelings. Major FZ on the contrary, used to display his personal everyday life to the public through several everyday pictures made using his mobile phone and tagging persons pictured in his photo⁴; he not only used to tagged people but also used Facebook along the lines of a poll, asking to citizens their opinions about a regulation proposal, an issue about his community or a recently issued law to test trends among his citizens. The actor and TV entertainer TG raised questions among his friends on political views and opinions about recent problems or facts. His strategy is to propose his own opinion or idea and to whether such idea is shared by his friends; when ideas were challenged by other persons through rational arguments or by ideology TG used many deceptive techniques (such as Ad Hominem, Red Herring, Asking for Support, and Appeal to Ridicule) to induce his opponent into error, to disappear from such discussion, or to change his opponent mind. Journalist GF, used Facebook for self-advertisement and for discussing about topics related to his generation (the '68, recent political mysteries, the Italian Red Brigades). All these people, with similar political viewpoints (they are not supporter of Italian Centre-Right Parties) started very harsh debates (with a huge number of discussants that often lasted long time) over Facebook.

Therefore, Facebook appears to be a new kind of arena, apparently open for debate but led by new technological opinion-makers (people who are already opin-

2. My friends list was populated by 168 person by 14.04.2009.

3. I will only use their initials for protecting their privacy.

4. In Italy and EU legislation is still matter of debate whether tagging is a breaking of personal privacy.

ion makers already because they have had popular prominence) who use Facebook to amplify their narcissism and which relegate other less popular individuals to gregarious or even passive audience roles. Famous people do not only become the center of attention by means of their personal content available in their Facebook profiles. In particular photo sharing between friends and virtual gifts sharing let them feel in the centre of their symbolic, networked, virtual stage. Facebook, in this very sense, realizes what Guy Debord calls “the invasive forces of the ‘spectacle’ - “a social relation between people that is mediated by images” (Debord, 1992) or the last ultra-capitalistic Trojan horse (that is using an obsolete cold war logic) to stop the new peer-sharing economy envisaged by open source methodologies. In the following paragraph we will first explain some concepts useful to understand the framework in which the concept of “spectacle” has originated before trying to parallel the old logic of “spectacle” with a new, more refined logic hidden into the networked technologies.

Marxist Theorists

Guy Debord is one of the most known French situationist Marxists. We will analyse the Facebook phenomenon in Italy with the help of his fundamental concepts regarding the society of spectacle. In the next paragraphs we will first analyse some fundamental tools from authors who started their analysis of contemporary culture by means of Marxism: such as the Frankfurt School (notably H. Marcuse), Post-modernism school (such as Lyotard and Baudrillard) and in the last paragraph we will deal with Debord’s Situationism. Debord’s theory cannot indeed be fully understood if not placed within the context of Marxist theories. Marxist theories too will offer in turn some useful tools to better understand computer mediated communications (CMCs), and in particular Facebook.

H. Marcuse: A general criticism of technology

H. Marcuse was a very trendy philosopher at the beginning of the 70s: his works such as *One-Dimensional Man* (1964) were largely read and discussed, and were the conceptual framework for ‘68 students’ movements. According to Marcuse the typical modern trait of contemporary civilizations is the way in which it is able to suffocate those needs which should be freed and at the same time it breeds and forgives rich societies destructive power and their repressive function both in west than in east countries. In a similar way, technology would have helped to reduce timework (that is – in Marxist terms – alienation) but actually (1) it has increased an individual’s production potentiality and (2) at the same time technology has created a diffusive, high level system of control over humans in which is no possible real criticism neither opposition to central power. Such control is not violent and dictatorial (such as in Nazism

and Fascism), but “a comfortable, smooth, reasonable, democratic unfreedom prevails in advanced industrial civilization, a token of technical progress.” (Marcuse, 1964, ch.1).

Very importantly, Marcuse continues:

“In the face of the totalitarian features of this society, the traditional notion of the «neutrality» of technology can no longer be maintained. Technology as such cannot be isolated from the use to which it is put; the technological society is a system of domination which operates already in the concept and construction of techniques. The way in which a society organizes the life of its members involves an initial choice between historical alternatives which are determined by the inherited level of the material and intellectual culture. The choice itself results from the play of the dominant interests. It anticipates specific modes of transforming and utilizing man and nature and rejects other modes. It is one “project” of realization among others. But once the project has become operative in the basic institutions and relations, it tends to become exclusive, and to determine the development of the society as a whole.” (Marcuse, Id., Intro).

In other words, no opposition is possible within this kind of society; but whenever such an opposition would arise, it would be de facto outside any symbolic universe accessible to humankind and therefore rejected as alien.

“As the great words of freedom and fulfillment are pronounced by campaigning leaders and politicians, on the screens and radios and stages, they turn into meaningless sounds which obtain meaning only in the context of propaganda, business, discipline, and relaxation. This assimilation of the ideal with reality testifies to the extent to which the ideal has been surpassed. It is brought down from the sublimated realm of the soul or the spirit or the inner man, and translated into operational terms and problems. Here are the progressive elements of mass culture. The perversion is indicative of the fact that advanced industrial society is confronted with the possibility of a materialization of ideals. The capabilities of this society are progressively reducing the sublimated realm in which the condition of man was represented, idealized, and indicted. Higher culture becomes part of the material culture. In this transformation, it loses the greater part of its truth (Marcuse, op.cit. §3).

J.F. Lyotard: Information as commodity, Information and exchange value

In *The Postmodern Condition* (1979), the French philosopher and literary theorist, J.-F. Lyotard – who coined the word postmodern – claimed that

“... technological transformations can be expected to have a considerable impact on knowledge. Its two principal functions – research and the transmission of acquired learning-are already feeling the effect, or will in the future. With respect to the first function, genetics provides an example that is accessible to the layman: it owes its theoretical paradigm to cybernetics” (Lyotard. Id., ch.1).

In particular, digital technological transformations will lead to a change into the humankind cognitive sphere. That is the first important step in order to producing and re-producing the reality. Everything which is not translatable through the logic of the digital will be marginalised and time by time expunged from the social body. According to Lyotard, indeed,

“The nature of knowledge cannot survive unchanged within this context of general transformation. It can fit into the new channels, and become operational, only if learning is translated into quantities of information. We can predict that anything in the constituted body of knowledge that is not translatable in this way will be abandoned and that the direction of new research will be dictated by the possibility of its eventual results being translatable into computer language. The “producers” and users of knowledge must now, and will have to, possess the means of translating into these languages whatever they want to invent or learn. Research on translating machines is already well advanced. Along with the hegemony of computers comes a certain logic, and therefore a certain set of prescriptions determining which statements are accepted as “knowledge” statements” (cit.).

In other words, the very nature of digital technologies is based on the logic of simulation (or as-if model) which becomes fundamental for digital rationality. No analogical information is provided into the digital, therefore, digital information will always be operationalised under a “simulation” or “self reproduced symbolic universe” form. No other levels of realities are accessible outside this paradigm. Lyotard concludes:

“We may thus expect a thorough exteriorisation of knowledge with respect to the “knower,” at whatever point he or she may occupy in the knowledge process. The old principle that the acquisition of knowledge is indissociable from the training (*Bildung*) of minds, or even of individuals, is becoming obsolete and will become ever more so. The relationships of the suppliers and users of knowledge to the knowledge they supply and use is now tending, and will increasingly tend, to assume the form already taken by the relationship of commodity producers and consumers to the commodities they produce and

consume – that is, the form of value. Knowledge is and will be produced in order to be sold, it is and will be consumed in order to be valorised in a new production: in both cases, the goal is exchange.”

J. Baudrillard: Simulacra and Consumption

The concept of simulation (that is charged with symbolic meanings, the very fabric of reality) is now embedded in a digital technology context. Simulation, in such technological context, becomes a simulacrum. According to J. Baudrillard, modern society has replaced all reality and meaning with symbols and signs, and that the human experience is of a simulation of reality rather than reality itself. The simulacra that Baudrillard refers to are signs of culture and media that create the perceived reality; Baudrillard indeed believes that society has become so reliant on simulacra that it has lost contact with the real world on which the simulacra are based:

“An alter-action which tends to diminish with increasing information and which will, in the end, be eliminated by absolute information; the world’s equivalence to the world – the final illusion, that of a world which is perfect, fully realized, fully effectuated, a world which is consummated and has attained the height of existence and reality, and also the furthestmost extent of its possibilities. It is God (this we cannot hide) who stands at the end of this process of increasing information and complexification, of verification of the world in real time. It is God who presides over this dissolution of the world as illusion and its resurrection as simulacrum and virtual reality, at the end of a process of extenuation of all its possibilities by the real” (J. Baudrillard, 1996, p.8).

In other words, media and digital technology, because they are based on imagery (which is not just a .jpg photo format, but mainly a symbolic universe conveying a “picture” of reality and/or even an individual’s reality), have absorbed the actual reality. Baudrillard continues,

“So long as an illusion is not recognized as an error, its value is exactly equivalent to that of reality. But once the illusion is recognized as such, it no longer is one. It is therefore the concept of illusion itself, and this alone, that is the illusion”(Baudrillard, op.cit., p. 51).

In this way digital technologies take reality’s place creating a destabilizing short-circuit. A simulacrum is therefore what is generated by such short-circuit: something which is not constructed according to truth and falsity. Digital technology itself is indeed creating reality. Baudrillard (1985) theorizes the lack of distinc-

tions between reality and simulacra originates in several phenomenon; in particular, in contemporary media including television, film, and the Internet, which are responsible for blurring the line between goods that are needed and goods for which a need is created by commercial images. On the contrary, we will see in the last paragraph that simulacra, finally, is annihilated in the knowledge economics productive lifecycle as commodity and therefore itself become a consumable good.

G. Debord: The society of Spectacle

Debord, as well as Baudrillard, starts his analysis of immaterial societies with Marx's analysis of fetishism of goods. According to Marx a good is a mix between use value and exchange value. While the first is the material consumption of a good, the latter is a good circulation power. In advanced capitalistic societies use value is less important than its exchange value. Any object counts not as such but as a good. A good symbolic characteristic takes advantage, so to speak, over its material characteristics and the very nature of production is changing. Therefore the concept of good becomes even more an abstraction. Marx attributes goods a fetish characteristic or even a magical one: similarly to religion in which gods socially created become others than humanity, become aliens and even human-independents.

Debord, agrees with Marcuse (see above), about technology rationality that keeps hostages both political-social model; in this way technological rationality homogenizes everything that it encloses. According to Debord pre-digital analysis such homogenization takes the form of a spectacle:

"In societies dominated by modern conditions of production, life is presented as an immense accumulation of spectacles. Everything that was directly lived has receded into a representation... The society based on modern industry is not accidentally or superficially spectacular, it is fundamentally spectaclist. In the spectacle — the visual reflection of the ruling economic order — goals are nothing, development is everything. The spectacle aims at nothing other than itself ... The spectacle is capital accumulated to the point that it becomes images. (Debord, cit., §1)

For Debord, spectacle is not what we see through media (that is what in English we mean with "show"), he says:

"The spectacle is not a collection of images, rather, it is a social relationship between people that is mediated by images." (op.cit.)

There are three kinds of "spectacles": concentrated spectacle, that

“belongs essentially to bureaucratic capitalism, even though it may be imported as a technique of state power in mixed backward economies or, at certain moments of crisis, in advanced capitalism. In fact, bureaucratic property itself is concentrated in such a way that the individual bureaucrat relates to the ownership of the global economy only through an intermediary, the bureaucratic community, and only as a member of this community. Moreover, the production of commodities less developed in bureaucratic capitalism, also takes on a concentrated form: the commodity the bureaucracy holds on to is the totality of social labor, and what it sells back to society is wholesale survival. (cit., ch.1 §64).

The diffused spectacle, on the other hand,

“accompanies the abundance of commodities, the undisturbed development of modern capitalism. Here every individual commodity is justified in the name of the grandeur of the production of the totality of objects of which the spectacle is an apologetic catalogue. Irreconcilable claims crowd the stage of the affluent economy’s unified spectacle; different star-commodities simultaneously support contradictory projects for provisioning society: the spectacle of automobiles demands a perfect transport network which destroys old cities, while the spectacle of the city itself requires museum-areas.” (cit., Ch.1 §65).

A third form of spectacle, Debord concludes, has been established, through the rational combination of these two, and on the basis of a victory of the form which had showed itself stronger: the diffuse. This is the integrated spectacular, which has since tended to impose itself globally. The integrated spectacular

“shows itself to be simultaneously concentrated and diffuse, and ever since the fruitful union of the two has learned to employ both these qualities on a grander scale. Their former mode of application has changed considerably. As regards the concentrated side, the controlling center has now become occult, never to be occupied by a known leader, or clear ideology. And on the diffuse side, the spectacular influence has never before put its mark to such a degree on almost the totality of socially produced behavior and objects. For the final sense of the integrated spectacular is that it integrates itself into reality to the same extent that it speaks of it, and that it reconstructs it as it speaks. As a result, this reality no longer confronts the integrated spectacular as something alien” (Debord, 1967, § IV).

Conclusions

Facebook is an alternative tool able to amplify an individual's alienation and narcissism, which, according to Debord, are more than an emotive description or an aspect of individual psychology. Rather, they are a consequence of the mercantile form of social organization which has reached its climax in capitalism. The development of modern society in which authentic social life has finally been replaced with its whole representation: "All that was once directly lived has become mere representation" (Debord, id.); life is actually meaningful life if and only if it is described and shared on Facebook". Debord argues that the history of social life can be understood as "the decline of being into having, and having into merely appearing" (Debord, id.) which is finally concluded with Facebook, in which private and public spheres are finally melted together. This condition, according to Debord, is the "historical moment at which the commodity completes its colonization of social life" (Debord, id.). In a similar way, Jaron Lanier claims that similar collaborative communities such as Flickr, MySpace, and Wikipedia represent a new form of "online collectivism" that is suffocating authentic voices in a muddled and anonymous tide of mass mediocrity (quoted in Tapscott and Williams, 2008). In these cases such as Facebook and MySpace we cannot follow Howard Rheingold when he claims that "Collectivism involves coercion and centralized control; collective action involves freely chosen self-selection and distributed coordination" (quoted in Tapscott and Williams, 2008).

The Italian example shows how Facebook (but also MySpace, and even YouTube) cannot be compared – as Tapscott and Williams (op. cit.) claim – to other open-source, collaborative and participative endeavors that Tapscott and Williams call "Wikinomics". They rather are – according to my debordian analysis – the latter development of advanced capitalism, leading to individuals' further alienation and narcissism by means of social relations between people mediated by (self-created) symbolic imaginary. It is not enough that Facebook is opening its platform to users and external developers when most people are, as in Italy, still using Facebook just to join networks, and to connect and interact with other people or adding friends and send them messages, and update their personal profile to notify friends about themselves. Directly parallel to Marx's notion of commodity - for Debord (op.cit.) – the spectacle (and in our particular case Facebook) making relations among people seem like relations among images/symbols (and vice versa). The spectacle is the form taken by society once the instruments of knowledge production have become wholly commoditized and exposed to free circulation. To sum up, many cases provided by Tapscott and Williams (op. cit.) (such as Facebook, MySpace, Flickr, Second Life and similar) do not promise any new shift from capitalism to a new form

of economy based on openness, peering, sharing and global action – what they called Wikinomics; but they must carefully be distinguished from genuine new forms of peer economy such as Open Source or Wikipedia. Facebook and similars are rather new disguised forms of advanced capitalism that is trying to colonize the Internet by reducing the networking model to a more controllable (by old advanced capitalism) broadcasting model. Facebook model can be labelled, using Debord's terminology, as distributed spectacle where the spectacle function is shared or distributed amongst those with the ability and experience necessary to ensure the function is carried out to the benefit of the most traditional organisations.

Such a distributed view of spectacle seems to be an attempt to colonize the web using an hidden form of integrated spectacle (based on the broadcasting model), into an intricate and complex web of spectacles (working on the network model), which appear to be an integrated spectacle in which simulacra of individuals become consumable goods into a capitalistic logic rather than being a new way for collaborative efforts.

REFERENCES

- J. Baudrillard**, 1985, *Simulacres et Simulation*, Paris, Galilée
- J. Baudrillard**, 1995, *Le Crime Parfait*, Paris, Galilée
- G. Debord**, 1992, *Society of the Spectacle*, OakLand: AKPress
- G. Debord**, 1967, *Comments on The society of Spectacle*, London: Verso
- C. Di Gennaro**, 2008, "How Facebook is changing Italian social and political life", <http://blogs.law.harvard.edu/idblog/2008/11/11/how-facebook-is-changing-italian-social-and-political-life/>
- B. Lorica**, 2008, "Facebook Growth By Country and the Slowdown in App Usage", <http://radar.oreilly.com/2008/07/facebook-growth-by-country-and.html>
- II Rapporto LUISS**, 2008, *Generare Classe Dirigente (Generating a Leading Class)*, Milano: Il Sole24Ore
- J.-F. Lyotard**, 1979, *La Condition Postmoderne*, Paris: Les Editions de Minuit
- H. Marcuse**, 1964, *One-Dimensional Man*, Boston:Beacon
- D. Tapscott & A.D. Williams**, 2008, *Wikinomics 2.0*, London: Penguin

The Commodification of the Individual in the Internet Era: Informational Self-determination or “Self-alienation”?

Lilian Mitrou*

University of the Aegean, Department of Information
and Communication Systems Engineering

Abstract

This contribution deals with the commodification of individuals in cyberspace. I discuss questions of informational privacy and the “alienability” of personal information. Privacy is perceived not as merely a right possessed by individuals but as a prerequisite for taking autonomous decisions, freely communicating with other persons and being included in a participation society. I examine the dignity-autonomy approach, which regards privacy as an inalienable individual right and the liberty-property approach, which treats personal information as a tradable property object. Emphasis is given on the concerns related to «Consent fallacy», i.e. the (in)ability of the individual-websurfer to form and express free, conscious and informed choices. I conclude by considering privacy as an inalienable right because one can never be permitted to freely dispose of one’s dignity and, moreover, dignity and privacy are inherent elements of a democratic constitutional order.

Keywords: Information privacy, dignity, property, autonomy, commodification, consent, alienability.

The commodified Websurfer

The Internet reveals information technology’s great promise, which is to form new links between people. People, increasingly, live their every-day lives in cyberspace. They run their errands and conduct their business online. They social-

* Dr. Lilian Mitrou is Assistant Professor at the University of the Aegean-Greece (Department of Information and Communication Systems Engineering) and Visiting Professor at the Athens University of Economics (Postgraduate Studies Program). She teaches information law and data protection law. She has served as a Member of the Hellenic Data Protection Authority (1999-2003) and as Advisor to the former Prime Minister K. Simitis in sectors of Information Society and Public Administration (1996 - 2004). She served as member of many Committees working on law proposals in the fields of privacy and data protection, communications law, e-government etc. Her research interests include: Privacy and Data Protection; e-Democracy and eGovernment services; Internet Law. L. Mitrou published books and chapters in books (in Greek, German and English) and many journal and conference papers.

ize with their friends in the virtual world and they are building social networks. People use the Internet to exchange messages, receive and produce knowledge, built up their participation to social and political life, play, purchase and exchange goods and services. The Internet is growing at a rate that outpaces any modern medium for communication. But exactly the aspects of the Internet that make it such a powerful information and communication medium all combine to transform it into a fertile ground for collecting personal data about surfers/users [Froomkin (2000)]. Technology in cyberspace affects privacy in ways that are dramatically different from anything previously possible [Schwartz (2000)].

The technical infrastructure of the Internet combined with ever-advancing computer technology make it easy and cheap to collect, combine and use vast amounts of personal data, i.e. information relating to an identified or identifiable natural person. Web sites, on line businesses and Internet Service Providers (hereafter, 'ISPs') providing goods, services and information on-line, typically collect, process and store immense amounts of information about consumers, without that consumers necessarily either knowing data is collected, or consenting to that collection. E-commerce is not, or -in any case- less, anonymous than the traditional commerce, since online merchants need to collect personal information such as names, addresses and credit card numbers.

Personal information collected falls broadly into two categories: "visibly" collected data and "invisibly" collected data. Most of the recent controversy around data collection on the Net has, however, revolved around «invisible» collection of data, held via devices such as "cookies" or Web beacons. Here both collection and use of data is non-transparent. ISPs can –at least in theory – amass complete data on every website a subscriber has visited. Suppliers of third party cookies, such as the online advertiser Double Click, can also collect comprehensive profiles of consumers' online activity across multiple websites. However, in many cases it is the surfer who supplies information or makes it public on the Internet. Many sites require some kind of registration, usually by an online form, before the consumer can start interacting with the site. Users who desire to use Google Analytics' Web-traffic monitoring service must submit specific personal data, such as user's phone number or country of residence, although it is not necessary in order to provide the service. This information serves as a prerequisite to account creation. The newly created account allows access to the desired service - "in this case, to valuable technology that tracks activity on an account holder's personal or business Web site"[Ciocchetti (2008)]. Such information is at least transparent in its collection, but not always in its subsequent use.

The more the people rely on the Web for their trading, shopping, communications and relationships the more likely it becomes that data about their interests,

preferences and behaviour are captured. One of the major reasons for the ever expanding processing of personal data is their commercialisation. The commercialisation really climaxed as soon as Internet changed from an - elitist - forum into information "bazaar". Any discussion of benefits and risks pertaining to the collection of personal information must recognize that the Web functions as the primary tool of e-commerce precisely also because of such collection [Ciocchetti (2008)]. In this expanding and borderless information market, the individual is both user and content-provider. Users develop inevitably into an inexhaustible source of personalised information. Personal data - such as names, addresses, ages, marital status and family, employment and income status, shopping habits, websurfing habits, nationality e.tc. - is of enormous commercial value, particularly when used to create profiles about "typical consumers".

Advanced technologies of data mining increasingly allow new information, both about individuals and collective populations, to be generated from personal information databases. Companies aggregate and mine information for customer preferences and then tailor marketing efforts specifically towards the consumer's interests. The newest trend in tailored marketing is the so-called "behavioral targeting", a concept where companies collect and monitor the external Web sites that their customers visit in order to develop behavioral advertising. Behavioral targeting is a method of tracking the online behavior of Internet users in order to serve those consumers with advertising targeted to the specific interests "expressed" through web-browsing activity [Hotaling (2008)].

With the growing importance of various "personalization services", it is clear that personal data and individual user profiles become the key instrument in realizing returns on the investment [Ciocchetti (2008)]. Companies view this information as a corporate asset and they invest heavily in software applications that facilitate the collection of users information. Personal data is one of the biggest (or sole) assets of "dot.com" companies, which often have little in the way of fixed assets or capital [Edwards (2003)]. With a few mouse clicks and a plethora of available buyers such companies generate additional revenue streams.

The growing commercialization of personal data has intensified the commodification of the individuals. They are increasingly viewed as the sum of a variety of data that can be arbitrarily aggregated or taken apart and profitably marketed. In the Moore case the controversy was not a dispute over the preservation of the plaintiff's integrity by securing the inaccessibility of his genetic data and at best allowing their use for purely therapeutic or scientific research purposes. The controversy concerned the conditions of their sale and the allocation of the gains [Simitis (1999)]. The Toysmart bankruptcy case was a far from unique example of the phenomenon that large amounts of personal data changed hands or "own-

ership", as part of merger-acquisitions, reorganizations and other strategic company movements. Customer databases figured as a controversial asset in the many dot.com bankruptcy proceedings that have fallen out after the bursting of the dot.com bubble in 2000/01 [Edwards (2003)]. The 2001 takeover by the American company eBay of the French auction sales operator iBazar is but one example of what is at stake when it comes to the acquisition of subscribers', users' and customers' lists [Gauthronet (2001)]. In the case of Google's acquisition of Double Click, a deal approved by the US Federal Trade Commission, the world's largest and most popular search engine is integrated with services from the world's largest online advertising firm.

When data typically collected from website browsers is combined with information held in various databases, then not only can its value increase dramatically, but the risks to personal privacy increase correlatively [Edwards (2003)]. Never before hence the behaviour of the individuals was so closely observed and recorded, the attempts to expand the use of the data collected so persistent, the proliferation of ever more detailed personal profiles so intensive [Simitis (1999)]. For individual consumers it is no longer possible to really find out what happens to their personal data, without referring to their ability effectively control the uses of these data.

Commodified personal data is a discrete package of personal information that can be exchanged for something else [Schwartz (2004)]. Many individuals understandably try to profit "from what is left of their privacy"[Prins (2006)]. Google Gmail initiative offered greater storage space in return for having Google monitor email and using the information for advertising. As Ciocchetti (2008) points out, even privacy-sensitive e-consumers appreciate the value these services provide and not rarely concede that most of the personal data collection is a price to pay in return for convenience and benefits provided. Individuals already participate in the commodification of their personal data [Schwartz (2005)]. The more however the commodification spreads and becomes a - possibly annoying but nevertheless - normal aspect of everyday life, the bigger the risk that an already clearly discernible modification in the perception of privacy and personal data protection may prevail.

Privacy, dignity and autonomy

Prima facie, privacy seems to be universally recognized as an individual right of philosophical and moral origins. However, it is easier to defend privacy than to describe it. The notion and content of privacy are often debatable. There are mainly two different ways to address privacy issues in the Western cultural and legal tradition. Continental privacy protection regimes are, at their core, a form

of protection of a right to respect and personal dignity. It has become common for Europeans to maintain that they respect a “fundamental right to privacy and/or (personal) data protection” [Whitman (2004)]. By contrast the other approach, which is mostly typical of the USA, is much more oriented toward values of liberty, and especially liberty against the State. At its conceptual core the American right to privacy still takes the form that it took in the 20th century. It is the right to freedom from intrusion by the State.

Traditionally, rights have often been understood as protecting the individual against the incursion of the community, based on respect for the individual's personhood or autonomy [Solove (2001)]. This understanding is strongly influenced by the approach of Kant, who attributed the dignity of human beings to their moral autonomy and concluded that they must be treated as “ends” and not as “means” [Neuman (2000)]. The heart of the modern liberal political project lies in the effort to promote human dignity by establishing the conditions for personal autonomy and self-determination [Kohler (2000)]. Dignity has become in the meantime a universal, fundamental and inescapable term of reference even though it should always be seen against the specific cultural and historical background [Rodota (2004), Litman (2000)]. In its landmark *Census Case* Decision on census law the German Federal Constitutional Court emphasized that the constitutional order focuses on the “value and the dignity of the person” [Simitis (1984)] while recent, highly ambiguous instruments, as the EU's Charter of Fundamental Rights, have brought about the constitutionalisation of the person starting from personal freedom and dignity.

Forty years in advance of the famous and passionate essay by Warren and Brandeis, Robert Kerr was referring to the “right to be left alone” and founded that the essence of privacy consisted in “mutual respect and intimacy” [Rodota (2004)]. The claim for (informational) privacy or data protection can be traced back to the concept of dignity. Dignity as related to privacy summarizes principles such as recognition of the personality and non-objectification: an individual cannot be viewed as the sum of a variety of data that can be ad libitum assembled or taken apart and profitably marketed [Simitis (1999)]. The collection and processing of personal data, especially without knowledge or consent of the person, irrespective of concrete “harm”, is an assault on the dignity of the person and a breach of her informational privacy rights [Edwards (2003)].

Informational privacy indicates much more as informational seclusion, a refuge for the individual. Such an approach would, in addition, serve as an incomplete constitutional paradigm in the online world and age [Schwartz and Reidenberg (1996)]. Informational privacy, which encompasses the claim for exercising control over one's own information [Westin (1967)] or – to express it in an

eurocentric way – the right to informational self-determination [Simitis (1984)], rests on the premise that information about ourselves is something over which individuals may exercise autonomy. Informational privacy responds to the requirement that everyone should be in control of the information concerning her so as to formulate conceptions of self, values, preferences, goals and to be protect her life choices from public control, social disgrace or objectification. Informational privacy offers safeguards to preserve an underlying capacity for autonomous decision - and choice-making.

Dignity and inalienable rights residing with the individual are the hallmarks of the European regulatory approach. Continental law has resisted the notion that one can definitively alienate one's dignity [Whitman (2004), Weichert (2004)]. This approach is expressed through the Convention 108 of the Council of Europe as well as the Data Protection Directive 95/46/EC requiring the EU Member States to adopt conforming regulation. The European legislation is comprehensive in that it refers to the collection and processing of all types of data capable of identifying a person and is applicable both to the public and the private sector. Three basic elements constitute the European approach to the protection of personal data: a) the establishment of conditions, obligations and responsibilities for the lawful processing of personal information, b) the maintenance of transparent processing, based not only on the notification system but mainly on the rights of individuals and c) the establishment of external, independent and effective oversight of the data processing activities in the public and private sector.

Despite the recognition of a right to informational privacy in constitutions and international legal instruments texts as well as in – mostly European - detailed laws covering the processing of personal data, day-to-day practice shows that informational privacy appears not at all efficiently protected, especially in the on-line environment. A conceptualization of informational privacy as a right that cannot be infringed appears not to be an undoubtedly dominant scenario in the reality of the on-line world.

Privacy, liberty and property

By visiting websites, creating online accounts or conducting transactions, Internet users disclose personal information in order to gain access to a service or to complete a transaction. Desiring to avoid exclusion from the use of a site, information or service, websurfers enter usually “whatever... information required, click submit and then forget about the process entirely” [Ciocchetti, (2008)] or they make “deals” for the disclosure, collection, use and reuse of their personal data. In certain situations they receive some form of compensation (which may

vary according to the type of data as well as use), and thus “exploit” and “sell” their habits, use-profile and personal data [Prins (2006)].

Individuals may assume as natural, that they do or should “own” data referring to them [Laudon (1996)]. John Locke’s statement that “every Man has a Property in his own Person” has grounded the suggestion that an individual has exclusive rights to the use of his person and can preclude its use or even knowledge of it from third persons [Mell (1996)]. However, at least in the context of European continental legal tradition, the recognition of constitutionally or legally protected rights in information was not deemed as grounded on a perception that individuals have property rights in personal data as such [Whitman (2004) Samuelson (2000)]. Treating privacy as property is a debate initiated primarily in the United States and lasting at least forty years as Westin (1967) supported an enforceable property right [Westin (1967)], a position further analysed by Epstein (1978) and Posner (1978) based on the economic analysis of law approach. With the growing economic importance of services based on the processing of personal data, it is clear that recognizing the right to alienate rights in personal data become a key instrument in realizing returns on the investment [Prins (2006)]. The commodification tendencies, enhanced through the expansion of the Internet, revitalized the discussion about treating personal information as a property object.

The proponents of a proprietary approach translate control of personal information into property concepts. Free information market advocates argue that if a consumer wants to sell or give privacy (rights) and personal data away then she is entitled to do it [Edwards (2003)]. Freedom of alienation is the paramount characteristic of liberal property rights. The argumentation relies on the choice of the individuals: “if they (the consumers) choose not to (prefer dignity), that is evidence that they do not want it in the first place” [Kang and Buchner (2004)]. Individuals can express their preferences about who should be allowed to use their data. Whereas technical infrastructure makes it remarkably simple and inexpensive for businesses to collect and use substantial amounts of personal data, granting property rights in personal information would allow individuals to bargain their right to their personal data against preferential services or other benefits. Laissez-faire and its outcomes represents in this approach individual freedom, because freedom is defined as free choices of an individual conceived as a trader [Radin (1987)].

Businesses collect and use personal data usually for free, gaining the full benefit without suffering financial losses or bearing the societal costs. Deighton compares the privacy invasion for the information economy to the pollution: social costs are borne not by those who benefit from the pollution, but by the rest of society [Deighton (2004)]. A property rights model would force businesses to internal-

ize certain costs of the widespread collection of personal data, which are now borne by the individuals [Samuelson (2000)], resulting in markets working more fairly and reasonably [Laudon (1996), Samuelson (2000)]. Such a privacy regime is deemed to be successful if it guarantees individuals the right to trade their personal information for perceived benefits and places the lowest transaction barriers in the way of consensual trades. To overcome the problem that separate negotiations would involve substantial transaction costs for the individuals, some commentators have suggested “informed intermediaries” to serve as intermediaries on behalf of individuals to represent their interests and negotiate with data-buyers [Cranor, (1999)].

The legislative policy that has been adopted in the USA reflects and favours market based solutions to personal information protection over strict regulatory approaches that characterize the European privacy policy. While there are fairly regulations that deal with the use of personal information by governmental agencies, comprehensive legislative solutions are eschewed with respect to the private sector. Legal protection of data privacy takes the form of ad hoc, narrowly circumscribed, sector-specific legislation, combined with recourse to litigation based on the tort of invasion of privacy and/or breach of trade practices legislation [Bygrave (2004)]. US has adopted a so-called sectoral approach to informational privacy [Reidenberg (2000), Cohen 2000]. Arising from particular economic considerations, this approach aims at keeping the data collection environment as business-friendly as possible, limiting the imposition of privacy regulation to specific situations so as to minimize the potential financial and administrative burden of businesses. The privacy as property model arises out of a historical American distrust of governmental solutions and preferences for market-based behavioural incentives. The American constitutional law also influences this approach. A first constitutional aspect concerns the Constitution’s placement of limitation generally on the government alone rather on private organisations. A second pertinent characteristic of constitutional rights in the US is that, even when applicable, these interests generally do not require the State to take positive action [Schwartz and Reidenberg (1996)]. Because of the ideological and constitutional heritage of the so-called negative liberty, Americans conceive all inalienabilities, the privacy-inalienability included, as paternalistic limitations of freedom. In the USA most discourse on privacy rights tends to focus only on the benefits these rights have for an individual qua individual [Bygrave (2004)], who accordingly should have the right to trade and sign away her privacy.

Moreover, creating stronger property rights is often thought to be a plausible way of securing interests in the era and the area of cyberspace. This is especially the case in legal orders (as for example the US), where a legal right of individuals to control uses or disclosures of personal data does not exist or, if provided, is very

restricted. Some authors propose a market/property rights model as an opportunity for setting default rules that would enhance consumers' control over their personal information [Litman (2000)]. By internalizing these costs firms could make better investment decisions about what data to gather [Samuelson (2000)] or they would be more inclined to collect and compile personal data than they currently do [Laudon (1996)]. Laudon argues that individual tradable ownership of personal data could enhance their protection because people would experience more control over the fate of their own information: "if personal data markets were allowed to function more effectively... the information flow would be more institutionalized and there would be less privacy invasion"[Laudon (1996)]. Legal scholars, however have been suspicious of the consequences of this approach for privacy rights, arguing that according individuals a property right would increase the value of information and thus the incentive for businesses to obtain (by whatever means) this information [Litman (2000)]. The resulting level of privacy would be seriously affected as "recognizing property rights in personally-identified data risks enabling more, not less, trade and producing less, not more, privacy" Cohen (2000)].

When assessing the propertization approach, we should also take into consideration some characteristics, which are inherent in the property rights systems, and mainly this of free alienability and further transferability. Free alienability thus prohibits an individual from limiting another party in the use or transfer of data. Once alienated, the consumer would have no more claim to it than a piece of sold real property [Hetcher (2001)] and cannot restrict those property interests that she signs away, which means that the buyer can freely transfer it to third parties. Propertizing personal information as a way of achieving information privacy goals might be proved not only an unnecessary but also a dangerous way. Alienability of personal information results in a situation that the individual, having sold her data, will be usually firstly precluded from engaging in further transactions to disclose or to sell the same personal data and secondly foreclosed from any control of these data in the hands of transferee or in the hands of third persons to whom the data might have been further transferred. Litman characterizes the assumption that ownership of data would enable individuals to restrain the downstream/further use of these data by negotiating conditions before disclosing them as a "fairy-tale picture" [Litman (2000)].

Freedom of contract, (in)equality and (mis)information

The basic assumption of property approach is that the individuals are able to exercise their (free) will with respect to their rights through the conclusion of contractual arrangements. This approach relies further on individual autonomy assuming that parties enter into contracts voluntarily, guaranteeing them a considerable degree of freedom to enter into contractual obligations. When applied to personal information, this argument implies that individuals have the choice to preserve their privacy or to negotiate the content of agreements to best suit their needs, and to exploit their privacy rights by using the instrument of contract.

The privacy/property debate is also a debate about freedom of choice and its necessary preconditions [Cohen (2000)]. The parameters and the conditions of choices within these given parameters are of crucial importance for evaluating the property rights approach. Apart from the difficulties with ascribing a value to personal information, it is highly contestable, if individuals do have effective choices in bargaining their privacy. The norm of choice is or should be interpreted in a way so as to include make a choice among a number of alternatives [Cohen (2000)]. Online transactions are increasingly governed by (standardized) contracts between providers and users, and less by a priori (default) entitlement structures. Individuals are free to accept or reject the terms and conditions offered but it is the vendor who specifies what terms to offer in the first place. In the practice of the online world, businesses and other users of personal data apply “take it or leave it” terms under the threat of exclusion or denial of access to digital services and information. Average consumers cannot bargain efficiently with their personal information because few, if any, have any idea of its actual or potential worth for the vending company or the subsequent acquirors, who –depending on the use - may value the data more highly.

The principle of freedom of contract applying to both contractors allows principally the stronger party to determine the contractual relationship in a way that is most unfavourable to the weaker party. In such a scenario the balance of powers between the contracting parties is clearly affected. Behavioural economics scholars have demonstrated that individuals general inertia toward default terms, specified by the vendor, is a strong and pervasive limitation on free choice [Korobkin (1998)]. Individuals might transfer ownership of their data exactly because they lack sufficient bargaining power.

The power inequality is strictly related to the information asymmetry, which characterizes the online personal information trading. Individuals are likely to know little about the conditions and the circumstances under which their personal data are collected, sold or transferred. The online industry has incentives to provide suboptimal information to guide individual decision-making about the

processing of personal data [Schwartz (2000)]. Websites think that it is in their interest to simulate privacy respect than to provide it. Posted terms and policies are cryptic or in small print no one can read. The analysis of communicative strategies in privacy policies has revealed that they contain vague statements. The opacity and vagueness contained in “terms and conditions” as well as “privacy policies” preclude people from understanding them in their entirety or may even deter them from reading these documents altogether [Pollach (2005)].

People are mostly not aware or informed of the technological context of data use [Schwartz (2000)]. Individuals face enormous difficulties in assessing the benefits and costs of an information trade. Usually they are not in a position to understand and assess the worth and uses of their data, especially in the case of secondary, tertiary or x-ry uses, which is also due to the fact that the categories of recipients are routinely specified in the most general terms [Cohen (2000)]. Consumers suffer from “privacy myopia”: they value their data at its marginal value in terms of lost privacy whereas a merchant will value it at or near its average value as a part of a profile [Froomkin (2000)]. Without information about the identity of users/acquirors and the nature of possible, future uses, individuals cannot determine if and what data to provide or withhold. The widespread ignorance regarding the terms that regulate the use of personal information prevents individuals from negotiating for their privacy interests.

Persistent (bargaining) inequality is reinforced through the lack of accurate information about choices and their consequences [Cohen (2000)]. The systemic disadvantage of individuals as far as it concerns their negotiating powers and abilities and the asymmetry of information underscores concerns about the vulnerability of the free to self-alienating individual. The knowledge gap undermines the freedom of contract as such. However, in most cases individuals thus appear to “gladly” agree to certain uses of their personal data. As they believe or feel that they must have access to services or benefits they submit easily the information required. The so-called “just click submit” phenomenon is “caused by the simple concepts of (1) must, (2) rush, (3) trust [Ciocchetti, (2008)]. Especially in cyberspace, consent to data processing is likely to turn into an empty, ritual process.

Consent and control

The contractual-property approach relies on interactions between individuals and data processors to generate and maintain appropriate norms for information privacy. One of the virtues of this approach is that it can accommodate the significance of consent as a factor of determining appropriate use [Samuelson (2000)]. Consent is considered to be a substantial, or even indispensable instrument because it safeguards the participation of the individual in the decision to the uses of her

data [Simitis (1999)]. In other words, under the condition that the individual is informed and it may express explicitly its free will and choice, consent guarantees primarily the right to express choices and/or exercise the right to informational self-determination. However, due to the weaknesses in the nature of agreement to data trade consent seems to result in a “fallacy” [Schwartz (2000)] for the individual. Some scholars argue that consent to data trade should imply not only the initial opportunity to accept or refuse trade, but also a later chance to exit from an agreement to trade [Schwartz (2005)], an argument that - at least in the pure property approach - seems to be a contradiction in terms.

Actually, consent was from the beginning a central element of the debate about privacy and personal data protection. The notion that informed consent alone is sufficient to protect individual interests in the uses of personal data is a peculiarly American one. American commentators in general seem to prefer consent in combination to market-based solutions to personal information protection over strict regulatory approaches [Samuelson (2000), Litman (2000) Prins (2006)]. In the context of European data protection legislation, consent of the individual serves also as a ground that legitimises by itself the processing of personal data. In this perspective the European data protection regime allows, in the final analysis, the individual to exploit her data [Prins (2006)]. However, in this context there is a crucial regulatory element, which distinguishes the European notion of consent as a legal basis for the legitimization of processing: consent is per definition informed and it has been surrounded by a vast array of mandatory requirements intended to secure the individual the chance to assess the consequences of processing and limit the stronger party's freedom of contract. The mandatory legal frame, consisting from the Council of Europe's Convention, the Data Protection Directive (D 95/46/EC) and the e-Privacy Directive (Directive 2002/58/EC), determines the range of processing and guarantees the respect of the individual's rights.

However, the new importance attached to an apparently well-established principle is all but fortuitous. The revitalisation of the discussion, both in the Anglo-Saxon and the Continental legal area, has in fact a clearly instrumental background. The upsurge in recent interest in a property right approach is fuelled by anti-regulatory (or in best case self-regulatory) culture approach. Stressing the attention to the uniqueness or primacy of the (necessity of) consent simulate an influence of the individuals that, if any, in reality depends solely on the specific relationship between the individual and the person or organisation, which will potentially use the information pertaining to the consenting individual. In this context, the consent is also seen as the probably best means to review legislative choices and reduce especially the powers of data protection supervisory authorities [Simitis (1999)]. “Self-empowerment” as substitute of binding rules

and institutional control reflects an approach of informational privacy as “private business” and serves the de-construction of privacy protective institutional instruments.

Fundamental rights for sale? Equality, Dignity and Democracy

Individual rights are conceived as closely linked to constituting and maintaining an individual's personal integrity. Can individuals waive the protection of their fundamental rights by means of a contract and “sell them to the highest bidder”[Edwards (2003)]? Should we consider privacy as an inalienable human right or as an item of property? Is it and, if yes, to what extent permissible to transfer the rights over our unique information to others or would this violate human dignity?

Dignity is not only an essential component of the human being but, moreover, also a condition for freedom and equality. In the market approach, tradability and alienability of privacy right is fully consistent with commitments to dignity and equality, because it considers each person as “an autonomous, rational actor” and supposes, “all individuals are equally capable of ascertaining and pursuing the goals that will maximize their own happiness”[Cohen (2000)].

As Laudon emphasizes, the property regime is inequitable. Individuals, irrespective of their economic situation, “possess” their identity, their own information and data image. However, poor people will sell their privacy more than the rich. Rich people, exactly because they are rich, most will not care to sell their privacy and simply withdraw from the market to preserve their privacy. Individuals from middle-income groups will probably seek the best prices, as their data will be presumably in high demand [Laudon (1996)]. In this context personal data is the “wedge that enables “scientific,” market-driven, and increasingly precise separation of “haves” from “have-nots”[Cohen, 2000].

The market approach, focusing on a tradable alienable privacy right, raises significant concerns in relation to equality. Reliance on data alienability places people who could not purchase “sufficient” privacy at a disadvantage when they trade for it [Schwartz (2004)]. If data privacy costs money or if giving up privacy saves money, than the preservation of privacy will become more unequal. In this case privacy would become a luxury right, its level depending mainly on socio-economic status. In the European approach the safeguards relating to the use of personal information aim at preventing discrimination against persons because of their opinions, beliefs, health or social status. Studies conveyed how profiling and the widespread collection and aggregation of personal information increase social injustice and generate even further discrimination against traditionally disadvantaged groups [Gandy (2000)]. Informational privacy is therefore regarded

as a condition for social inclusion and as a key component of an equality society [Rodota (2004), European Data Protection Supervisor (2005)].

The perspective that informational privacy will become simply “a status that can be chosen (and paid for) the way one might choose a neighborhood, a health club, or a brand of automobile” [Cohen (2000)] affects not only distributive justice principles and claims but also the dignity of the individual. Persons possess by virtue of their humanity an intrinsic worth that should be recognized and respected [Neuman (2000)]. Human rights are rooted in a non-commodified understanding of personhood and the attributes and context necessary to constitute and maintain personal integrity. In 1867, the Paris Appeal Court ruled in the case *Alexandre Dumas père* that one’s privacy, like other aspects of one’s honor, was not a market commodity that could be simply and definitely sold and stated that the “very publication of (such) photos” could put such a person on notice “that he had forgotten to take care for his dignity and remind him that the private life must be walled off in the interest of individuals, and often in the interest of good morals as well” [Whitman (2004)]. Nowadays privacy constitutes a response to the requirement for control of personal information so as to be free and able to decide how to behave in the frequently strained relations between individuals and society [Rasi (2002)].

Autonomy (self-determination), identity and privacy are seen as major, interrelated aspects of dignity, which - from Kant to Rawls - has been a central strand of Western philosophical tradition, embedded in many constitutional texts. The individual’s autonomous status concerning the elements of her identity, her informational self-determination and, do not allow principally appropriation of her “persona” by others, “because the decision to be associated with a certain commercial product is not entirely a commercial issue, but part of the inner core of a person’s personality” [Prins (2006)].

If an individual can be “objectified” and judged out of context, diminished is not only freedom but also the dignity of a person [Rodota (2004)]. Radin argues that the commodification approach undermines personal identity by conceiving personal attributes, relationships, and philosophical and moral commitments as monetizable and alienable from the self. Consistent with this argumentation she reveals the inherent contradiction of the property approach, stating “the person cannot be an entity exercising free will if it is a manipulable object of monetizable value” [Radin (1987)]. The “sanctity of personality” [Kang and Buchner (2004)] is inconsistent with trading and selling privacy in the information marketplace.

Privacy has ceased to be a right simply aimed at protecting the rights of the happy few bourgeois of the 18th and 19th century, who seek to foster their “private

sphere" [Simitis (1987)]. Privacy has value beyond its usefulness in helping the individual maintain her dignity or develop private relationships. It stands in a dialectical relation with dignity and personality: privacy is a right grounded on dignity and personality but, at the same time, it has become a pre-requisite for the development of everyone's personality. Personality, like liberty as well, does invoke a kind of freedom, the purpose of which is to allow each individual to realize her potential as an individual, to give expression to her peculiar capacities and to participate into social life [Rodota (2004), Whitman (2004)]. Privacy protects individuals against practices that erode individual freedom, their capacity for self-determination, and their autonomy to engage in relationships and foster social appearance and behaviour. Prevailing property-based approaches treat privacy as a matter of individual taste and ignore that informational privacy serves vital individual and collective purposes.

Development of the capacity for autonomous choices and decisions is an indispensable condition for free action in society and in the political arena [Rodota (2004)]. Privacy is necessary for both "individual self-determination" and "democratic deliberation" [Schwartz (2004)]. The German Federal Constitutional Court in the National Census Case [Bundesverfassungsgericht (1983)] stressed the so-called democracy-functional aspect of the individual's right to an "informational self-determination": unrestricted access to personal data imperils virtually every constitutionally guaranteed right. "Neither freedom of speech nor freedom of association nor freedom of assembly can be fully exercised as long as it remains uncertain whether, under what circumstances, and for what purposes, personal information is collected and processed"[Bundesverfassungsgericht (1983)]. If individuals fear that information pertaining to them might lead to false incrimination, reprisals or manipulation of their data, they would probably hesitate to engage in democratic participatory activities, a position which might shape the public's willingness to voice critical and constructive idea. The vitality of democracy depends exactly on an autonomous citizenry [Schwartz (2004)]. The autonomy fostered by informational privacy generates collective benefits because it promotes "reasoned participation in the governance of the community" [Cohen (2000)].

Democratic participation presupposes constant interaction between public and private life [Simitis (1987)], which requires an information allocation policy including the guaranteed access to the information necessary to follow and evaluate social processes. Granting individuals property rights would lead to the "division" and "monopolization" of personal information. Simitis advocated for a privacy approach, which does not eliminate access to information, arguing, "open or hidden 'sanctifications' of property sacrifice the citizen and reduce the *constitutio libertatis* to a mere guarantee of the bourgeois' refuge" [Simitis (1987)]. In this approach informational privacy promotes the development of both indi-

viduals and society and its value lies exactly in its systemic effects on power and powerlessness in society.

Conclusion

Without a coherent conception about the nature of a person's interest in information pertaining to her, it is difficult to design a legal regime to protect this interest appropriately. Privacy has instrumental value in terms of psychic well-being [Kang and Buchner (2004)] and the construction and deepening of social relationships. Privacy has value beyond its usefulness in helping the individual maintain his or her dignity or develop personal relationships. To the extent that privacy is related to the concepts of dignity and autonomy, it refers not only to principles such as the protection of an individual's personality or non-interference with other's life choices, but also to the non-commodification of the individual as well as to the secured possibility to act autonomously and freely [Rodota 2004]. One can freely dispose of one's liberty but one can never be permitted to freely dispose of one's dignity. Hence, the property approach cannot provide a strong and privileged ground for informational privacy protection. Relying on private property rights to serve as a basis for privacy hardly gives this right the privileged standing that individuals claim for it [Prins (2006)] and undermines collective ends and interests.

Information privacy, in its both nature as individual good and condition for a democratic participatory society, requires "some degree of social and legal control to construct and then maintain" [Schwartz (2004)]. One of the strengths of the European approach is that the regulatory regime it embodies is consistent with its underlying conception of information privacy as a fundamental human right grounded in dignity and as a condition for active participation to public life. In academic discourse on both sides of the Atlantic it is increasingly recognised that data privacy rules are valuable not simply for individual persons but for the maintenance of pluralism and a flourishing democracy.

REFERENCES

- Bundesverfassungsgericht (1983). Volkszählungsurteil, 65, 1.
- Bygrave, L. A., (2004). Privacy Protection in a Global Context – A Comparative Overview. *Scandinavian Studies in Law*. Vol. 47, 319-348
- Ciocchetti, C., (2008). Just Click Submit: The Collection, Dissemination and Tagging of Personally Identifying Information. *Vanderbilt Journal of Entertainment and Technology Law*. Vol. 10 (Spring 2008), 553-642

Cohen, J.E., (2000). Examined Lives: Informational Privacy and the Subject as Object. *Stanford Law Review*. Vol. 52, 1373-1438.

Cranor, L., (1999). Internet privacy, *Communication of the ACM*. Vol., 42 (2), 28-38

Deighton, J., (2004). The Presentation of Self in the Information Age. *Harvard Business School Marketing – Research Papers No 04-02*.

Edwards, L., (2003). Consumer Privacy, On-Line Business And The Internet: Looking For Privacy In All The Wrong Places. *International Journal of Law and IT*. IJL&IT 2003.11(226)

Epstein, R. A., (1978). Privacy, property rights and misrepresentations, *Georgia Law Review*. Vol. 12, 455-480

European Data Protection Supervisor, Public access to documents and data protection. Background Papers Series No 1, July 2005, Brussels

Froomkin, M.A, (2000). The Death of Privacy? *Stanford Law Review*. Vol.52, 1461-1543.

Hetcher, S., (2001). Changing the Social Meaning of Privacy in Cyberspace. *Harvard Journal of Law and Technology*. Vol. 15, 149-209.

Hotaling, A., (2007-2008). Protecting Personally Identifiable Information on the Internet: Notice and Consent in the Age of Behavioral Targeting. *CommLaw Conspectus*. Vol. 16. 529-565

Gandy, O.H. Jr., (2000). Exploring Identity and Identification. *Notre Dame Journal on Law, Ethics and Public Policy*. Vol. 14 (2), 1085-1111.

Gauthronet, S. (2001). The future of personal data in the framework of company reorganizations. *Proceedings of 23rd International Conference of Data Protection Commissioners*. Paris.

Kang J. & Buchner B., (2004). Privacy in Atlantis. *Harvard Journal of Law and Technology*. Vol. 18, 229-266.

Kohler, T.C., (2000). Autonomy and Personhood: The implications for Labor and Employment law, in D. Simon/P. Weiss, *Zur Autonomie des Individuums – Liber Amicorum Spiros Simitis*. Nomos Verlag. Baden-Baden. 164-170

Korobkin, R., (1998). Inertia and Preference in Contract Negotiation: The Psychological Power of Default Rules and Form Terms. *Vanderbilt Law Review*. Vol. 51, 1583 -

Laudon, K.C., (1996). Markets and Privacy. *Communications of the ACM*. Vol. 39(9), 92-104.

Litman, J., (2000). Information Privacy/Information Property. *Stanford Law Review*. Vol. 52 (5), 1283-1313.

Mell, P. (1996). Seeking in a Land of Perpetual Sunlight: Privacy as Property in the Electronic Wilderness. *Berkeley Technology Law Journal*. Vol. 11, 21-82.

Newman G.L., (2000). Human Dignity in United States Constitutional Law. D. Simon/P Weiss (Hvsg). *Zur Autonomie des Individuums - Liber Amicorum Spiros Simitis*. Nomos Verlag, Baden. Baden, 249-271

Pollach, I., (2005). A Typology of Communicative Strategies in Online Privacy Policies: Ethics, Power and Informed Consent. *Journal of Business Ethics*. Vol. 62, 221-235

Posner, R.A., (1978).The Right of Privacy. *Georgia Law Review*. Vol. 12 (Spring) 1978

Prins, C., 2006. When personal data, behavior and virtual identities become a commodity: Would a property rights approach matter? *SCRIPT-ed*. Vol. 3 (4), 270-303

Radin, M. J., (1987). Market Inalienability. *Harvard Law Review*. Vol. 100, 1849 –1937.

Rasi, G. (2002). What is changing in production activities. *Garante per la Protezione dei Dati Personali, Da costo a risorsa – La tutela dei dati personali nelle attività produttive*. Roma, 17- 28.

Reidenberg, J.R., (2000). Resolving Conflicting International Data Privacy Rules in Cyberspace. *Stanford Law Review*. Vol. 52, 1315-1375

Rodota, S. 2004. Privacy, Freedom and Dignity – Closing remarks. *Proceedings of 26th International Conference on Privacy and Personal Data Protection*. Wrocław.

Samuelson, P., (2000). Privacy as Intellectual Property? *Stanford Law Review*. Vol. 52, 1125-1173.

Schwartz P.M. & Reidenberg, J.R., (1996). Data privacy law – A study of United States data protection. *Michie Law Publishers*, Charlottesville.

Schwartz, P.M., (2000). Privacy, Participation and Cyberspace – An American Perspective. D. Simon/P. Weiss (Hrsg.). *Zur Autonomie des Individuums – Liber Amicorum Spiros Simitis*. Nomos Verlag, Baden-Baden, 337-352

Schwartz P.M., (2004). Property, Privacy and Personal Data. *Harvard Law Review*. Vol. 117, 2055-2128

Schwartz P.M., (2005). Privacy Inalienability and the Regulation of Spyware. *Berkeley Technology Law Journal*. Vol. 20, 1269-1282

Simitis, S., (1984). Die informationelle Selbstbestimmung Grundbedingung einer verfassungskonformen Informationsordnung. *Neue Juristische Wochenschrift* 1984 (8), 389-405

Simitis, S., (1987). Reviewing privacy in an Information Society, *University of Pennsylvania Law Review*. Vol. 135, 707-732

Simitis, S., (1999). Die Erosion des Datenschutzes. Von der Abstumpfung der alten Regelungen und den Schwierigkeiten, neue Instrumente zu entwickeln. In Hrsg. Bettina Sokol: *Neue Instrumente im Datenschutz*. Wuppertal, 5-40

Solove, D. (2001). Privacy and Power: Computer Databases and Metaphors for Information Privacy. *Stanford Law Review*. Vol. 53: 1393-1462

Weichert, T., (2004). Wem gehören die privaten Daten?. In Taeger, J and Wiebe A., *Informatik – Wirtschaft – Recht . Regulierung in der Wissengesellschaft –Festschrift für Wolfgang Kilian*, Nomos Verlagsgesellschaft, Baden-Baden, 281-298

Westin, A. (1967). *Privacy and Freedom*. Athenaeum, New York.

Whitman, J. Q. (2004). The Two Western Cultures of Privacy: Dignity versus Liberty. *The Yale Law Journal*. Vol. 113, 1153-1221

The Joy of Excellence: Intellectual Property and Hackers' Virtue

Scott J. Molony*

Boston College, Lawrence House
Presidential Scholars Program

Abstract

The legitimacy of intellectual property (IP) rights is an ongoing controversy in Computer Ethics. In this paper, I argue that the debate over the legitimacy of IP rights has been incorrectly focused. I argue that it has been both theoretically demonstrated and practically established that retaining IP rights is ethically permissible. However, I argue for a virtue ethics where renouncing IP rights is an act of supererogation rather than obligation. Particularly as concerns software, I argue that this cultivates a new, MacIntyrean virtue, which I refer to as, "hackishness."

Spinello outlines three primary arguments for IP rights. These derive from:

1. Locke
2. Hegel
3. Bentham and Mill

The Lockean argument has received a great deal of attention already (Tavani 2005), but essentially argues that one is entitled to derive benefits from the fruits of one's labors. Insofar as one is an author or programmer, IP rights are the only way to secure those profits, and are therefore legitimate. However, Locke ties these profits to one's livelihood – they do not extend beyond the level needed for sustenance (which, broadly interpreted, may mean a modest living).

Hegel's argument, as outlined by Spinello, is considerably different from Locke's. Hegel contends that an artisan invests some portion of his or her personality in the work that the artisan creates. Because the artisan has a right to autonomy and

* *Scott J. Molony* is an undergraduate at Boston College in the Presidential Scholars Program, pursuing a degree in Philosophy and Theology. His research interests are in all areas of computer ethics, particularly the intersection spirituality and professional ethics. He is the former winner of the Siemens Competition in Math, Science and Technology for a project in bioinformatics.

self-determination, and the objects of one's craft are a part of one's personality, one must retain a right of control over a work's use, or at least a right to maintain the integrity of the work. Insofar as IP rights are the only way to guarantee this control, they are legitimate.

The Utilitarian (or pragmatist) argument is the view articulated in the US Constitution. "The Congress shall have power ... to promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries" (Art. 1 §8) It argues that the limited monopoly given to creators of works promotes innovation ("the Progress of Science and useful Arts"), which is in the best interest of all concerned. Insofar as IP rights spur innovation and discovery, then, they are legitimate.

The novel virtue, "hackishness", I am arguing for is first articulated by Stephen Levy (1984) and continued by writers like Eric Raymond (2008) and Cory Doctorow (2007). Levy gives several examples of what he terms the "Hacker Ethic" and I later use as the basis for the "hackishness" virtue. Eric Raymond presents a similar description, but while Levy writes from a perspective outside the Hacker culture, Raymond writes as an elder statesman from within the culture. Doctorow's work is intended for young adults, and, in addition to its concerns about Privacy, his novel *Little Brother* is a hortatory work, encouraging young adults to be involved with programming by giving a description of its joys.

I present a brief overview of the phenomena associated with the exercise of the virtue, and I give a detailed description of the novel combination of attitudes this virtue entails, particularly as regards Computer Programming and Software Creation.

I then argue that programming fulfills the criteria set forth by Alasdair MacIntyre (1984) for a "practice."

1. The activity is coherent and complex.
2. The activity is socially-established and cooperative.
3. There are goods internal to the activity.
4. There are standards of excellence internal to the activity.
5. These goods are realized in performing the activity excellently.
6. Human powers and goods are extended by performing the activity excellently.

Complexity and coherency are trivial to establish. I refer to modern code-sharing websites like SourceForge as well as historical communities like the Homebrew Computer Club in Levy's work to establish the second criteria. The existence of

goods internal to the activity is typified in Levy's assertion that the sole driving motivation of the hacker is to "create awesome software." (1984; 381) The standards of excellence in hacking are extensively discussed by Raymond (2003), who gives an exceedingly thorough treatment of what constitutes good UNIX software. All three authors give an excellent treatment of the changes which hacking instills in a person. Far from merely instilling a technical skill set, the truly hackish person derives intense joy from solving interesting problems, and takes joy in enabling other people to solve interesting problems. Thus, the hackish person extends human powers and the human good by conquering difficult computer problems and providing the means for others to do so as well.

Thus examined, I conclude that computer programming is a practice, and its most excellent exercise, hackishness, is a virtue. I then examine three major contemporary IP frameworks to determine their consonance with the virtue of hackishness:

1. Full United States Copyright Protection
2. GNU GPL
3. Creative Commons

Next, I examine how a person demonstrating this virtue might act within each of these three frameworks. I conclude that Full Copyright Protection, as established earlier, is legitimate to protect economic well-being of a person (programmers must eat, after all), but traditional proprietary programming becomes problematic because it doesn't encourage the social component of code-sharing identified above. The GPL suffers from precisely the opposite dilemma – although it encourages the virtue of hackishness exceedingly well, it becomes problematic in that, with a few exceptions (e.g., those fortunate enough to work for companies like Red Hat), it does not answer Bill Gates' objection: "Who can afford to do professional work for nothing?" (Levy 1984; 226) It also falls prey to the Hegelian objection that a programmer has no control over the portion of his or her personality invested in the code. I argue that Lessig's Creative Commons framework appears to take the best approach – by permitting, for example, free use of software for noncommercial activity, but requiring corporations to deal individually with the programmer, the programmer can share his code with the world at large, and still derive a basic living. The attribution and (if necessary) no derivative works clauses allow the Hegelian to maintain control over works, if desired.

Finally, I demonstrate how the hackish virtue leads to acts of supererogation for each of the three justifications for IP rights. Against the Lockean, the hackish person replies that, a modest living having been secured (through either full IP protection or Creative Commons frameworks), the joy derived from programming is the only payment required – the proviso has been fulfilled. Against the

Hegelian, the hackish person replies that fulfillment can also come from the creative evolution of works, something that is impossible if rigid controls on use are maintained. Finally, the pragmatist's argument has already been answered, by the description of what it means to be hackish – those who cultivate this virtue are driven by the desire to innovate. It doesn't need to be induced – merely supported. Insofar as there is adequate support available, hackish people will innovate, regardless. Lastly, I speculate on future work, including applying this virtue to other forms of IP.

Keywords: Intellectual Property, IP, Hackers, Hackishness, Virtue Ethics, Supererogation

Background: Justification for Intellectual Property

How can one own an idea? Such a question has been raised in society since the time of Gutenberg – and, until recently, had been satisfactorily answered. The advent of the Internet, however, has changed the entire frame of the question. When books, inventions, and other creative works heretofore were physical objects – things that could be touched, made, and controlled by scarcity – society developed copyrights and patents, which worked admirably. However, the digital revolution of the Internet changed the rules of the game when two major advances changed the nature of creative works:

1. Creative works are digital – composed on computers. This means that a work has gone from being markings on a piece of paper to being bits in a data stream. It is therefore now possible to reproduce flawless copies of any work composed on the computer. While this capability existed in the hands of book printers up until now, the widespread adoption of computers and desktop publishing has enabled even average citizens to produce work on the level of professional typesetters. Software, in particular, is a medium which is exclusively digital – code outside of a computer environment is useless.

2. The Internet connects personal computers together in a global network. While the capability to move files from one computer to another was one of the first major advances of the computer industry, Geography had heretofore posed a definite limitation to any attempt to share files; Users could put a file on portable storage media like a floppy disk, say, and share it with neighbors, but only realistically able to share information between people who are geographically close. Companies could control the distribution of software by being the only people who could ship uses copies of software. With the advent of the Internet, however, works can now be placed online for anyone to access and copy.

This has thrown back open the question of intellectual property (IP) rights: are the rights given by traditional intellectual property regimes morally legitimate?

Some, like noted activist and president of the Electronic Frontier Foundation (EFF) Richard M. Stallman argue that intellectual property rights, at least applied to software, constitute a “betrayal” (Babe 2004) of other users; an idea which he calls “morally sickening” and argue that “It is not ethical to use non-free software.” Other authors, however, like Richard Spinello (2006) argue that at least the idea of some sort of copyright is completely legitimate, if not the current US copyright regime.

The debate appears to be somewhat intractable. However, in some ways, the debate is also moot. It is highly unlikely that the logical conclusion of abolitionists like Stallman’s reasoning will come to pass, and copyright will (even if merely for software) be abolished. The debate, then, needs to be refocused – given that copyright exists, what is acceptable conduct within this framework? What is the most morally superior way of conducting oneself under the scheme? It is these questions which have the most direct bearing on moral lives, and deserve the most scrutiny.

I will argue that a virtue-based ethics of supererogation is the most appropriate way to deal with these questions. In order to deal with these questions, however, it is necessary to first examine what justifications are given for IP rights in the first place in order to be able to adequately answer the question. There are three sources for justifications of IP rights. Briefly, they are:

1. A Lockean theory, commonly referred to as the Labor-Desert theory
2. A Hegelian theory, referred to as the Personality theory
3. A Utilitarian theory

A brief discussion of the three follows.

Locke: Labor-Desert Theory

The Lockean argument is perhaps the one most documented. Briefly, the argument emphasizes that Labor makes Right – an agent has a legitimate claim to only that which the agent works for.

Locke begins with a rights-based approach; declaring, as noted above, that everyone has rights to life, liberty, and property. His main moral thrust comes in his assertion that people are entitled to as much property as may be necessary to sustain themselves’ and their dependents’ right to life. (The text lends itself to a broader reading of “sustenance” to include a modest, though not extravagant, living.) Locke also asserts that the earth was, originally, held by people in common; property rights derive from people claiming certain small portions of land and other resources to exclude other’s access to, thus increasing the number of people

who can subsist on the remaining common land. (Ironically, he essentially justifies private property on altruistic grounds.)

Locke derives his theory entirely from a supposition that all tangible property is held in common; the Lockean argument for IP rights extends this to creating a parallel immaterial commons (Spinelis & Bottis 2009, Kimppa 2003; 64). One key feature of the immaterial commons, however, is that every object in it is both unique and nondiminishing. Hamlet is only one particular combination of words in the entire English language, and so can be said to be unique, but my reading Hamlet does not permit my next-door neighbor from doing so as well.

Hughes identifies two alternate readings of Locke's theory:

1. Labor-Avoidance: This reading holds that labor is, by definition, an implicitly unpleasant activity that people pursue only reluctantly. It is just, therefore that those who labor be rewarded with the fruits that are thence derived. Applying it to Hamlet Shakespeare deserves the IP right to Hamlet because writing plays is painful (or, at least, less pleasant than alternate options).

2. Value-Added: Those whose labor creates value should be rewarded with its profits. Rather than stressing the unpleasantness of the work, this theory stresses the value that is created. Applying this to the Hamlet example, Shakespeare is entitled to the IP rights to Hamlet not because it is painful to write plays, but because Hamlet is a great play.

Hegel: Personality Theory

The Hegelian (Hughes 1988; 330-365) position holds that an author invests some facet of the author's personality in all created works. The works created by a given author embody a part of the author's personality. Therefore, an author is entitled to some amount of control over the created work.

Note that the IP rights associated with this theory are separate and distinct from those associated with the Lockean argument. Rather than stressing compensation for pain or value, the Hegelian instead demands veto power over derivative works, and seeks to ensure that created works accurately reflect the author. Unlike either the Lockean or the Utilitarian, who argue that creation must be incentivized, the Hegelian does so almost from a defensive posture – to define the agent itself concretely in the world, rather than allow others to define the agent. This argument assumes that creation will happen, and even encourages it, but insists that the rationale behind creative works – a self-actualization solely by an agent – must remain only within the agent's sole control.

Mill/Bentham: Utilitarianism

The utilitarian's argument is perhaps the simplest of the three, and is most prominently articulated in the US Constitution, which reads, in part "The Congress shall have power ...To promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries." (Art. I, §8, Cl. 8)

The utilitarian's argument is merely that creating works takes time, and effort, and without sufficient protection, authors will have no incentive to create, and will, accordingly, not. Important to this idea is that creation: artistic, literary, scientific, or technical, must be incentivized. Agents within the system do not create IP works absent these incentives, because, economically, the "costs of expression" outweigh the "costs of distribution" – in effect, because it does not make economic sense to create IP works, agents will, accordingly, not create IP works. Because IP works are held to be a nonmoral good, the utilitarian argues that the only moral route for society is to give some sort of economic incentive for IP works to be created. Further, the utilitarian argues, the best method to incentivize the production of IP works is the limited monopoly granted by IP rights; ergo IP works are moral.

Hackishness: A New Virtue

Against this background of justifications for IP rights, I argue for a new virtue ethic of supererogation. The question I mean to answer is not, "Are IP rights legitimate?" but rather, "What is the most morally superior action when dealing with IP?" To that end, I point to a novel virtue, which I deem "hackishness." This virtue describes supererogatory conduct – that is, conduct that cannot be obligatory, but is nevertheless morally superior to alternate actions.

Origin of Articulation: Hacker culture, especially writings of E. S. Raymond

"Hacker" is an unfortunate term in modern culture. Thanks to sensationalist journalism, it has left its former meaning, and come to mean "computer criminal." This is unfortunate, because it is a radical departure from the former meaning – "enthusiast programmer." A "hacker" in the former (enthusiast) sense was a person who engaged deeply with a given system, in order to learn how it works, and control it in the optimal way – which often meant innovative departures from what was presumed to be standard operating procedure for the system.

This is most clearly seen in the writings of E. S. Raymond, a self-proclaimed "hacker" and major developer of Linux – something of an elder statesman in this subculture. "Hackers" have a number of characteristics, most clearly addressed in Raymond's manifesto, "How To Become A Hacker" (2001).

Hackishness: A Description

Hackishness derives, then, directly from hacker culture. Raymond's manifesto describes an ethic of constant innovation, with admonitions like "The world is full of fascinating problems waiting to be solved" and "boredom and drudgery are evil" (Raymond 2001), Hackish people, then, are those who are constantly internally driven to innovate, because they derive value from the act of innovating itself, not from that which results from it. Those who do attempt to use hackish techniques for personal gain – the cyber criminals who give the modern meaning of "hacker" its name – are disparagingly referred to by Raymond as "crackers" and dismissed.

The other great source for a qualitative description of hackishness comes from Steven Levy's work "Hackers: Heroes of the Computer Revolution", which has acquired the status of a major work of folklore in the subculture. While the book is inundated with descriptions of hackish people, a couple of key passages leap out:

"Art, science and play had merged into the magical activity of programming" (Levy 1984; 120).

The virtue of hackishness necessarily merges these three activities into one – the satisfaction that comes from the exercise of hackishness is equal parts aesthetic experience, novel discovery, and joy of creation.

"The people in Homebrew were a mélange of professionals too passionate to leave computing at their jobs, amateurs transfixed by the possibilities of technology, and techno-cultural guerrillas devoted to overthrowing an oppressive society in which government, business, and especially IBM had relegated computers to a despised Priesthood" (Levy 1984; 200).

Here we see how this is a true manifestation of the virtue of hackishness – the passion within these people melds their work and their play, into the same thing, and they stare, awed at the possibilities of what machines can do for them. Most telling, though is the description of an ideal company by Mark Duchaineau – a young, hackish programmer at Sierra On-Line who, faced with the increasing bureaucratization of his company, dreams of founding his own, utopian workplace:

"Duchaineau's company would be a hacker paradise, with programmers having every conceivable tool at their disposal to create awesome software. If a programmer felt the company needed a piece of equipment, say some supercalibrated oscilloscope, he would not have to get permission from unconnected management channels...he and his fellow hackers would have a large say in the process" (Levy 1984; 378).

It's a telling description of paradise – one where the people have all of the tools that they need to create awesome software. Duchaineau doesn't dream of "enough royalties to buy cherry-red Trans-Ams and Caribbean trysts with hot-blooded software groupies" (Levy 1984; 389). Rather, his ideal company merely gives him the ability to pursue the joy of programming unimpeded.

The virtue is not confined to these older works, however. Cory Doctorow, a noted blogger and author, also describes the virtue of hackishness in his young adult novel, "Little Brother."

"If you've never programmed a computer, you should. There's nothing like it in the whole world. When you program a computer, it does exactly what you tell it to do. It's like designing a machine – any machine, like a car, like a faucet, like a gas-hinge for a door – using math and instructions. It's awesome in the truest sense: it can fill you with awe" (Doctorow 2007; 49).

In this way, Doctorow's work has a horatory facet; the novel is aimed at young people, and here he describes the joy associated with the virtue of hackishness in order to entice them to practice it themselves.

Programming: A MacIntyrean "Practice"?

Alasdair MacIntyre, noted moral philosopher, proposed in his work "After Virtue" a new frame of morality. Fed up with emotivism and the relativism which pervaded modern moral theory and led to intractable problems of dueling axiomatic assertions, MacIntyre turns to an Aristotelian moral theory. He frames virtues in terms of "practices" – the excellent performance of a certain activity which tends to improve the agent's character as a result. MacIntyre specifies six criteria for a "practice," which I detail below, and then apply to Computer Programming.

1. The activity is coherent and complex.
2. The activity is socially-established and cooperative.
3. There are goods internal to the activity.
4. There are standards of excellence internal to the activity.
5. These goods are realized in performing the activity excellently.
6. Human powers and goods are extended by performing the activity excellently.

The activity is coherent and complex

This assertion is almost trivial to establish. Coherency is the defining feature of computer programming; the “object” of writing a program could be said to produce a string of alphanumeric characters which fit into a particular grammar of a programming language. Complexity is also trivial – computer programming contains an almost bewildering variety of separate languages, each of which have many, many commands associated with them. The pursuit of programming, even at the present state of the art, may well said to be endless, since it is doubtful a single individual could become proficient in every language and every environment.

The activity is socially-established and cooperative

This is less trivial to establish, but not much more so. Computer programming can be practiced as an entirely solitary pursuit, but most programmers operate within some sort of larger community. Different kinds of communities have existed throughout the different ages of programming; the earliest computer hackers, descendants of the MIT Model Railroad Club, all shared a single machine, and so formed a community surrounding a single terminal. (Levy 1984; 26) A later generation formed clubs, the most famous being the Homebrew Computer Club which flourished in San Francisco in the late 1970s. (Levy 1984; 195-218) Modern hackers use websites like sourceforge.net and the Mozilla foundation (Mozilla Foundation 2008) to collaborate on common coding projects from around the globe.

There are goods internal to the activity

Most of the previous section detailed the goods associated with the activity of programming – equal parts aesthetic experience, novel discovery, and joy of creation. These goods are potent, and, for hackish people, sufficient to justify their pursuit to the exclusion of others.

There are standards of excellence internal to the activity

There are many, many books written on what constitutes good computer programming; two major ones are E. S. Raymond's *The Art of Unix Programming* (2003), and Donald Knuth's *The Art of Computer Programming* (1997). Knuth's monograph is largely a technical work describing the fundamentals of algorithmic theory – what makes good code on a technical level. Raymond's work is much more philosophical – it deals with aspects of system design, and is a normative theory of how computer programs ought to interact with the user and each other.

These goods are realized in performing the activity excellently

There can be little doubt that those who take programming seriously are changed by it. As noted by Levy,

“People would work all day long, rush home, eat their dinner and come back,” MITS executive Eddie Currie later recalled. “You could go in there any hour of the day or night and there would be twenty or thirty people, a third to half the staff [excluding manufacturing]. And this went on seven days a week. People were really caught up in this because they were giving computers to people who were so appreciative, and who wanted them so badly. It was a grand and glorious crusade”(Levy 1984; 221).

The only reason people could be caught up in such a “grand and glorious crusade” is that they derive some deep, inner satisfaction from the work. One characteristic of hackishness is a passionate, single-minded pursuit of excellence, particularly in programming. By becoming immersed in programming, and thereby cultivating this virtue, hackers gain this good.

Human powers and goods are extended by performing the activity excellently

This is implicit in the entire premise of hacking. Raymond admonishes, “Creative brains are a valuable, limited resource. They shouldn’t be wasted on re-inventing the wheel when there are so many fascinating new problems waiting out there.” (Raymond 2001) Thus, the point of exercising this virtue is to extend human capabilities – to allow people to do new, innovative, wonderful things with their computers that they had heretofore not been able to do.

The Open-Source Movement: The Cultivation of “Hackishness”

There are three primary IP schemes for software (there are many licenses, on the order of a hundred, but the three most archetypical for software are the following):

1. Full United States Copyright Protection (“All Rights Reserved”)
2. GNU General Public License (GPL) (One must distribute source code with the license, and all derivative works must be licensed with the GPL)
3. Creative Commons series of licenses (Attribution mandatory, with optional restrictions on derivative works, commercial use, and relicensing under the same Creative Commons license)

Full US Copyright Protection

It is difficult to make full US Copyright protection consonant with the virtue of hackishness. The key obstacle is that it explicitly cuts code off from the social

aspect, identified above as being key to hackishness's expression as a practice. It is true that someone might still find a way to reconcile traditional, proprietary programming with a social community – a programmer might operate in a group in a company, for example. However the full expression of hackishness is nigh-impossible to do within this scheme.

GNU GPL

The GNU GPL suffers from the opposite problem. While it cultivates this virtue excellently by encouraging collaboration (the explicit point according to its author, Stallman (2008)), the GPL's proponents have difficulty answering a question posed by the planet's richest man: "Who can afford to do professional work for nothing?" (Gates 1976) Here, Gates has a point – who can work for nothing? Insofar as he is a programmer wanting to earn a livelihood to continue programming, his desire is praiseworthy, but insofar as he only desires money, and is doing this to make money, it is only permissible, not praiseworthy.

Creative Commons

This is the licensing scheme which best fosters the virtue of hackishness. By permitting free copying of software for personal use, it fulfills the hackish imperative that drives an author to get as many people to use it as possible. However, by reserving the right to deal with corporations, it also allows an author to make something of a living, and thus continue creating works.

"Hackish" virtue's interaction with IP

Having now identified the Creative Commons framework as the one which best fosters the virtue of hackishness, I now examine how this virtue of supererogation interacts with the philosophical justifications given for IP rights earlier.

Locke: Labor-Desert Theory

Recall that the Labor-Desert theory had two main expressions:

1. Labor-Avoidance: This reading holds that labor is, by definition, an implicitly unpleasant activity that people pursue only reluctantly.
2. Value-Added: Those whose labor creates value should be rewarded with its profits. Rather than stressing the unpleasantness of the work, this theory stresses the value that is created.

The virtue of hackishness short-circuits the former objection. To a truly hackish author, the act of creation is not something unpleasant; rather, it is a deep, fulfilling joy. "You couldn't stop me from writing unless you broke my fingers." says Doctorow (Molony 2008). To the second objection, the hackish person replies that, after satisfying the Lockean proviso of a modest living (mostly to facilitate further creation), no more reward need be given. Indeed, the hackish person is

perhaps the Lockean exemplar — a person driven to work, and to contribute rather than take away from the intellectual commons.

Hegel: Personality Theory

The Hegelian has a whole other set of concerns. Unlike the Lockean, who could probably adopt the GPL if sufficiently well off to afford to work for free, the Hegelian could not, because it provides no protection for the integrity of the work. However, one of the Creative Commons licences explicitly disallows unauthorized derivative works, which affords the Hegelian the best of both worlds — unrestricted distribution of verbatim copies of a particular work

Mill/Bentham: Utilitarianism

The utilitarian's arguments can only really be countered with an alternative business plan. One of the things that makes Doctorow notable is that, while being a successful author, he releases his books online under the Creative Commons framework. One book, the aforementioned *Little Brother*, is now about to go into its 8th printing. This would appear to be a contradiction, but in an article released in *Forbes* magazine (Doctorow 2006) he argues that releasing the book online actually improved sales; one might even argue that it therefore makes economic sense for author to release their work under the Creative Commons licenses.

In addition, the virtue of hackishness refutes some of the basic reasoning associated with the Utilitarian argument. Rather than only create because of incentives in some sort of dreadful economic plan, the hackish person creates for the joy and satisfaction inherent in creation, with the fruits at best a distant second consideration. The cultivation of this virtue can “promote the Progress of Science and useful Arts” far better than any incentive scheme.

REFERENCES

Anarcho Babe, (2004). Interview with Richard Stallman. Free Software Foundation. Available from <http://www.gnu.org/philosophy/rms-interview-edinburgh.html> [Accessed 4-15-09]

Doctorow C., (2006). Giving It Away. *Forbes*. Available from: http://www.forbes.com/2006/11/30/cory-doctorow-copyright-tech-media_cz_cd_books06_1201doctorow.html [Accessed 4-15-09]

Doctorow C., (2007). *Little Brother*. New York: Tom Doherty Associates. Available from http://craphound.com/littlebrother/Cory_Doctorow_-_Little_Brother.pdf [Accessed 4-15-09]

Gates, W.H., (1976). An Open Letter to Hobbyists. Available from: <http://www.blinkenlights.com/classiccmp/gateswhine.html> [Accessed 4-15-09]

Hughes, J., (1988). The Philosophy of Intellectual Property. The Georgetown Law Journal, 77 (2), 287–366.

Kimppa, K., (2003). Problems with the Justification of Intellectual Property Rights in Relation to Software and Other Digitally Distributable Media, (PhD). Turku Centre for Computer Science.

Knuth, D.E., (1997). The Art of Computer Programming, Volumes 1-3. New York: Addison-Wesley.

Levy, S., (1984). Hackers: Heroes of the Computer Revolution, 1st ed. New York: Doubleday.

MacIntyre, A., (1984). After Virtue, 2nd ed. Notre Dame: University of Notre Dame Press

Molony, S. (scott.molony.1@bc.edu), 16 Dec 2008. Re: Little Brother and IP rights. e-mail to Doctorow, C., (doctorow@craphound.com).

Mozilla Foundation. (2008). What is Mozilla? Available from: <http://en-us.www.mozilla.com/en-US/about/whatismozilla.html> [Accessed 4-15-09]

Raymond, E.S., (2003). The Art of Unix Programming. Available from <http://www.catb.org/~esr/writings/taoup/html/> [Accessed 4-15-09]

Raymond, E.S., (2008). How To Become A Hacker. Available from <http://www.catb.org/~esr/faqs/hacker-howto.html> [Accessed 4-15-09]

Spinello, R, (2006). Cyberethics: Morality and Law in Cyberspace. 3rd ed. Sudbury, MA: Jones and Bartlett.

Spinello, R, & Bottis M., (2009). A Defense of Intellectual Property Rights, London: Edward Elgar Publishing.

Stallman, R., (2006). Don't Let 'Intellectual Property' Twist Your Ethos. Free Software Foundation. Available from <http://www.gnu.org/philosophy/no-ip-ethos.html> [Accessed 4-15-09]

Stallman, R., (2008). The GNU Manifesto. Free Software Foundation. Available from: <http://www.gnu.org/gnu/manifesto.html> [Accessed 4-15-09]

Tavani, H.T., (2005). 'Locke, Intellectual Property Rights, and the Information Commons', Ethics and Information Technology, 7 (2) June, 87–97.

Warwick, S., (1999). Is Copyright Ethical? An Examination of the Theories, Laws, and Practices Regarding the Private Ownership of Intellectual Work in the United States' In: Spinello, R., Tavani, H.T., ed. Readings in Cyberethics. Sudbury, MA: Jones and Bartlett, 263-279.

From Public Data to Private Information: The Case of the Supermarket

Vincent C. Müller*
Anatolia College/ACT

Abstract:

The background to this paper is that in our world of massively increasing personal digital data any control over the data about me seems illusionary – informational privacy seems a lost cause. On the other hand, the production of this digital data seems a necessary component of our present life in the industrialized world. A framework for a resolution of this apparent dilemma is provided if by the distinction between (meaningless) data and (meaningful) information. I argue that computational data processing is necessary for many present-day processes and not a breach of privacy, while collection and processing of private information is often not necessary and a breach of privacy. The problem and the sketch of its solution are illustrated in a case-study: supermarket customer cards.

Keywords: Privacy, data, information, meaning, digital world, supermarket, customer card

Privacy and private information in the digital world: A framework

The general problem: Privacy in the digital world is a lost cause

Concerns about informational privacy have become more urgent with the advent of widespread digital computer technology. This concerns two types of activities. First, matters that I, as an individual, now do with computers were normally done with other media some 25 years ago (e.g. word processing, mail, telephone, photographs & videos, reading news, address book, diary, bibliography, music, e.tc. e.tc.). Apart from these activities that are obvious to me, since I myself per-

* Vincent C. Müller is Associate Professor of Philosophy at ACT/Anatolia College, where he has been teaching since 1998. He studied Philosophy with General Linguistics and History at the universities of Marburg, Hamburg, King's College London and Oxford. He received his PhD from the Centre for Cognitive Science of the University of Hamburg. Müller's publications focus on the problems of categorization, vagueness, symbol-grounding, theory of computing, computational intelligence and computer ethics. He is working on a book project about basic problems of artificial intelligence. Müller is the coordinator of the FP7 project EUCogII, a network of researchers on artificial cognitive systems <http://www.eucognition.org>

form them on a digital system, a very large amount of data is produced by other non-computational activities of everyday life, e.g. driving (car tracking [GPS, navigation systems, toll systems], road tolls, video surveillance, insurance data, police data, ...), speaking on the telephone (connection data, cost data, location, content storage, ...), taking photographs (image, time & date, technical data, location, ...) or just buying things (credit card data, bank data, customer data, location, ...) – not to mention the data produced by as yet uncommon applications, such as ‘ambient intelligence’, remote health services, e.tc. These developments are clearly increasing: more systems become integrated (e.g. telephone-computer-camera-music player) and more systems become computerized. In addition to this production of data, the possibilities of analyzing and storing them forever have greatly increased.

Let me make it quite clear that I do not generally deplore these developments: I am quite glad that I can now learn of an interesting academic conference, get further information, register, read literature, send an abstract, receive reviews, communicate with colleagues, write a formatted text, revise it, e.tc., all without leaving my desk, at negligible cost and with no delays. Of course, some of the developments sketched above are driven by profit interests and political interests that I happen not to approve of (e.g. state surveillance), but it is equally evident that a lot of these are in my personal interest and in that of others.

On the other hand, this digitalization means that a lot of information about me is now accessible from other people and that I have lost control of this information. In other words, the developments bring with them a loss of informational privacy. I use this term in the classic sense: “Privacy is the claim of individuals, groups or institutions to determine for themselves when, how and to what extent information about them is communicated to others” (Alan F. Westin 1967, 7, quoted from Rosenberg, 2004, 349).

This is where the dilemma lies: I want privacy and I want a digital life. It seems that I can’t have both and it seems that I can’t even reject the latter. In other words, it seems that privacy is a lost cause in the digital world.

The framework

In this paper I wish to sketch a general framework that could provide at least a partial way out of this dilemma. The framework relies with a philosophical distinction and I will then investigate in a case-study whether that framework holds water in practice.

A digital computer system operates according to algorithms on tokens of types that are in specific digital states and produces further tokens of these types (for details on digital states, see Müller, 2008). In current systems these are typically

just two states; the systems are binary. These basic states are used to represent higher level states, e.g. numbers, letters or truth-values, and these higher level states are again used to represent higher level properties, e.g. bank account balance, a beep, or an error message. So, if I send an SMS message saying “I will be back at 5”, this message is in digital states on several levels and at its destination it will hopefully produce a particular pixel image on a certain mobile phone screen. Is there a privacy problem here? If the message is just a sequence of binary data that is algorithmically processed by several conventional computing machines, then this message has no meaning for these machines, nothing is understood. If there is no meaning, then no information about me is conveyed and my privacy is wholly unaffected by the process. Our right to privacy is geared to human relations, i.e. it concerns what another human (or at least intelligent agent) can understand, and nothing of the sort is taking place here. Contrast this computational processing with the case of a Morse telegraph where human operators on both ends must encode and decode the message – normally understanding it in the process.

Generally, I think we must distinguish between data, meaningless digital states that are processed, and information, which is meaningful for someone. The qualification ‘for someone’ is needed to capture cases where something is data for one person but information for another – if I look at a road sign written in Georgian script, this is just data for me, for a Georgian speaker it is information. (Incidentally, in this and many other cases, it might not even be possible to distinguish the digital data, e.g. to copy the sign in such a way that it would distinguish the letters properly.)

If this is the right framework, then privacy is only an issue if persons come into play, at least at some point. This point might be after a lot of digital data processing, e.g. if the NSA processes my SMS as part of data mining for counterterrorism surveillance – an issue I investigated in (Müller, 2009). Given that the data is turned into personal information at some point, a privacy issue does arise in these cases.

Quite simply, mere data becomes information when it has meaning to someone. This can happen to personal data that I would like to keep private (in surveillance), but it can also happen to data that I control only to some extent, or even not at all. What happens when I buy something at the local supermarket is data that I do not control, so I will investigate the general framework of data vs. information in the case study of supermarket customer cards.

Case study: the local supermarket

What supermarkets do for marketing

Supermarket “customer cards” or “loyalty cards” are given away to customers with a promise of participation in special rebates or cash-returns to loyal customers. Of course, no guarantee is given that the owners of the card actually benefit. In fact, some studies have found that rebates were typically lower than before the introduction of the cards (Albrecht, 2001, 536). Crucially, customers are not informed – and often not aware – that their shopping data is collected.

It is apparently far from clear that these schemes do much work in terms of customer loyalty (McIlroy and Barnett, 2000), especially in areas like supermarket shopping where customers are not strongly related to a particular business and thus ‘relationship marketing’ is likely to be less successful (Boedeker, 1995). So, the actual point of the ‘loyalty’ card is less customer loyalty than data collection. This point the customer is typically not informed about.

The supermarket will store two kinds of data, and link them: data on the customer and data on the purchases.

The card stores identifying data, as far as this is accessible to the supermarket. This will typically include full name and address, often a telephone number and e-mail address, the date of birth and gender. This set of personal data is sometimes supplemented by financial data, if the card is combined with credit or debit card functions. The card may also be combined with further uses, e.g. insurance coverage or employee benefits, that imply further personal data. In practice, any kind of card can be used that identifies the customer, so a credit card or even an ID card would do. The more certain the identification, the better the integration with other data.

At the moment of purchase and with the help of the card, the supermarket will store the customer ID, transaction date & time, item bought (with number of items, classifications, price levels, promotions, relations to other items, e.tc. e.tc.), means of payment and other information that it finds relevant. It may combine that information with other data, such as personal movement through the store or even the fate of purchased items outside it (e.g. through RFID).

What the supermarket is trying to achieve is something that other businesses already have, by design. If sales take place over electronic devices, e.g. web sites then customers are (at least partially) identifiable through customer numbers, credit cards, ‘cookies’ or static IP numbers, and thus the business already has all that data on customers and purchases available. This is standard practice on the Internet. (If you have any doubts, check the cookies in your web browser.)

Apart from being used, this data is also stored indefinitely, for uses still to be determined. This work is normally outsourced to specialized IT firms, who also sell the data to other interested parties.

A sketch of the problem: Data mining on customer and purchase data

Imagine that there was a person at your neighborhood store who noted down the name of each customer and what they bought and when. I presume you would not be happy. – This is precisely what the customer card does, at least on the level of data. Here is an example of resulting transaction data, depicted in a binary matrix (Hand, Mannila and Smyth, 2001,8):

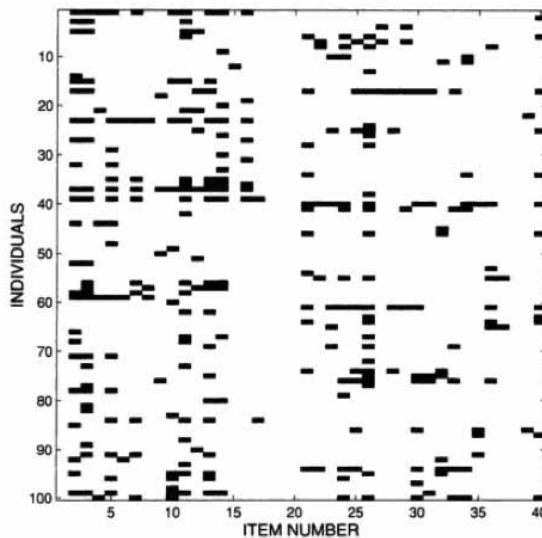


Figure 1.1 A portion of a retail transaction data set displayed as a binary image, with 100 individual customers (rows) and 40 categories of items (columns).

A typical response to this kind of activity from privacy concerned academics is this very useful study: “The shift from a paper-based to an electronic-based society has dramatically reduced the cost of collecting, storing and processing individuals’ personal information. As a result, it is becoming more common for businesses to ‘profile’ individuals in order to present more personalized offers as part of their business strategy. While such profiles can be helpful and improve efficiency, they can also govern opaque decisions about an individual’s access to services such as credit or an employment position. In many cases, profiling of

personal data is done without the consent of the target individual.” (Camenisch, Sommer, Fischer-Hübner et al., 2005, 20).

Having said that, ‘profiling’ particular individuals is not normally the point of customer cards, but rather the support of general and particular business decisions about store supplies, store location, product presentation, pricing, e.tc. For these purposes, the data sets produced must be subject to analysis, typically using techniques of data mining. “Data mining is the discovery of interesting, unexpected or valuable structures in large datasets” (Hand, 2007, 621) in particular, it is used in data that does not yet provide the structure that one is looking for. It looks roughly for two kinds of things:

A global model in data mining is a statement that applies to all the data, thus says something about each data set. A global model does not turn any particular data into information but says something about all data – e.g. whoever bought item 2 had a high probability of also buying item 3.

What is more interesting for our purposes is what is called a local pattern where a specific set of data is singled out because of ‘interesting characteristics’ (this technique will also produce “false negatives” and “false positives”). One aim of pattern data mining is to find; “association rules” (who buys one type of items also buys another type), the classical algorithm for which is described in (Jong Soo, Ming-Syan and Philip, 1995). These ‘patterns’ can help the business decide which items to stock in which store and how, it can help planning for seasonal items, e.tc. e.tc.

A problem that remains with this technique is that it only finds correlations, not causal relations. It is thus never certain that a certain correlation is not accidental – though the probability of such findings can be minimized by large and diverse enough data sets that offer enough repetitions (rather than one-off events) (Hand, Mannila et al., 2001, esp. 119). Many technical papers deal with the details of handling these problems in the supermarket environment, e.g. (Lawrence, Almasi, Kotlyar et al., 2001).

Who’s data? Who’s information?

This data is property of the supermarket, it can and is thus sold if this provides a profit. (Some of the data is anonymized before the sale, which may or may not assure its anonymity.) It may also fall into other hands, e.g. lawyers, law enforcement, secret agencies, e.tc. The data is in the hands of supermarket employees, not especially protected by state agencies or procedures, so misuse and errors are certain to occur. What can be done with this data is unlimited. For example, the UK Internal Revenue Service has demanded access to customer data in order to verify income statements in tax records (Albrecht, 2001, 539).

The urgency of our problem can be underlined by the combination of two remarks: “You are your information’, so anything done to your information is done to you, not to your belongings” (Floridi, 2006, 111) and the old saying “Tell me what you buy and I tell you who you are!”

Having said that, is there any personal information in the matrix depicted above? No, we are just given a set of unidentified individuals and a set of unidentified items. This is just data. Presumably it is necessary for the purposes of the supermarket to identify the items, so the system will contain a list which maps the numbers to products for sale. Still, as long as the individuals are not identified, this is not personal information. However, if there is a database of all customer cards, then the supermarket has a matching tool for combining this information in such a way that it does become personal information (Mr Smith buys a lot of alcohol, e.tc.). Thus, if there is such a database – and in all schemes I know of there is – then common privacy concerns apply, i.e. individuals must be informed and asked for consent to the spread of this information. In this case, the collection of personal information does not even seem particularly useful, except that it helps the supermarket find out the location of customers (for store location). All other aims can be reached with an anonymous card that give the benefits to the person who happens to present it. The collection of data is useful, the collection of information is superfluous.

Conclusion: The right to information, not data

The increasingly digital world poses serious challenges with its production of large amounts of personal digital data. However, there are technical and political means to reduce the accumulation of this data, and, more importantly, to prevent the turning of this data into personal information. This distinction between data and information has implicitly been used already in many cases.

The little case study of supermarket customer cards has shown that this distinction can be made in practice and that it can provide for a means to achieve the aims of the data-collector while respecting the right to privacy of the customers. This suggests that political work should go into the direction of turning information into data, rather than just fighting data collection per se (which remains an important aim, however). Philosophically and practically, we will need to sharpen the distinction in such a way that it can cover the many diverse cases and needs.

REFERENCES

Albrecht, K., (2001) Supermarket cards: the tip of the retail surveillance iceberg. *Denver University Law Review*, 79(4), 534–539, 558–565.

Boedeker, M., (1995) Relationship marketing and regular customer cards: daily product retailing in Finland. *Marketing Intelligence & Planning*, 15(6), 249 - 257.

Camenisch, J., Sommer, D., Fischer-Hübner, S., et al. (2005). Privacy and identity management for everyone. DIM'05 (Digital identity management), Fairfax, Virginia, ACM, New York, 20-27.

Floridi, L., (2006) Four challenges for a theory of informational privacy. *Ethics and Information Technology*, 8, 109-119.

Hand, D.J., (2007) Principles of data mining. *Drug Safety*, 30(7), 621-622.

Hand, D.J., Mannila, H. & Smyth, P., (2001) Principles of data mining. MIT Press, Boston.

Jong Soo, P., Ming-Syan, C. & Philip, S.Y., (1995). An effective hash-based algorithm for mining association rules. Proceedings of the 1995 ACM SIGMOD international conference on management of data, San Jose, California, ACM, 175-186.

Lawrence, R.D., Almasi, G.S. & Kotlyar, V., et al. (2001) Personalization of supermarket product recommendations. *Data Mining and Knowledge Discovery*, 5(1-2), 11-32.

McIlroy, A. & Barnett, S., (2000) Building customer relationships: do discount cards work? *Managing Service Quality*, 10, 347-355.

Müller, V.C., (2008). Representation in digital systems, in *Current issues in computing and philosophy*. eds. A. Briggie, K. Waelbers and P. Brey). IOS Press, Amsterdam, 116-121.

Müller, V.C., (2009) Would you mind being watched by machines? Privacy concerns in data mining. *AI & Society*, 23(4), 529-544.

Rosenberg, R.A., Ed. (2004). *The social impact of computers*. Elsevier, San Diego.

Information ethics in business organisations

Kioshi Murata^{*}

Centre for Business Information Ethics,
School of Commerce, Meiji University

Abstract

Recently, a large majority of development, deployment and use of information and communication technology (ICT) has been carried out by business organisations. Because their behaviour has had strong influence over a wide range of people in the modern market economy and ICT can function as an amplifier of power imbalance, they have to carefully engage in activities of development, deployment and use of ICT and take their social responsibility regarding intentional as well as unintentional outcomes of the activities. As major providers of the technological architecture of the ICT-dependent society, they are expected to overcome nosism of them and to address ethical and social issues caused by their development, deployment and use of ICT proactively. However, the fact that business organisations operate in competitive markets often makes it difficult for them to develop an ethical outlook appropriately. In order to get over the difficulty, correct understanding of business ethics and corporate social responsibility is necessary, which requires reconsidering definitions and roles of business organisations, the market economy, liberalism and capitalism from the social viewpoint.

Keywords: Information ethics, Business organisations, ICT-dependent society, Market economy, Business ethics, Corporate social responsibility

Introduction

Recently, especially in developed countries, a large majority of development, deployment and use of information and communication technology (ICT) has been carried out not by governmental or military organisations but by business organisations. Because their behaviour has had strong influence over a wide range of

* Kioshi Murata is a director of Centre for Business Information Ethics and a professor of management information systems at School of Commerce, Meiji University, Tokyo, Japan. His research interest is in information ethics in business organisations including privacy, ICT professionalism, surveillance and gender issues. He is an international research associate at Centre for Computing and Social Responsibility, De Montfort University, Leicester, UK and an editorial board member of *Journal of Information, Communication and Ethics in Society*, *Journal of the Japan Society for Management Information* and *Journal of Information and Management*.

people in the modern market economy and ICT can function as an amplifier of power imbalance, they have to carefully engage in activities of development, deployment and use of ICT and take their social responsibility regarding intentional as well as unintentional outcomes of the activities. As major providers of the technological architecture of the ICT-dependent society, which is one of the four modalities of regulation of human behaviour Lessig (1999) identifies, they are expected to behave accountably and address ethical and social issues caused by their development, deployment and use of ICT proactively. Unfortunately, however, it is not necessarily easy for business organisations to recognise ethical and social issues related to ICT correctly and to address them properly. This study attempts to examine why it is difficult for business organisations to cope with ethical and social issues brought about by development, deployment and use of ICT and how businesspeople can be motivated to address the issues appropriately getting over the difficulty.

The remainder of this paper is structured as follows. The next section describes the necessity and importance of examining information ethics in business organisations. After that, difficulties for business organisations in addressing issues of information ethics are discussed followed by the section which examine challenges for information ethics in business organisations and considers ways of surmounting them.

Why information ethics in business organisations?

Business Organisations as Entities of Influence in the ICT-dependent Society

We have observed the advent of the information society in which business organisations play a central role. Since the early days of the development of computers, governmental and military organisations had led the development of ICT and computerisation of society. Indeed, computers such as Colossus, Harvard Mark I and ENIAC were products of the WWII and ARPA Network was one of the Cold War; a huge amount of investment from military budgets had been justified by the threat of wars. However, a large majority of investment in order to develop and deploy ICT/ICT-based information systems is now carried out by business organisations. Simultaneously, it seems that almost all modern businesses premise use of ICT as an integral part of their business processes and operations owing to the rapid development and widespread availability of ICT.

Due to logical malleability and information enriching nature of ICT (Moor, 1985; 1998) and seemingly unlimited permeation of business activities into every corner of society, which has been enhanced by the market-based economic policies many industrial nations' governments have adopted, development, deployment and use

of ICT by business organisations have significantly influenced the way and quality of life of a wide range of people. Even though it is often alleged that widespread availability of the Internet-related services including search engine services and consumer/user generated media (CGM/UGM) has enabled individuals to enjoy freedom of speech, the freedom is limited by the technological architecture business organisations involved in the services have developed. Accordingly, business organisations have become “entities of influence” in today’s ICT-dependent society through development, deployment and use of ICT and their power can be amplified by using ICT.

Modern Practical Syllogism and Responsibility of Business Organisations

Imamichi (1971; 1989; 1990; 1998) described the necessity of developing appropriate ethics for the current technological society in his insightful studies of eco-ethics. He points out that the practical syllogism described in Aristotle’s *Nicomachean Ethics* must be reconsidered in the modern eco-environment or human habitat, which is composed not only of nature, but of “technological conjunction”. The classical form of practical syllogism is as follows¹.

Major premise: A is desirable.

Minor premise: p, q, r and so on realise A.

Conclusion: For some reason, I choose p as the means to achieve A.

Here, the ideal goal is obvious and the minor premise is the horizon of the freedom of choice, the object of which is a means to achieve the goal. Because the major premise is the peremptorily given condition, worth of practical human action is dependent on which means is chosen.

This form of practical syllogism remains valid in our individual decisions even today. However, the logical superiority of goals over means had morally justified the rapid development and autonomous deployment of technology. Consequently, the continuous development and deployment of technology as a heteronomous means had created the autonomous technological conjunction which includes impersonal institutionalisation of human beings and the restraining force of means over goals had grown.

Due to the rapid progress of technology and the advent of the technological society, means is now more important than goals. For people living in the modern age, the obvious given condition has been technological functions which can realise a variety of goals and the logical structure of the practical syllogism has been reversed.

1. See 1111b 26-27, 1112b 15-17 and 1144a 31-32 of *The Nicomachean Ethics*.

Major premise: We have means or power P.

Minor premise: P can realise goals a, b, c and so on.

Conclusion: We choose a as the goal of P for some reason.

Obviously, powerful means exist, including nuclear, electric and electronic technology. Goals attainable using these means are considered analytically, and hence the means control the goals. The means are so powerful that a selection of a particular goal may have considerable influence on society. However, these sorts of means are usually controlled by groups or organisations, not by individuals; the subject in the modern form of practical syllogism described above is “we”. This tends to result in a confusion of responsibilities for goal setting. The problem here is the nosism of organisations, not egoism of individuals.

The Imamichi’s argument can be interpreted in the context of modern business organisations as follows. ICT is obvious means for almost all business organisations in the modern business environment where ICT is an almost indispensable component of successful business processes. Because of the versatility of ICT, business organisations can set various goals development, deployment and/or usage of ICT can realise subject to budgetary constraints. The goals can be, for example, cost saving, increased customer satisfaction, strengthen business partnership or, even, profit improvement through illicit business activities. Indeed, in the incident of Kanebo’s account rigging in 2005, the application software packages for auditing and making out annual security reports was misused and the software made it easier for four certified public accountants working for Chuo Aoyama Audit Corporation, who were arrested, to conduct their crimes and made it harder for the crimes to be detected². Due to the invisibility of computing (Moor, 1985), ICT can enhance asymmetry of information with respect to intentions and activities of business organisations between organisational members and non-members, which has been considered as a major cause of business organisations’ moral hazards.

In addition, situations in which utilisation of ICT can be considered as reflection of the reversed practical syllogism and are ethically controversial have been observed in recent days. For example, a vehicle video system with a recording facility, which was originally developed to provide forensic evidence to resolve disputes with respect to car accidents or car-related crimes quickly, has been introduced to office space as part of information security systems. The vehicle video system is used to improve labour productivity of office workers as well through workplace monitoring with the lovely tiny eye and ear of a video camera to prevent the workers from making a wasted motion and an idle talk with their colleagues.

2. Kanebo is a major chemical and cosmetic company established in 1887 and Chuo Aoyama Audit Corporation is one of the big four audit corporations in Japan.

The individual tracking system using a GPS locator-equipped mobile developed for target marketing, which is now in the process of implementation, is another example. With other types of ICT such as RFID, a contactless IC smart card and personal information databases, this system is expected to provide business organisations with really efficient and effective functions of personalised target marketing at all hours, whereas this system may cause invasions of privacy of individual mobile users.

As entities of influence, modern business organisations have to take social responsibility for their development, deployment and use of ICT suitable for their power. They are required to address ethical issues and problems with respect to ICT proactively. The modern form of practical syllogism suggests that they have to deliberate carefully about what they develop, deploy and use ICT for considering social influence of their choice.

Difficulties for business organisations in addressing issues of information ethics

Ethics in the Market Economy as an Autonomous System

The development of ethical behaviour of modern business organisations with respect to development, deployment and use of ICT is vital to overcoming nosism of them and eliminating the ambiguity over their social responsibilities for ICT. However, the fact that business organisations operate in competitive markets often makes it difficult for them to develop an ethical outlook appropriately.

Ohba (2004) points out that in the market economy system, which has been separated from communities as an autonomous system long before, information related to a price signal such as of profit, cost, productivity and competition tends to be considered as a matter of concern and, in contrast, any other information as just a noise. In the current business environment where market-economy principles have spread globally, a vast majority of business scholars as well as practitioners has seemed to believe that just information concerning a price signal is significant for business organisations. In fact, Coase (1937) considers that the nature of business organisations can be explained by the two kinds of cost, transaction and opportunity cost, and standard textbooks of microeconomics describe that a business organisation is a profit maximiser subject to their own production technology, initial endowments and budget constraints. In the logic and vocabulary of market economy, the spirit of fair play based on sympathy which Adam Smith, the founder of economics, considered every player in a competitive market was required to possess seems to have no place and addressing ethical/social issues may be considered as a cost factor. Actually, the term of compliance depression has recently been used among businesspeople in Japan as a consequence of the enforcement of the Act on Protection of Personal Information (Act No. 57 of 2003) and the Revised Financial

Instruments and Exchange Act (Act No. 65 of 2008) which is called J-SOX Act and requires business corporations to submit internal control reports in addition to annual security reports.

Such belief can be observed in the arguments about ethics and economics. Sen (1967) demonstrates that to get out of Pareto-inferior, which depends on what each individual expects about the others' action, all that is necessary is that each individual is assured that the others are doing the "right" thing, and then it is in one's own interest also to do the "right" thing. Arrow (1971) describes that norms of social behaviour, including ethical and moral codes, are reactions of society to compensate for market failures and might be interpreted as agreements to improve the efficiency of the economic system through lowering transaction cost. Casson (1993) maintains that good culture like integrity, honesty and altruism brings about decrease in transaction cost and, in the result, activates business activities and economy.

In the study area of business administration, a close relationship between trustworthiness and reputation of business organisations and economic performance of them is examined theoretically (For example, Barney and Hansen, 1994) as well as empirically (For example, Waddock and Graves, 1997; Dowling, 2001). Porter and Kramer (2002) propose the idea of the context-focused approach to philanthropy; "corporations can use their charitable efforts to improve their competitive context, the quality of the business environment in the location or locations where they operate" and "using philanthropy to enhance context brings social and economic goals into alignment and improves a company's long-term business prospect". They consider that "the acid test of good corporate philanthropy is whether the desired social change is so beneficial to the company that the organisation would pursue the change".

These arguments would be true descriptively. However, these may be somewhat misleading from the normative viewpoint; the arguments may lead businesspeople to hold the belief that ethics and social responsibility can be an instrument for business organisations to gain profit or to save cost.

Confusion and Delusion over Business Ethics and Corporate Social Responsibility

Business organisations in the market economy system usually tend to be thought to make decisions based on their productivity and economic efficiency in order to improve their profit, even though business ethics and corporate social responsibility (CSR) have recently become a popular topic of conversation. In the past, these topics had often recognised as oxymoron or, even, hypocrisy. Carr (1968) claims that violations of the ethical idea of society are common in business and business organisations have the legal right to shape their strategy without reference to anything but their profit as long as they don't violate the rules of the game set by law.

Friedman (1962) pronounces that the doctrine of social responsibility is a fundamentally subversive one in a free society and one and only one social responsibility of business is to use its resources and engage in activities designed to increase its profit so long as it stays within the rules of the game.

Friedman's argument has not been confuted completely to date. Under this somewhat muddled situation, confusion and delusion over business ethics and CSR among researchers as well as practitioners have been observed. They seem to fail to understand correctly why business organisations are required to behave ethically and take their social responsibility and what the nature of business ethics and CSR is. For example, CSR Initiative Committee (2005) proposes the following CSR management philosophy:

Corporations are expected to make a positive contribution to the advancement of the society while working to prevent the eruption of corporate scandals with a view to encouraging the sound development of their own organisations and society. In order to execute their commitment to CSR, corporations strive for forging favourable relations with, and thus winning enhanced confidence from, a variety of stakeholders existing around themselves, including consumers, suppliers and customers, employees, shareholders, local communities and international communities at large, through sincere corporate practices that adhere to the CSR Charter of Conduct and CSR Code of Conduct stipulated below, in pursuit of sustainable development for both themselves and the society (pp. 20-22).

This statement reflects their recognition that the core of CSR is compliance with laws and rules. However, forcing socially responsible business organisations or corporate "citizens" to blindly comply with law and rules is sometimes dangerous from the social standpoint; compliance may be used as an excuse of ethically controversial behaviour of the organisations or they may be encouraged to struggle to find good ways of slipping through the meshes of laws and rules. In addition, citizens' blind compliance with laws and rules may permit arbitrary law-and rule-making.

The misunderstandings about business ethics and CSR can be classified into two categories. One is the "long-term profit doctrine" which is based on a mindset that business organisations have to engage in business ethics and CSR in order to maximise or ensure a long-term profit. For example, the Japanese Federation of Economic Organisations (Nippon Keidanren, 2004), one of the strongest lobbying groups in Japan, expresses their view that CSR is a source of corporate competitiveness and brings about improvement of corporate value. This is directly related to the concept of enlightened self-interest and cannot be accepted as rationale as well as a practical foundation of business ethics and CSR, because, as a logical consequence, this doctrine should recommend business organisations to get involved in unethical behaviour when doing so ensures a bigger long-term

profit than one attained by behaving ethically. This reminds us of Thrasymachus' claim that "perfect immorality is more rewarding than perfect morality" and "criminals are clever, good people if their criminality is able to manifest in a perfect form and they are capable of dominating countries and nations" described in Plato's Republic (347e-348e).

The other is the "social responsiveness doctrine" or "external pressure doctrine" which asserts that business organisations have to behave ethically and take their social responsibility because, unlike in the past, society has already become to require business organisations to do so. This doctrine is also unacceptable. Were business organisations permitted to behave unethically in the past? Should they be admired when they respond to an immoral requirement of society? The doctrine fails to recognise that business organisations propose values to society through providing their products and services and, therefore, they are intrinsically moral entities.

Indeed, behaving unethically is self-defeating for business organisations. When immoral or ethically questionable business behaviour is observed, regulations on business activities would be strengthened. Otherwise, decent organisations would exit from markets. Anyhow, the free-market economy system would collapse. Even when unethical behaviour of a business organisation is not recognised by people outside the organisation, those who commit the behaviour recognise it and, because of invisibility of other organisations' behaviour, they would have to be suspicious of whether the other organisations behave fairly. This would result in lack of trustworthiness of business organisations and malfunction of laws, rules and social norms in business transactions.

Challenges for information ethics in business organisations

The confusion and delusion over business ethics and CSR expose challenges for information ethics in business organisations. Sincere engagement in both business ethics and CSR on the basis of correct understanding of them would provide business organisations with good clues to how they can approach ethical and social issues caused by development, deployment and use of ICT. However, the confusion and delusion described above may be an obstacle to encouraging businesspeople to cope with issues and problems of information ethics appropriately. For example, necessity of development of a professional outlook of ICT engineers would not be recognised because it is costly and return or benefit of establishing ICT professionalism in business organisations can be hard to be evaluated in terms of money. Simultaneously, development of ICT professionalism in business organisations has not clearly been required by society.

Another challenge is derived from asymmetry of information inherent in business activities. Because they operate in competitive markets, they cannot necessarily

disclose every piece of information regarding development, deployment and use of ICT, though these may cause irreversible transformation of societies. This implies that business organisations should develop a proper information disclosure scheme concerning ICT as well as a satisfactory level of professionalism among employees.

In order to surmount these challenges, we need to think anew about definitions and roles of business organisations, the market economy, liberalism and capitalism from the viewpoint of flourishing the whole society. Information ethics in business organisations has to shift its focus of examination from micro ethics to macro one.

Conclusions

In today's ICT-dependent society, business organisations have become entities of influence through development, deployment and use of ICT. Therefore, they are required to take social responsibility with respect to ICT suitable for their power. However, because of the fact that they operate in competitive markets and confusion and delusion over business ethics and corporate social responsibility, it is difficult for business organisations to cope with issues and problems of information ethics appropriately. In order to overcome the difficulty, correct understanding of business ethics and corporate social responsibility is necessary, which requires rethinking definitions and roles of business organisations, the market economy, liberalism and capitalism from the social viewpoint.

Acknowledgement

This study was supported by an open research centre project for private universities entitled "Quality-oriented human resource development and smart business collaboration: management quality science": a matching fund subsidy was provided by MEXT (the Ministry of Education, Culture, Sports, Science and Technology), 2007-2012.

REFERENCES

- Arrow, K. J.** (1971) Political and economic evaluation of social effects and externalities, in *Frontiers of Quantitative Economics* (ed. Intriligator, M. D.). North Holland, Amsterdam, 3-25.
- Barney, J. B. and Hansen, M. H.** (1994) Trustworthiness as a source of competitive advantage. *Strategic Management Journal*, 15, 175-190.
- Carr, A.** (1968) Is business bluffing ethical? *Harvard Business Review*, 46 (1), 143-153.
- Casson, M.** (1993) Cultural determinants of economic performance. *Journal of Comparative Economics*, 17, 418-442.

- Coase, R. H.** (1937) The nature of the firm. *Economica*, 4 (16), 386-405.
- CSR Initiative Committee (2005) CSR Initiative: CSR Management Philosophy, Charter of Conduct and Code of Conduct. Japanese Standards Association, Tokyo.
- Dowling, G.** (2001) Creating Corporate Reputations: Identity, Image and Performance. Oxford University Press, Oxford.
- Friedman, M.** (1962) Capitalism and Freedom. The University of Chicago Press, Chicago, IL.
- Imamichi, T.** (1971) Self-plasticity of Identity. The University of Tokyo Press, Tokyo (in Japanese).
- Imamichi, T.** (1989) The concept of an eco-ethics and the development of moral thought, in *Man and Nature: The Chinese Tradition and the Future* (eds. Tnag, Y., Lie, Z. and McLean, G. F.). The Council for Research in Values and Philosophy, available at http://www.crvp.org/book/Series03/III-1/chapter_xv.htm (accessed 6 February 2008).
- Imamichi, T.** (1990) *Eco-Ethica*. Kodansha, Tokyo (in Japanese).
- Imamichi, T.** (1998) Technology and collective identity: Issues of an Eco-Ethica, in *The Humanization of Technology and Chinese Cultures* (eds. Imamichi, T., Wang, M. and Liu F.). The Council for Research in Values and Philosophy, available at http://www.crvp.org/book/Series03/III-11/chapter_i.htm (accessed 6 June 2008).
- Lessig, L.** (1999) *Code and Other Laws of Cyberspace*. Basic Books, New York.
- Moor, J. H.** (1985) What is computer ethics? *Metaphilosophy*, 16 (4), 266-275.
- Moor, J.H.** (1998) Reason, reality, and responsibility in computer ethics. *Computers and Society*, 28 (1), 14-21.
- Nippon K.** (2004) The Basic Concept of Promotion of Corporate Social Responsibility. Available at <http://www.keidanren.or.jp/japanese/policy/2004/017.html> (accessed 29 November 2008; in Japanese).
- Ohba, T.** (2004) *The Myth of Ownership*. Iwanami Shoten, Tokyo (in Japanese).
- Porter, M. E. and Kramer, M. R.** (2002) The competitive advantage of corporate philanthropy. *Harvard Business Review*, 80(12), 57-68.
- Sen, A. K.** (1967) Isolation, assurance and the social rate of discount. *The Quarterly Journal of Economics*, 81 (1), 112-124.
- Waddock, S. A. and Graves, S. B.** (1997) The corporate social performance-financial performance link. *Strategic Management Journal*, 18 (4), 303-319.

Blogs and privacy in Seken as a Japanese life-world including indigenous moral norms

Makoto Nakada*
University of Tsukuba

Abstract

This paper examines the positions of Japanese blogs and Japanese views on privacy in a broad cultural context called Seken as a Japanese life-world, focusing on empirical findings of the researches the author and his research group conducted in 2007 and 2008. These two researches are a continuation of our previous researches in the last ten years in which the author tried to understand the characteristics of Seken and Japanese indigenous morality based on it. Through these researches including the newer ones, we found that Japanese people still live within Seken which provides Japanese people with implicit and/or explicit indigenous moral norms. In the author's view, Seken is based on Buddhism, Shintoism, Confucianism and Japanese traditional culture. Within this framework or a horizon, a spiritual interpretation of nature is predominant together with cultural memory of wars and disasters as well as an orientation to good human relations founded on such values as sincerity and purity of mind. But, according to the author, Japanese people live simultaneously also in Shakai world which is a modernized and rationalized aspect of Japanese life-world influenced by Western culture(s) particularly since the 19th century.

This dualism has been empirically confirmed by researches done by the author and his group (Nakada et al., 2004; Nakada, 2005; Nakada and Tamura, 2005; Nakada, 2007a; Nakada, 2007b). For example, in the case of the research done in 2006 among some 500 male and female respondents between 20-49 years old, the percentage of the respondents who said "agree or somehow agree" to one of the Seken-related views, "Within our modern lifestyles, people have become too distant from nature" is 78.2%. The majority of the respondents showed sympathy with the Seken-related views likewise; "People will become corrupt if they become too rich" (83.2%), "People have a certain destiny, no matter what form it takes" (81.6%), "In our world, there are a number of things that cannot be explained by science" (89.6%), "Doing your best for other people is good for you" (76.4%), "The frequent occurrence of natural disasters is due to scourge of heaven" (scourge of heaven)" (55.0%).

* *Makoto Nakada*: Professor of University of Tsukuba (Graduate School of Humanities and Social Sciences, Doctoral Program in International and Advanced Japanese Studies). Member of ICIE. Editor of IRIE.

One of the most impressive findings about Seken in terms of its relation with information ethics is that Seken-related views have strong or fairly strong correlations with privacy-related views in the information era. In the case of the 2006G Research, 'criticism of material wealth' (one of Seken-related factors, i.e. a combination of several Seken-related views) has (statistically significant) correlations with a set of concerns for violation of privacy including 'abuse of name list,' 'exhibition of private diary without permission,' 'abuse of the resident registry network system,' 'exhibition of home address without permission,' 'abuse of cookies.' On the other hand, 'individualism' (sympathy with individualism and independency) has negative or no significant correlations with these concerns for violation of privacy. These findings suggest us that Seken-related views and privacy-related views/concerns occupy the same place or adjoining places within Japanese minds in the information era.

Our research in 2007 (2007G) among some 1200 male and female respondents between 20 and 49 years old, provides us with additional findings to the meanings of Seken in Japanese information society. Through our research findings on the relations between 'experiences and feelings when being engaged in writing blogs' and 'views on Seken' we could see that 'meanings of blogs' lie within a broad cultural context which might called Seken and that 'meanings of blogs' are closely tied with people's views on Seken.

According to the author's interpretation, in the case of Japanese respondents whose responses will be examined later in this paper, people's experiences of blogs or their attitudes towards blogs are (at least partly) determined by their past (or present) experiences in a Seken-related life-world. And we can also expect that at the same time such experiences might change the inner states of Seken or the inner relations of meanings and values within Seken at least to a certain extent.

In a way, we observe here two different levels of life-world experiences 'coming together,' namely the moral life-world experiences based on Japanese past culture, history, social-political structures, and traditional world views, on the one hand, and the experiences and values particularly enabled by modern information and communication technology on the other hand. If we use the terms by Rafael Capurro who has been introducing some Heideggerian or hermeneutics perspectives in the information ethics discussion, this coming together of different levels of experiences is a typical phenomenon of our being-in-the-world-with-others as a being-in-the-networked-world (Capurro, 2005). In Gadamer's terms, we might be able to regard this 'coming together' as a sort of "Horizontverschmelzung" or "fusion of horizons" in the information era.

What we have to take into consideration in regard to this 'coming together' is the foundation or a broader horizon within which it can take place. In a way, such foundation might be interpreted as the world openness itself (see: Capurro, 2005) but we have to look further this openness or universal horizon in order to relate these views with concrete and complex phenomena as shown in the tables in this paper. According to the author's views, Japanese attitudes towards blogs and privacy can't be divided from this openness or universal horizon in the information era.

In this paper, our discussions about these points will be largely dependent upon our two recent researches in 2007(2007G Research) and 2008(2008G Research).

Keywords: Seken, Privacy, Blogs, Life-world

Introduction

Japanese live in two different worlds or two different aspects of world/society: Seken and Shakai. In a very simple and predigested way, we can say that Shakai is a modernized and 'westernized' realm of Japanese life-world and Seken is a traditional realm of meanings and experiences based on 'orientation to traditional human relations' including highly evaluated values such as sincerity, purity of minds, modesty, a sense of shame, a sense of self-restraint, 'orientation to human-nature interrelations' (in this case 'nature' is considered to be sources from which beauty, productivity of culture, not-artificial ethical attitudes such as sincerity, purity of minds as well as obedience to nature come from), 'orientation to self-realization' (including self-restraint, cultivation of personality, return to oneness of person and nature).

In contrast to Seken, Shakai as a different horizon, another realm of Japanese life-world includes different meanings, attitudes, orientation to values, ethical views, although Seken and Shakai are related with each other in some ways (in fact, Japanese development in the modernized era wouldn't have been possible without combination of Seken and Shakai in a certain way).

In my own view, Japanese ways of communication or self-presentation through use of the Internet, particularly blogs, SNS or BBS, reflect this dichotomy and also the inner structures of Seken (which seem to influence the ways of 'fusion' of different meanings, views, values within it, i.e. Seken). This is the central point which I (the author) intend to emphasize in this paper theoretically and empirically. Therefore the contents of this paper consist of two separated and at the same time interrelated sections: one is about explanation or interpretation of Seken from theoretical viewpoints and another is the empirical examination based on the research data about the relations between 'the characteristics of

Japanese communication or self-presentation by using ICTs/the Internet' and 'the ways of their life in Seken.'

According to the findings of our previous researches (see Nakada, 2008), Japanese communication through the Internet is related with their attitudes towards privacy which seems to be influenced by the dichotomy, Shakai and Seken. This appears to be one of the major characteristics of Japanese communication through the Internet. (According to our previous researches) Japanese have contradictory attitudes towards privacy in the so-called information era: Japanese believe in the importance of respects for privacy in the informatized society and they also feel that openness of their privacy is something which 'activates' their communication with others on the Internet.

If we can see in this paper that Japanese views on privacy as well as their views on blogs are related with Seken-related views, we might have a chance to know in detail the broad cultural horizon within which Japanese understanding of privacy and views on blogs as well as their understandings of other ethical and value-related problems 'come together.' This means that we can understand the inner structure of Japanese Seken more deeply and broadly.

Nothingness and coming together of meanings

According to Rafael Capurro (Cappuro, 2003) who is in the tradition of Heidegger and Gadamer and is continuously trying to relate these ontological traditions to the problems of information society, especially ethical aspects of CMC or of people's mutual understanding through use of ICTs, 'messages can't be divided from messages/messengers relationships.' In my own view, aside from those sharing ontological understanding of this world, more specifically the meanings of information that can't be apart from 'angeletic' ('Capurro's term) aspects of communication, with him, the most qualified ones who are in the position from which they can understand this somewhat mysterious (at least for those with ontic attitudes towards this world or those with belief in 'reductionistic' world views) remark by Capurro, 'messages can't be divided from messages/messengers relationships,' might be those who can have strong interests in the following poem (Haiku) by Matsuo Basho(1644-1694).

Furu ike ya
kawazu tobikomu
mizu no oto
(an ancient pond / a frog jumps in / the splash of water)

In my own interpretation which, I believe, reflects Japanese traditional world-views on 'Mu'(nothingness) as a kind of fundamental source of meanings in our

everyday activities, within this poem or this event appearing in front of the poet through his own poetic expression, messages (in this case, the poetic expression of the situation) and messengers (in this case, the poet himself) can't be divided from each other. And it is also clear that the fact and the expression can't be divided from each other in this case, because we can imagine that the fact in this case occurs only through poetic connection(s) of flog, old pond, Basho himself, sound of jumping flog, Basho's ears and perhaps our own ears too. To put this in a different way, we can see here the connection(s) between message/messenger, facts/expressions, mono (the objects or beings)/koto (language, expressions, objects expressed by words), person (artist or audience)/ objects.

Whether we are clearly aware of these connections or not, it is true that these connections occur in front of our eyes, and in the case of Basho, his trial to make this awareness of these connections more clearly seems to let him come near to the point where these connections themselves happen from Mu(nothingness). If we follow Toshihiko Idsutsu's suggestions, in the case of Basho's poem, these connections are always within the fluid or active process of interchange of articulation of and non-articulation of things (mono and koto, or Mu and the beings).

One of the most important efforts or wishes for Japanese scholars, authors, artists, poets, thinkers in previous eras and also even in the modern era, I believe, has been to express these phenomena intuitively or accurately.

According to Toshihiko Idsutsu, Dougen, a famous Zen Buddhism priest (1200-1253) in Kamakura era of Japan, tried to bring Being, which is dried up by process of articulation of beings or by grasp of essence based on the process of articulation of beings, into the state of 'articulation of beings without grasping of essence' and he (Dougen) also tried to bring Being into its original fluency (Idsutsu, 1991).

Yujiro Nakamura (Japanese philosopher who attempts to combine Japanese traditional thoughts with western modern thoughts) suggests us that Kitaro Nishida tried to regain meanings of beings based on Mu (nothingness) or 'predicative substrata' which is in contrast with subjective substrata. In this sense, Mu is understood not as mere emptiness but as source of beings on which articulations of beings are founded (Nakamura, 2001).

Bin Kimura, Japanese psychiatrist who is influenced by Kitaro Nishida, Zen-Buddhism and Heidegger, says that in every case of our perception, we feel, if we carefully see what happens, that the objects of our perception have some kind of active selfness (a sense of presentation/appearance of its uniqueness or existence) and that this selfness of the objects is (considered to be) our own self-experiences on the other hand (Kimura, 1975, p. 6).

In my own view, what I have tried to focus on in some of my previous papers on perception, agnosia, imagination (Einbildungskraft and common senses combined with it) is an additional trial of this kind of approaches based on the interests in the 'emergence' of meanings in this world, taking a form of 'coming together' of different meanings and experiences (Nakada and Capurro, 2009).

Although the interests in these phenomenon of the connections of mono/koto, objects/observers, fact/expression are not confined to Japanese scholars, thinkers, artists(in fact, we can recall the works of Merleau-Ponty, Wolfgang Blankenburg, Max Sheler, Goethe, Jung and so on in this respect), but, on the other hand, as we will see in the discussions about Seken later in this paper, Japanese people are likely to be 'actively' aware of these phenomena, i.e. connections of beings, mono/koto, outer world/inner world, Mu/experiences, person/person.

In fact, i believe that, Seken is the realm of meanings or experiences where these connections as well as the newer connections of meanings of blogs, views on privacy and interests in politics can occur.

Openness and Closedness of Seken and Mu

As we will examine in a detailed way later in this paper, based on our empirical researches, if we use the term 'horizon' as suggested by Gadamer and also by Capurro, Seken, as a kind of 'horizon' of meanings or experiences, is characterized in two contrasting ways: it's openness and it's closed-ness. It is clear that Seken is an open horizon or a realm of meanings in which various levels of meanings or experiences are coming together or fusion of these meanings or experiences is made possible through a certain process. (In my own view, even within Gadamer's frame of discussions in regard to fusion of different horizons or combined experiences in a horizon, how and why 'fusion' or 'coming together' occurs still remains 'not-visible' to a certain extent or explicitly. In this sense, examination about 'fusion' or 'coming together' within Seken might enable us to see the phenomena of 'fusion,' 'coming together' more deeply and concretely). For example, as we will see later, 'warnings from heaven' and 'experiences in writing (in) blogs' are 'coming together' within Seken and 'communication in a self-telling form spurred by one of the most modernized technology' are found to 'come together' with different views within Seken too. On the other hand, within Seken, every thing (meaning, experience, evaluation, approval, denial, sympathy, interests, concerns) appears to be viewed from certain limited ethical points or values-related views: we should refrain from doing wrongdoings in order to avoid heaven's anger; we have to follow our somewhat pre-determined destiny. In this sense Seken is considered to be closed in regard to diversity of viewpoints within it. In fact, Japanese tend to see political problems from Seken-related viewpoints or Seken-related ethical eyes, typically in the case of judgement of actions of poli-

ticians ('Is this politician sincere?' and so on), but sometimes these tendencies are liable to narrow their (Japanese) political eyes, because some serious political problems can't be solved by depending (just) on purity of minds of politicians. (Japanese are sometimes unable to determine what they should do when they are facing serious political problems such as the dilemma between 'orientation to equality' and 'orientation to efficiency' in the global world.)

In this sense, we need to see the inner structures of *Seken* and *Mu*. (*Seken* and *Mu* seem to have some fundamental relations or similarities between them as broader horizons within which different meanings or experiences 'come together'.)

In the following passages we will see the findings of our two researches conducted in 2007 and in 2008 in Japan (the respondents are the Internet users), i.e. 2007G Research and 2008G Research. By seeing the findings of these researches we can see (at least partly) the inner structures of Japanese *Seken* as a sort of horizon or a cultural context.

Findings of 2007g research and 2008g research

The following figures of Table 1 show that majority of Japanese (strictly speaking, majority of the Japanese respondents) regard a variety of *Seken*-related views (for the process of re-finding of *Seken*-related views in Japanese minds in the modernized era as well as our previous researches about *Seken*, see Nakada, 2005 and 2008) as important ones and as something worthy of respect in their everyday life. We can easily see what *Seken*-related views mean for the Japanese respondents on the whole. It is without doubt that Japanese in the modern informatized era still live in the world filled with *Seken*-related meanings, views and values. (For the comparison, the following table includes the findings of our previous researches.) (Table 2 shows the findings on quasi-*Seken*-related views and anti-*Seken*-views for the purpose of comparison and statistical analysis.)

Table 1. Sympathy with *Seken*-related meanings in Japan

	1995 G	2000 G	2002 G	2003 G	2005 G	2006 G	2007 G	2008 G
Distance from nature	73.6%	-	82.6	79.0	82.2	78.2	-	79.8
Honest poverty	83.7	81.5	84.4	80.3	82.6	83.2	-	84.0
Destiny	84.4	79.0	77.9	76.0	80.8	81.6	-	81.2
Denial of natural science	88.5	88.3	90.7	88.7	89.8	89.6	-	86.2
Criticism of selfishness	85.5	88.3	90.0	90.3	84.8	84.2	-	90.2

Powerlessness	71.9	64.8	69.2	62.0	62.4	60.8	-	73.4
Superficial cheerfulness	73.3	65.6	70.8	62.7	60.4	60.4	-	71.0
Belief in kindness	-	68.1	73.1	71.5	73.8	76.4	84.3	77.2
Scourge of heaven	62.7	49.5	-	-	-	55.0	-	-
Warnings form heaven	-	-	-	-	-	-	63.1	67.4

1) Table 1 shows the percentages of the respondents who said “agree or somewhat agree” to Seken-related statements. These statements are such as: “Within our modern lifestyles, people have become too distant from nature”(Distance from nature); “People will become corrupt if they become too rich”(Honest poverty); “People have a certain destiny, no matter what form it takes”(Destiny); “In our world, there are a number of things that cannot be explained by science”(Denial of natural science); “There are too many people in developed countries (or Japan) today who are concerned only with themselves” (Criticism of selfishness); “In today’s world, people are helpless if they are (individually) themselves” (Powerlessness); “In today’s world, what seems cheerful and enjoyable is really only superficial” (Superficial cheerfulness); “Doing your best for other people is good for you” (Belief in kindness); “The frequent occurrence of natural disasters is due to scourge of heaven” (Scourge of heaven); “Occurrences of huge and disastrous natural disasters can be interpreted as warnings of heaven to people”(Warnings from heaven).

Table 2. Sympathy with (quasi) Seken-related meanings and anti-Seken views in Japan

	2007 G	2008 G
Decision depending upon one’s own views without relying on others’ opinions is the best way to do a good thing.	40.3%	39.6
To assert one’s demands and desires is very important for social life.	47.5	53.6
To keep public morals such as not to rump rubbish on the street even in cases of other peoples’ absence is very important for us human beings.	94.9	76.8
If it doesn’t put the companies or colleagues to much trouble, it isn’t too much bad to take equipments at the office such as notebooks or writing implements home and use personally.	22.9	10.2
To keep a good society, kind mutual help is required .	88.1	87.2

1) This table shows the percentages of the respondents who said “agree or somewhat agree” to each statement.

It might be quite strange at first glance (at least for the 'western' people and also for the Japanese who can see only *Shakai* and not *Seken*) to see that the majority of Japanese people of today have this kind of views in their minds, because everyone knows that Japan of today is a very modernized country. But as our researches in the past 10 years show, it is quite clear that Japanese people still live in an aspect of world /life-world which might be called *Seken*.

Originally, *Seken* consists of two different meanings, *Se* and *Ken*. *Se* means this world and *Ken* means 'between.' So *Seken* means the Between World, i.e. the world between heaven and the vulgar earth or the world between gods and people or the world between person and person.

In my (the author's) own view, this 'Ken' is characterized by the constant tensions between two contradictory forces: centrifugal force and centripetal force; or remoteness and closeness. In fact, Japanese people have two different attitudes towards nature: fear (feeling or mood combined with remoteness) and affection (feeling or mood combined with closeness). Our research findings show that this remark by myself is not just a gratuitous assumption. The figures of Table 1 show that Japanese feel remoteness and closeness towards nature/natural disasters (nature and natural disaster seem to lie near each other in Japanese minds) at the same time (these combined feelings consist of part of *Seken*). Remoteness is symbolized by 'scourge of heaven' or 'warnings from heaven' and closeness is symbolized by '(fear of) distance from nature.' Likewise, Japanese feel remoteness and closeness towards human relations too. In this case, closeness is symbolized by 'belief in kindness' and remoteness is symbolized by complicated orientation to privacy (as we will see later).

Japanese complicated attitudes towards 'scourge of heaven' seem to be visible from these views on two contradictory forces. Traditionally, natural disasters, in particular earthquakes, were considered to be sources of evil as well as hope by Japanese (see: Nakada, 1982). In this case 'hope' means a kind of 'disaster utopias' which were (are) thought to be realized by some sort of tremendous natural disasters. In this sense, natural disasters symbolize remoteness (fear for destruction) and closeness (hope for something that enables extra-ordinary and exiting events to happen) at the same time.

Even in the modernized era, Japanese attitudes towards natural disasters are not far from those in the previous eras. When a great earthquake hit Tokyo and the neighbour districts in 1923 and killed more than 100,000 people, a famous writer (Junichiro Tanizaki) wrote about some kind of disaster utopia in one of his essays. (He wrote that this earthquake might be able to give 'us' a chance to rebuild the complicated and half-feudalistic city Tokyo). Of course, most of

Japanese people felt tremendous fear and sad during this catastrophe which was 'interpreted' as 'scourge of heaven' by the majority of Japanese people.

In my view, Seken's power as a broad horizon in which different meanings and experiences 'come together' (at least partly) derives from these constant tensions between two contradictory forces: centrifugal force and centripetal force; or remoteness and closeness.

Findings about Blogs in the case of 2007 Research

Table 3 shows the results of our 2007G Research about 'experiences and feelings when being engaged in writing blogs.'

**Table 3. Degrees of agreement on various views
on writing (in) blogs (data: 2007G) (N=439)**

	Agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Disagree
1. (Blog for reduction of stress) Writing (in) my blog makes me feel less stressed.	3.2%	21.2	39.0	23.9	12.8
2. (Blog for reduction of conflict) Writing (in) my blog makes me forget my discontent or conflict.	4.6	30.3	33.5	19.4	12.3
3. (Blog for my true feel-ing) Writing (in) my blog enables me to know my true feeling.	5.5	3-.5	38.3	14.1	11.6
4. (Blog for reflection of my feeling) Writing (in) my blog enables me to reflect on my feeling and thinking.	9.8	42.4	29.6	8.9	9.3
5. (Blog for acquaintance with others) Writing (in) my blog enables me to become acquainted with others interested in me.	7.5	32.1	36.4	13.0	10.9
6. (Blog for exchange of opinion) Writing (in) my blog enables me to honestly and frankly exchange opinions with others.	5.2	28.5	41.5	13.7	11.2

7. (Blog for knowing others) Writing (in) my blog enables me to know others through their responses to my writing.	3.9	30.3	39.0	15.0	11.8
8. (Blog for impressing my own beings) Writing (in) my blog enables me to impress my own being to others.	3.6	26.0	40.8	16.4	13.2
9. (Blog for others' better understanding of me) Writing (in) my blog enables me to gain others' better understanding of my own thoughts.	3.0	24.6	42.6	17.5	12.3
10. (Blog for expressing my feeling) Writing (in) my blog enables me to express my feeling and thinking.	10.9	44.4	30.3	6.4	8.0
11. (Blog for writing with satisfaction) .I often feel that writing (in) my blog makes me satisfied.	12.1	35.5	41.2	6.2	5.0

1) The percentages show degrees of agreement on various views on writing (in) blogs. The respondents were asked to respond to 'Do you agree or disagree with the following statements as your own experiences or feelings when you are engaged in writing your own blogs?'

We need to transform the findings shown in Table 3 into a simpler form, because it is impossible for us to see what these findings mean as a whole. For the purpose of this transformation, we conducted a factor analysis on the findings of Table 3. Factor analysis is a method to reduce a set of items (in this case a set of respondent's responses to the views) into several inner-related groups (factors) of items. As a result, we could get two factors (Table 4).

These two factors mean that, in the case of our Japanese respondents, writing (in) blogs is mainly motivated by two different orientations: 'orientation to communication' and 'orientation to self-understanding.' So far as our research findings suggest, the respondents' engaging in (writing) blogs includes two different experiences. In my (the author's) own view, we can understand very easily why writing (in) blogs includes these two different experiences or orientations, when we consider that blogs are a kind of personal diaries (in which the writer express-

es his own thoughts, feelings as well as he tries to know himself) on the one hand and they are written for communication with others on the other hand.

Table 4. Factor Analysis (principal factor analysis, Varimax rotation) for ‘various views on writing (in) blogs’ (data:2007G)

Factors	Contributing Values and Factor Loading
Orientation to communication	Blog for knowing others (.824) Blog for others' better understanding of me (.785)
	Blog for exchange of opinion (.763) Blog for impressing my own beings (.763)
	Blog for acquaintance with others (.754) Blog for expressing my feeling (.558)
Orientation to self-understanding	Blog for reduction of conflict (.831) Blog for my true feeling (.807)
	Blog for reduction of stress (.719) Blog for reflection of my feeling (.684)

(‘Contributing value(s)’ mean that each factor consists of these items(views).)

One of the main topics in this paper is how and to what extent ‘writing (in) blogs’ is related with other experiences, meanings or ‘horizons.’ In our case we are particularly interested in the relations between blog-writing experiences and Seken-related views. In order to know these relations, we did a factor analysis on Seken-related views in a similar way as we did on ‘blog-writing experiences.’ But in this case, because our research data on Seken-related views are limited (see Table 1), in order to perform factor analysis, we add ‘views on individualism,’ ‘views on publicness’ to the set of items used for calculation. (These views are shown in Table 2.), (This means that the results of factor analysis depend on the set of items used for calculation, therefore factor analysis always provides us with just a tentative model for understanding), as a result, we could get two Seken-Individualism-Publicness-related factors shown in Table 5.

Table 5. Factor Analysis (principal factor analysis, Varimax rotation) for Seken-related views, ‘individualism’, ‘publicness’-related items

(data:2007G)(N=1200)

Factors	Contributing Values and Factor Loading
Orientation To Seken and mutual reliance	To keep a good society, kind mutual help is required (.784)
	Doing your best for other people is good for you. (.621)
	To keep public morals such as not to rump rubbish on the street even in cases of other peoples' absence is very important for us human beings.(.557)
	Occurrences of huge and disastrous natural disasters can be interpreted as warnings of heaven to people. (.480)

Orientation to individualism and self-benefit	<p>If it doesn't put the companies or colleagues to much trouble, it isn't too much bad to take equipments at the office such as notebooks or writing implements home and use personally.(.528)</p> <p>Decision depending upon one's own views without relying on others' opinions is the best way to do a good thing.(.499)</p> <p>To assert one's demands and desires is very important for social life.(.493)</p>
--	--

By using these two factors, we tried to relate two Seken-Individualism-Publicness-related factors with blog-writing experiences in order to know our central purpose in this paper: to know how different meanings or experiences on different horizons come together. Table 6 shows the results. It is clear that a variety of blog-writing experiences have a strong and fairly strong relations with 'orientation to Seken and mutual reliance' factor and 'orientation to individualism and self-benefit' factor as well. In my (the author's) interpretation, these findings might be able to be explained in the following way. 1) As we saw before, blog-writing activities fulfill two purposes: to express (or to know) one's own personal minds and to communicate with others through this expression (or understanding) of his personal inner minds. In this sense, it is not strange that blog-writing experiences (which consists of two aspects: orientation to communication and orientation to self-understanding) are found out to be related with Seken-Individualism-Publicness-related factors' two aspects. 2) What makes these relations be difficult to interpret is the fact that one of the Seken-Individualism-Publicness-related factor, 'orientation to Seken and mutual reliance,' includes a variety of views: 'To keep a good society, kind mutual help is required,' 'Doing your best for other people is good for you,' 'Occurrences of huge and disastrous natural disasters can be interpreted as warnings of heaven to people' and so on. In order to interpret this complicated inter-related meanings or views, we need to see the ethical implications that these Seken-Individualism-Publicness-related views appear to have. As we suggested before, we might be able to interpret that Seken is a horizon or a realm of meanings in which everything is ethically evaluated in some ways. Nature, human relations, views on society, common senses, destiny have a lot of common in that they are seen from ethical view-points within Seken. And one aspect of blog-writing experiences is likely to be included within Seken as an ethically pre-interpreted horizon. Probably, Seken is a horizon which provides people with ethical viewpoints as well as a broad interpretative context for relating things in this world to one's minds or experiences. In regard to the latter aspect, views, interests or experiences within Seken seem to be inter-related within it, as a kind of interpretative context, by the constant tensions: the tensions between two contradictory forces; centrifugal force and centripetal force; or remoteness and closeness.

Table 6. Relations between ‘various views on writing (in) blogs’ and ‘Orientation to Seken and mutual reliance factor’ as well as ‘Orientation to individualism and self-benefit factor’

(data:2007G)

	Orientation To Seken and mutual reliance	Orientation to individual- ism and self-benefit
1. Blog for reduction of stress	.090*	.202**
2. Blog for reduction of conflict	.137**	.190**
3. Blog for my true feeling	.231**	.172**
4. Blog for reflection of my feeling	.294**	.164**
5. Blog for acquaintance with others	.205**	ns
6. Blog for exchange of opinion	.164**	.211**
7. Blog for knowing others	.190**	.174**
8. Blog for impressing my own beings	.116*	.267**
9. Blog for others’ better understanding of me	.127**	.296**
10. Blog for expressing my feeling	.319**	.099*
11. Blog for writing with satisfaction	.249**	ns

1)**= $p < 0.01$, *= $p < 0.05$, ns= non (statistically) significant

Findings about Blogs, Seken, Privacy in the case of 2008 Research

The aim of our 2008G Research is the continuation of 2007G Research, although 2008G Research includes some additional questionnaires on privacy, roboethics, business ethics and so on. As we said before, if Seken is a horizon or a realm of meanings in which everything is ethically evaluated and also is located according to some cognitive map(s), and if we can see that such views as views on privacy, roboethics, business ethics can be found to merge into part of Seken, then we

might be able to know (at least partly) how some different horizons of meanings come together. According to Hideo Kobayashi, the Buddhist wisdom tells us that the world emerges in front of us as a way of the world's responding to our own views on life (Kobayashi,1954). It seems that, at least for Japanese, connections of things, experiences and meanings emerge in front of them as a way of response of these connections or Mu to the inner structures of Senken.

Table 7 shows the results of the respondents' views on 'what are the good blogs?' It is clear at the level of impression that the criteria of good blogs are related with two points: the possibility of sharing inner experiences among the writer and the readers; the writer's activity in regard to 'operating' his blog as a communication tool. This impression is partly endorsed by the factor analysis on the findings in Table 7. Table 8 shows us the contents of two factors gained through this factor analysis.

Table 7 Views on good blogs in Japan(data:2008G)(N=384)

	Agree	Somewhat agree	Somewhat disagree	Disagree
1. I'm interested in blogs which show some part of privacy or inner private experiences of bloggers (writers) to a certain extent.	14.6%	67.4	15.6	2.3
2. The blog where the blogger tries to enable the readers to share the blogger's personal experiences is a good one.	15.4	69.8	13.0	1.8
3. The blog where the blogger tries to express his personal experiences or thoughts in a modest way is a good one.	11.7	64.6	20.3	3.4
4. The blog where the blogger tries to show objective data is a good one.	10.9	58.1	27.6	3.4
5. The blog of which contents are frequently renewed is a good one.	20.0	62.5	15.1	2.1
6. The blog where the blogger regularly responds to the comments or trackback is a good one.	14.6	63.0	20.2	2.3
7. The blog where the blogger express his own thinking clearly is a good one.	15.9	65.1	17.4	1.6

1) The percentages show degrees of agreement on various views on good blogs. The respondents were asked to respond to 'What are your thoughts about various views on good blogs shown in the following list?'

Table 8. Factor Analysis (principal factor analysis, Varimax rotation) for 'evaluation of good blogs' (data:2008G)(N=384)

Factors	Contributing Values and Factor Loading
Inner communication as evaluation of good blogs	The blog where the blogger tries to enable the readers to share the blogger's personal experiences is a good one. (.796)
	The blog where the blogger tries to express his personal experiences or thoughts in a modest way is a good one. (.678)
	I'm interested in blogs which show some part of privacy or inner private experiences of bloggers (writers) to a certain extent. (.584)
Maintenance and constant response as evaluation of good blogs	The blog where the blogger regularly responds to the comments or trackback is a good one. (.742)
	The blog where the blogger express his own thinking clearly is a good one. (.619)
	The blog of which contents are frequently renewed is a good one. (.587)

In order to know the position(s) of blogs in Japanese minds or Japanese life-world, we tried to find the relations between 'views on (good) blogs' and other views/values, by using two factors of 'good blogs.' The following tables (including the tables showing the findings about 'views on privacy' and 'views on business ethics') show these results. As these tables show, blog's power as a horizon(if we regard blogs-related activity as something related with a kind of horizon) is rather limited when we compare these findings on blogs with those on Seken as a sort of horizon. (The table about correlations between 'Evaluation of good blogs-factors' and 'Interests in politics' is omitted here because of limited space. We could not find any statistically significant correlations between 'Evaluation of good blogs-factors' and 'interests in politics.')

**Table 9. Degrees of agreement on various views on privacy
in Japan (data: 2008G) (N=500)**

1. (Crime camera) Watching people through Crime-prevention-camera (security camera) in the streets is very important to keep public order. 55.4%
2. (Google) Collecting someone's personal information through search engine like Google without his being aware of it is a controversial issue because this information is used without permission. 64.4%
3. (Violation of privacy among friends) To ask someone about his income, occupation, family in detail might be regarded as violation of privacy even among personal friends. 52.4%
4. (Privacy and democracy) Respect for privacy is among the most important presuppositions to build a free and democratic society. 61.4%
5. (Poor consciousness about privacy) The fact that firms and companies in Japan are not so careful about protection of privacy or personal information shows the relatively lower level of Japanese consciousness about privacy. 59.2%
6. (Respect for collectivism) Too much respect for personal privacy might not harmonize well with virtues in collective life culture which emphasize shame, modesty, consideration for others. 39.6%
7. (Friendship over privacy) When we worry too much about privacy, we can't honestly and frankly talk about matters with our good friends. 48.2%
8. (Privacy protection by mass media) When the newspapers or TV reports on crimes, they should pay careful attention to privacy of the culprits or suspects in order not to violate it. 25.8%
9. (Need of information on privacy of victims) Detailed reports on victims of serious crimes like homicide including victims' occupations, human relations, life history or personality which are sometimes presented in the newspapers or on the TV are important and necessary in some cases in order to know the meaning of the incidents. 31.2%
10. (Need of information on privacy of culprits) Photos or real names of culprits of crimes under 20 years old which are sometimes presented illegally on the Internet might be important information in some cases, although these items of information are controversial. 45.2%
11. (Openness of personal information) To open part of my afflictions of illness or failure to my friends sometimes makes our distance of relations closer and better. 46.8%

(The percentages are added figures of 'strongly agree' and 'somewhat agree'.)

Table 10. Correlations between ‘Evaluation of good blogs-factors’ and ‘various views on privacy’*(data: 2008G) (N=384)*

	1. Crime camera	2. Google	3. Violation of privacy among friends	4. Privacy and democracy	5. Poor consciousness about privacy	6. Respect for collectivism
Inner communication as evaluation of good blogs	.158**	ns	ns	.159**	.107*	.149**
Maintenance and constant response as evaluation of good blogs	.113*	ns	ns	ns	.133**	.152**

1)**= $p < 0.01$, *= $p < 0.05$, ns= non (statistically) significant**Table 11. Correlations between ‘Evaluation of good blogs-factors’ and ‘various views on privacy’***(Data:2008G)(N=384)(continuation)*

	7. Friendship over privacy	8. Privacy protection by mass media	9. Need of information on privacy of victims	10. Need of information on privacy of culprits	11. Openness of personal information
Inner communication as evaluation of good blogs	.184**	ns	.116*	ns	.229**
Maintenance and constant response as evaluation of good blogs	ns	ns	ns	ns	ns

1)**= $p < 0.01$, *= $p < 0.05$, ns= non (statistically) significant

Table 12. Degrees of agreement on various views on business ethics in Japan (data: 2008G) (N=500)

1. (Consumers' bad attitudes) One of the reasons why Japanese firms or companies try to do unfair and anti-moral practices such as disguised foods-producing areas is due to consumers' attitudes that demand the firms and companies to sell products as cheap as possible. 46.4%
2. (Need for reduction of informal employment) Japanese firms or companies should try to stabilize employment by their effort to reduce informal employment and so on. 57.0%
3. (Doubt about principle of competition) Japanese firms or companies should be blamed if they try to follow American style of capitalism and to obey the principle of competition as well as free (unbridled) market mechanism without question. 40.0%
4. (Orientation to cooperation with others) The attitudes of Japanese firms or companies that lead them to obeying the principle of competition as well as free (unbridled) market mechanism without question don't harmonize with Japanese values which are associated with evaluation of modesty or cooperation with others. 37.6%
5. (Workplace for self-improvement) The firms or companies which just seek profits and don't provide employees with any occasions for character-formation or self-improvement can't get high social valuation in the long run. 72.0%
6. (Orientation to profits) The best thing that firms or companies can do for society is to gain as many profits as possible and to return the profits to society as taxes. 27.6%
7. (Contribution to communities) Japanese firms or companies should actively try to do contribution to the regional communities where they are located. 59.0%

(The percentages are added figures of 'strongly agree' and 'somewhat agree'.)

Table 13. Correlations between 'Evaluation of good blogs-factors' and 'various views on business ethics'

(data:2008G)(N=384)

	1. Consumers' bad attitudes	2. Need for reduction of informal employment	3. Doubt about principle of competition	4. Orientation to cooperation with others	5. Workplace for self-improvement	6. Orientation to profits	7. Contribution to communities
Inner communication as evaluation of good blogs	ns	ns	ns	ns	ns	ns	ns
Maintenance and constant response as evaluation of good blogs	ns	ns	ns	ns	ns	ns	.102

1)**= $p < 0.01$, *= $p < 0.05$, ns= non (statistically) significant

One the other hand, the power of Seken as a horizon of meanings/experiences seems to cover a wide range of Japanese life-world. Table 14 indicates 4 factors which we could gain through factor analysis on 'Seken-related views' in Table 1 and 'views on individualism' as well as 'views on publicness', just as in the case of factor analysis for 2007G data.

**Table 14. Factor Analysis (principal factor analysis, Varimax rotation)
for 'Seken-related views,' 'individualism,'
and 'views on publicness'**

(data: 2008G)(N=500)

Factors	Contributing Values and Factor Loading
Orientation to non-rationality and sense of destiny	To do one's job as a kind of destiny or divined duty is important for a good person, no matter what the results are. (.790)
	People have a certain destiny, no matter what form it takes (.717)
	In our world, there are a number of things that cannot be explained by science. (.613)
	People will become corrupt if they become too rich (.558)
Denial of mod- ern civilization	In today's world, people are helpless if they are (individually) themselves. (.641)
	In today's world, what seems cheerful and enjoyable is really only superficial (.611)
	There are too many people in developed countries (or Japan) today who are concerned only with themselves (.590)
	Because human power is limited, we can't prevent damages of natural disasters, no matter how we try to do. (.565)
Orientation to sincerity	Sincerity enables us to understand each other, no matter what kind of person the other is. (.752)
	Doing your best for other people is good for you (.547)
Orientation to individualism	Decision depending upon one's own views without relying on others' opinions is the best way to do a good thing. (.629)
	To assert one's demands and desires is very important for social life. (.480)

As the following tables show, these 'Seken-Individualism-Publicness-related factors' have strong or fairly strong relations with such views or interests in as views on privacy or interests in politics and so on, although we need an additional and complicated interpretation about 'what the correlations mean' in the case of 'orientation to individualism' factor (this factor seems to be in contrast to other factors). (The table about correlations between 'Seken-Individualism-Publicness-related factors' and 'Interests in politics' is omitted here because of limited space.

We could find strong correlations between ‘interests in politics’ and ‘Seken-Individualism-Publicness-related factors.’)

Table 15. Correlations between ‘Seken-Individualism-Publicness-related factors’ and ‘various views on business ethics’

(data:2008G)(N=500)

	1. Consumers' bad attitudes	2. Need for reduction of informal employment	3. Doubt about principle of competition	4. Orientation to cooperation with others	5. Workplace for self-improvement	6. Orientation to profits	7. Contribution to communities
Orientation to non-rationality and sense of destiny	.166**	.297**	.215**	.216**	.405**	ns	.370**
Denial of modern civilization	.159**	.294**	.303**	.311**	.331**	.104*	.318**
Orientation to sincerity	ns	ns	ns	ns	.130**	.147**	.211**
Orientation to individualism	ns	ns	ns	ns	-.165**	.167**	-.130**

1)**= $p < 0.01$, *= $p < 0.05$, ns= non (statistically) significant

Table 16. Correlations between ‘Seken-Individualism-Publicness-related factors’ and ‘various views on privacy’

(data:2008G)(N=500)

	1. Crime camera	2. Google	3. Violation of privacy among friends	4. Privacy and democracy	5. Poor consciousness about privacy	6. Respect for collectivism
Orientation to Non-rationality and sense of destiny	.262**	.220**	.183**	.221**	.256**	.159**
Denial of modern civilization	.294**	.217**	.207**	.320**	.346**	.136**
Orientation to sincerity	ns	.115*	ns	.100*	ns	.235**
Orientation to individualism	ns	-.111*	ns	ns	ns	ns

1)**= $p < 0.01$, *= $p < 0.05$, ns= non (statistically) significant

**Table 17. Correlations between ‘Seken-related-factors’
and ‘various views on privacy’**

(Data:2008G)(N=500)(continuation)

	7. Friend- ship over privacy	8. Privacy protection by mass media	9. Need of information on privacy of victims	10. Need of information on privacy of culprits	11. Open- ness of personal information
Orientation to Non-rationality and sense of destiny	.177**	ns	ns	.195**	.222**
Denial of modern civilization	.148**	-.107*	ns	.282**	.099*
Orientation to sincerity	.222**	.124**	.107*	ns	.289**
Orientation to individualism	ns	.104*	.113*	ns	ns

1)**= $p < 0.01$, *= $p < 0.05$, ns= non (statistically) significant

Conclusions for further discussions

One of the most important problems for us in this paper is about ‘how and to what extent different views or values come together’ or about ‘fusion of different horizons of meanings or experiences (when we are engaged in communication or self-presentation by using ICTs).’ As we saw in the tables in this paper, Seken seems to have such a strong power as to enable different views, variety of interests or ‘unexpected’ experiences to ‘come together’ within it. It is also clear that, through examination on the phenomena of this ‘coming together’ within Seken, some aspects of (meanings of) privacy and (activities of) blog-writing(or blog-reading) lie within Seken as a sort of horizon.

This means that we already know ‘to what extent’ different views, interests or experiences ‘come together’ within Seken. But the latter half of our problems, i.e. ‘how’ these different views, interests or experiences ‘come together’ within Seken still remains unclear in many respects at least at the present time. We’ve already seen some of the possible reasons of this ‘coming together’ within Seken: Seken provides people with a sort of cognitive(interpretative) map in which two competing forces(remoteness and closeness) are at work; Seken is a context in which everything is ethically valued in some ways. But these are just tentative schema for explanation and we have to see how these schema have validity through our examination in the future.

In addition to examination on this point, we have to see whether our discussions on *Seken* will be useful or not for people or scholars who are outside in the cultural tradition filled with *Seken*-related views.

According to Kurt Goldstein (Goldstein, 1934) and Merleau-Ponty (Merleau-Ponty, 1942 and 1945), the patient of aphasia and agnosia whom Goldstein and Merleau-Ponty analysed in their books is unable to do many things: to imitate soldiers' salute; to understand metaphors; to understand verbal orders and to obey the orders (for example, to touch some part of his body by obeying someone else's verbal commands); to take a walk without particular purpose; to classify things with abstract rules. And according to Japanese psychiatrist Takeshi Utsumi (who follows the tradition of S. Freud), mental patients of schizophrenia can't relate things (for example, his experiences) in two separate points: the present time and the past (Nakada and Capurro, 2009).

In these cases, things, experiences, meanings based on body, meanings on minds, the past memories, the present experiences can't 'come together.' In short, in these cases, the horizons or the cultural contexts are not at work. If we use Capurro's expression, some forces based on 'the world as a common place and as world-openness' are not at work in these cases (Capurro, 2006b).

At this point, it seems that discussions on *Seken* intersect with discussions on our existence in this world or as 'In-der-Welt-sein,' just as Capurro suggests by saying in one of this articles: Human existence is characterized by its 'being-outside' sharing implicitly or thematically with others the 'sense' and 'meaning' of things in changing contexts (Capurro, 2006b). As we can easily imagine, when we are engaged in writing or reading blogs, our engaging in these activities are not exceptions of this characterization. In my view, users of blogs, SNS as well as those facing the problem of 'how to find good balance between openness of oneself and concealment of privacy' experience this process of sharing implicitly or thematically with others the 'sense' and 'meaning' of things in changing contexts. And in these cases, people appear to be experiencing 'coming together of different levels of experiences,' partly as a typical phenomenon of our being-in-the-world-with-others and also as a newer phenomenon of a being-in-the-networked-world (See: Capurro 2005). In the case of our Japanese respondents, they seem to experience these two phases of 'coming together' in the information era in the same manner. But in this case, on the other hand, within *Seken*, things and meanings are likely to 'come together' in a different manner than scholars in 'western' culture(s) as well as Japanese scholars influenced by 'reductionistic views on information society' might expect. In fact, within *Seken*, 'warnings from heaven' come together with 'views on blogs' or 'views on privacy' in a somewhat 'mysterious' and at the same time 'ethical' way.

REFERENCES

- Blankenburg, Wolfgang** (1971). *Der Verlust der natuerlichen Selbstverstaendlichkeit: Ein Beitrag zur Psychopathologie symptomarmer Schizophrenien*, Stuttgart: Ferdinand Enke Verlag.
- Capurro, Rafael** (2006). Towards an ontological foundation of information ethics. *Ethics and Information Technology* (8), 175-186.
- Capurro, Rafael** (2005). Privacy: An Intercultural Perspective. *Ethics and Information Technology*, 7 (1), 37-47.
- Capurro, Rafael** (2003). Angeletics-A Message Theory. In Hans H. Diebner, Lehan Ramsay (eds.): *Hierarchies of Communication*. Karlsruhe: ZKM, Center for Art and Media, 58-71.
- Goldstein, Kurt** (1934). *Der Aufbau des Organismus*. Haag: Martinus Nijhoff.
- Idsutsu, Toshihiko** (1991). *Ishiki to honshitsu*, Tokyo: Iwanami.
- Kimura, Bin** (1975). *Bunretsubyou no genshogaku*, Tokyo: Kobundo.
- Kobayashi, Hideo** (1954). *Watashi no Jinseikan (My views on life)*, Tokyo: Kadokawasyoten.
- Merleau-Ponty, Maurice** (1945). *Phénoménologie de la perception*. Paris: Presses Universitaires de France.
- Merleau-Ponty, Maurice** (1942). *La structure du comportement*. Paris: Presses Universitaires de France.
- Nakada, Makoto** (2008). Privacy and Seken in China, Inner Mongolia and Japan: Interrelations between Privacy and other dominant views in China, Inner Mongolia (China) and Japan. In F. Sudweeks, H., Hrachovec, & C. Ess (Eds.), *Cultural Attitudes towards Technology and Communication 2008*. Perth: Murdoch University, 684-697.
- Nakada, Makoto** (2007a). Japanese traditional values behind political interests in the information age: Japanese political minds within Seken as old and indigenous world of meaning in Japan. In T. W. Bynum, S. Rogerson, & K. Murata (eds.), *ETHICOMP 2007 (Globalization: Bridging the Global Nature of Information and Communication Technology and the Local Nature of Human Beings)*. Tokyo: Meiji University, 427-435.
- Nakada, Makoto** (2007b). The Internet within Seken as Old and Indigenous World of Meaning in Japan. In Rafael Capurro, Johannes Frühbauer and Thomas Hausmanninger (eds.), *Localizing the Internet. Ethical Issues in Intercultural Perspective*. Munich: Fink Verlag, 177-203.

Nakada, Makoto (2005). Are the Meanings of Japanese Popular Songs Part of Seken-Jinseikan Meanings? *Gendaibunka-Koukyouseisaku-ronnsyu*, vol.1, 105-130.

Nakada, Makoto (1982). Saigai to Nihonjin (Disasters and Japanese). *Nennpou Shakai Sinnrigaku*(Annual Journal of Social Psychology)23:171-186.

Nakada, Makoto and Capurro, Rafael (2009). The Public/Private Debate: A Contribution to Intercultural Information Ethics. In Rocci Luppigini and Rebecca Adell (eds.), *Handbook of Research in Technoethic*, Hershey PA:IGI Global, 339-353.

Nakada, Makoto and Tamura, Takanori (2005). Japanese conceptions of privacy: An intercultural perspective, *Ethics and Information Technology*, 7(1), 27-36.

Nakada, Makoto, Tamura, Takanori, Tkach-Kawasaki, L., & Iitaka, T. (2004). Does Old Japan Determine New Japan? The Relationship between Seken, the Internet, and Political Consciousness in Japan. In F. Sudweeks, C. Ess (eds). *Fourth International conference on cultural attitudes towards technology and communication 2004*. Perth: Murdoch University, 143-157.

Nakaumura, Youjiro (2001). Nishida Kitaro I, Tokyo: Iwanami.

Utsumi, Takeshi (2005). *Tougoushityousyou no seishinryouhou kanousei ni suite* (About the possibility of psychiatric treatment for schizophrenia). *Seishin-Ryouhou*, 31(1), 9-18.

The counter-control revolution: silent control over individuals with dataveillance systems

Yohko Orito*

Department of Comprehensive Policy Making,
Faculty of Law and Letters,
Ehime University

Abstract

Although the use of personal information may produce a high level of customer satisfaction, it may also result in a situation in which organisations consciously or unconsciously decide what kinds of information individuals can receive. This is social sorting based on personal information collected and processed by dataveillance systems. Under these circumstances, individuals' freedom of access to information, as well as their freedom of speech and thought, could be constrained by the system architecture. Consequently, the power of individuals to control the transmission and flow of information caused by the control revolution has already shifted to business organisations. This study examines dataveillance systems that business organisations have developed, and proposes the concept of the counter-control revolution, which has been progressing almost unnoticed, based on observations of relevant business cases.

Keywords: Personal information, Dataveillance, Freedom of speech, Architecture, Consumer-generated Media (CGM)

Introduction

Based on the tremendous advance of information and communication technology (ICT), in particular Internet technology and its rapid penetration into societies, Shapiro (1999) proposed the concept of the control revolution; "The real change set in motion by the Internet may, in fact, be a control revolution, a vast transformation in who governs information, experience, and resources. Increasingly, it seems that we [not large institutions but individuals] will" (p. 10). In fact, as

* Yohko Orito is an assistant senior professor at Faculty of Law and Letters, Ehime University, Matsuyama, Japan. She received her PhD in commerce at Meiji University, Tokyo, Japan in 2007. Her research interest is in information ethics in business organisations, particularly use of personal information in businesses and protection of the right to privacy and freedom. She has served as a reviewer for *Journal of Information, Communication and Ethics in Society* and *Journal of Information and Management*.

Shapiro emphasised, it could be that the ability to collect and transmit information is greater than ever due to the Internet. One typical example is a search engine website. When individuals access a site such as Google, Yahoo, or Microsoft network (MSN) and enter the query keywords about which he/she wants to know into the search box on the site, it seems that they receive relevant information in a most efficient manner. Individuals can also easily share information about the goods and services that public and private organisations provide with others, and transmit his/her opinions or remarks about them on the Internet. Bulletin board systems (BBS), blogs and social networking services (SNS) are the consumer-generated media (CGM) enabling such information sharing.

In many cases, this sort of web-based service for individuals is based on collecting personal information to be used to provide customised or personalised services. While these personalised services can be of great benefit to the individuals, they require the construction of dataveillance systems that can be used by large organisations and businesses in particular to exert silent control over individuals. If this sort of service has a negative impact on individuals and involves an unacceptable degree of social risk, countermeasures must be considered as soon as possible.

Based on the recognition of this possibility, this study attempts to analyse the social risks caused by web-based services for individuals that are based on dataveillance systems. Any dataveillance system has benefits as well as harm for individuals; this study focuses on the possible harm caused by dataveillance systems. In particular, this study considers whether individuals can autonomously determine what kind of information they acquire in the dataveillance environments through examining the situation opposite to the control revolution, which can be characterised as the counter-control revolution, with a focus on the individual customers of business organisations (hereafter “customer” in this paper means “individual customer” if not otherwise specified).

The structure of this study is as follows. The next section describes three examples of customised services and dataveillance systems and analyses the nature of dataveillance systems, which continually collect personal information to provide customised services. After that, the existence of the social risk of the silent control by the architecture in the dataveillance environment is described followed by the section that proposes the concept of the counter-control revolution, which is the risky phenomenon the individuals face in the today’s information society.

The nature of dataveillance systems

In recent years, many business organisations have shifted the focus of their business strategies toward the provision of customised or personalised products and/or services to respond to diversified market needs. The importance of customised and personalised approaches to each customer's needs is recognised in variety of management innovation concepts such as customer relationship management (CRM). This type of service is based on the real-time surveillance of customers or the continual collection and use of customers' personal information using ICT-based systems, i.e., dataveillance systems. This approach takes it for granted that customers accept the business organisations' access to their relevant personal information in exchange for enhancement of their customer satisfaction.

Dataveillance systems are controversial in terms of protection of the right to information privacy. The importance of protecting personal information has recently been acknowledged, and relevant regulations for it have been enforced in many developed countries. Business organisations that construct dataveillance systems promote personal information management systems that meet these regulations. Even though business organisations may try to extract some profit from use of personal information, personal information management systems are aimed at improving customer satisfaction. The systems were originally based on good intentions on the part of the business organisations. However, regardless of the good intention, the dataveillance systems may have negative effects on societies. The following three cases are helpful in investigating the nature of dataveillance systems to consider the ramifications of the systems and the resulting dataveillance environment.

Search engine services

Individuals use search engine services by entering keywords into the search box to acquire information they need. Users consider this process as simply obtaining relevant information from the search engine sites. However, this relevancy is made possible by the provision of general information such as keywords in addition to the personal information of a massive number of users accumulated in the search engine companies' databases and manipulated by proprietary algorithm. Each user's online behaviour is always monitored automatically. It is alleged that the more information that users provide to search engine systems, the better the quality of information is that they acquire from the search engines (e.g., Battelle, 2005). It is also alleged that the quality of information the user can acquire is dependent on the behaviour of other users or the wisdom of crowds (Surowiecki, 2004).

Amazon.com

Amazon's web site provides users with a variety of information services including recommendations, book searches, and page customisation, all of which can enhance the users' satisfaction. These types of services are provided using the each user's own personal information as well as that of a large number of other Amazon users stored in Amazon's databases. For example, recommendations are based on data matching between personal information (e.g., past activity on the website) and databases that permit the analysis of user tendencies. Many users recognise the benefit gained by the dataveillance system that Amazon has developed and the information provided by individual customers is a key factor in the quality of Amazon's services. However, as in the previous case, Amazon's algorithm that determines what kind of information users can acquire is not disclosed, and it appears that Amazon's users do not care about how this algorithm works or the mechanism behind Amazon's recommendations to customers. In contrast to librarian code of ethics which require librarians to keep book records of library users in confidential and not to utilise the records, Amazon attempts to analyse users' records and tendencies, based on personal information, in order to enhance customer satisfaction.

Mobile marketing

In today's communication-oriented society, individuals can use mobile phones as information terminals to acquire useful information in a personalised fashion. In this system, the users' locations are determined from global positioning system (GPS) satellite signals and combined with personal information to provide useful personalised services. With the advent of multifunctional mobile phones (they are called "Keitai" in Japan), mobile marketing methods have become quite diverse, and many business organisations are trying to create mobile marketing systems to access the market in even more diverse ways. Mobile marketing is especially-pronounced in Japan where mobile market much further developed than those in other countries and Internet access through mobile phones are increasing. For example, personal information such as a user's dietary preference and pattern of activity is registered in mobile marketing databases, so that the system can transmit information about desirable restaurants nearby or even send restaurant coupons directly to the user's mobile phone. In this system, users can acquire useful information anywhere and at any time from the dataveillance systems, which are always monitoring users. The algorithm in this system is based on dataveillance systems that continually collect personal information through mobile terminals to track individual customers and determine their trends and behavioural patterns.

The nature of the dataveillance systems in the examples above can be summarised as follows. First, in dataveillance systems, a human being is a collection of data stored in databases of the systems. This data double in the databases is continually updated through real time surveillance. The personal information of the customer must remain accurate to maintain and increase customer satisfaction. Second, customer satisfaction with dataveillance systems can be characterised as standardised happiness. Here, so called “wisdom of crowds” is the decisive factor in determining what our happiness is, and the undisclosed algorithm used in the systems realises the provision of the happiness. Services provided by the dataveillance systems are statistically derived and, on a superficial level, it seems that there is no room for human intervention in the process of providing this standardised happiness.

Silent control by the architecture

As mentioned in the previous section, human beings are translated as a data double in dataveillance systems, and these databases are used for the evaluation of standardised happiness. Because the processes are automated and have no human intervention, the problem of invasion of privacy does not seem to occur in the systems. This raises questions as to whether the dataveillance environment poses an ethical issue at all; or, what the ethical and social issues concerning dataveillance systems should be recognised. This study proposes several new perspectives regarding the social risks posed by dataveillance systems.

The provision of personalised/customised service is possible only because of the personal data stored in databases. These database records are the key to what kind of information individuals can acquire and therefore to what they know, what they say, and how they can think and behave. Two forms of invisibility characterise dataveillance systems. The first is the invisibility of data doubles: we cannot know exactly what data about us are stored in databases or how they are collected and accumulated, even though the quality of these data may directly affect our identity in the modern society. Moreover, personal information databases are always necessarily incomplete reflections of the individuals they represent. Even though business organisations attempt to keep tabs on each customer's behaviour with real-time monitoring, the database cannot reflect the totality of real human beings, and we do not know which data are missing.

The second area of invisibility is the invisibility of the algorithm used in the dataveillance systems, which directly affect our quality of life. The personalised/customised approaches taken to individual customers by a business organisation are based on a certain algorithm and statistical analyses that are not understood by the customers. For example, although most of us know that Amazon has developed a profile or data doubles of us, nobody knows or even seems interested

in how this is used. Nevertheless dataveillance algorithm is very important in that they affect information processing and decision making about individuals.

These situations suggest that we are under silent control by the architecture of dataveillance systems. Lessig (1998) maintains that architecture is one of the four modalities of regulation of human behaviour in the modern information age, the other three being law, market, and norms (Figure 1). In fact, every activity of an individual as an Internet user operating in a dataveillance system is more or less affected by the web architecture.

The dual invisibility is such that users are not aware of the existence of the control exerted by the architecture. Such type of control is not easily recognised and, so, can be characterised as silent. In an environment with such silent control, we can recognise several controversial points that may cause serious social issues. The first of these is that, whereas the convenience provided by the dataveillance systems should make individuals feel that the “standardised happiness” is truly their own personal happiness, the statistical significance cannot necessarily guarantee the appropriateness of the services provided to us.

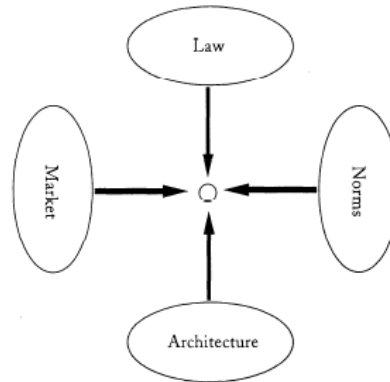


Figure. 1 Four modality model (Lessig, 1998).

The second controversial point concerns Moor's concept of the invisibility of computing, which includes the invisibility of computer abuse, the value judgments that IT professionals make, and the complexity of computation (Moor, 1985). It is doubtful that dataveillance systems are completely free from any human intervention and social risk. Moor explains the three invisibilities as follows. First, “invisible abuse is the intentional use of the invisible operations of a computer to engage in unethical conduct”. Second, “invisible programming values are those values which are embedded in a computer program ... In order to implement a program which satisfies the specifications, a programmer makes some value judgements about what is important and what is not”. Finally, “a third variety of the invisibility factor, which is perhaps the most disturbing, is invisible complex calculation ... Computers today perform, and certainly supercomputers in the future will perform, calculations which are too complex for human inspection and understanding”.

As Moor suggests, because computing may contain invisible impacts on us, a specific sense of values and functions may be buried somewhere in the architecture. When we make decisions based on the information provided by business organi-

sations without being conscious of the invisibility of computing, can our actions really be considered autonomous? Or, what guarantee is there that our Internet activities are really based on our own autonomous judgment? We do not have the ability to read all of the web pages extracted by the search engine algorithm due to cognitive limitations on our information processing. In addition, it is impossible for individuals to browse entire websites due to limits of time and economic constraints. Therefore, we cannot but depend on the effectiveness and other advantages of the web architecture. This raises the serious question; “are we autonomous beings in the dataveillance environment?”

The counter-control revolution

The social risk caused by the dataveillance systems is related with ethical issues concerning personal autonomy and freedom of speech. This sort of social risk can be seen as a type of social sorting. Lyon (2003) pointed out the risk of social sorting as follows: “For surveillance today sorts people into categories, assigning worth or risk, in ways that have real effects on their life chances. Deep discrimination occurs, thus making surveillance is not merely a matter of personal privacy but of social justice” (p.1). This suggests not only the risk of the invasion of information privacy but also the constraints on a range of information to which individuals have access and the opportunities they can exploit in society. These situations may affect an individual’s autonomous decision making and freedom of speech. In fact, in the current dataveillance environment, the more we use customised information services, the more we become dependent on them in our decision making processes. Individuals may not even consider whether the information they are provided is actually accurate. Even if the method of information sorting is not clear to the individual, some kind of unnoticed arbitrariness exists.

Whitaker (1999) called this new type of surveillance as a “Participatory Panopticon”. He states that “The strength of this new Panopticon is that people tend to participate voluntarily because they see positive benefits from Panopticon, and are less likely to perceive disadvantages or threat. They are not necessarily wrong to think this way, for the benefits are straightforward, real and tangible. Disadvantages are less tangible, more indirect and more complex. They ought not, however, to be ignored” (p. 140). In fact, because beneficial aspect of dataveillance is recognised by customers, it seems they voluntarily join in the dataveillance systems. However, the social risk caused by the new Panopticon may be more serious due to its invisibility for individuals.

Individuals in dataveillance systems are silently sorted based on their data doubles in various organisational databases, and the information they can access is determined by the black-box architecture. When social sorting occurs, individu-

als may be governed by the architecture and can be controlled by the values implicit in that architecture. Individuals will suffer from restrictions in the scope of information they can acquire and may lose sight of their identity, while their freedom of speech and thought seriously deteriorate. Such a situation appears as a counter-control revolution. While individuals appear to handle information on their own initiative, they are actually controlled by the architecture that large institutions and business organisations have developed based on dataveillance systems.

Similar arguments can be found in the study on a democratic system. In particular, Sustein (2001) puts free expressions on the Internet in the perspective of deliberative democracy. In his argument, "a well functioning of free expressions must meet two distinct requirements"; one is shared experiences among citizens, the other is that people should be exposed to materials that they would not have chosen in advance (pp. 8-9). On that basis, he suggests that as customisation or personalisation based on Internet technology advances, an opportunity to acquire shared experience is decreased and group polarisation is promoted among citizens and, consequently, deliberative democracy would be hard to come to function. His analysis implies that the dataveillance systems which have enabled customisation can become a risk factor to develop social consensus formed by enough deliberation among autonomous citizens. In this context, there exist individuals as citizens under the situation opposite to the "control -revolution".

The silent organisational control over individuals through the architecture is in progress, and it is not a matter of whether the organisations constructed the architecture with good or bad intentions. Instead, even good intentions may cause harm through a fallacy of composition; behaviour based on good intentions of each business organisation may collectively bring about a harmful outcome from the standpoint of society as a whole. Indeed, individuals speak freely online using websites and blogs. However, their online freedom of speech may be only an illusion. For example, who listens to our speech is determined in part by search engines. It is important to consider the situation where access to a particular website depends on the web architecture. If that is the case, then the architecture may make us the Cassandras of the information age; we can say anything we wish online, but nobody listens. Even if we convey useful information, like the prophecy of Cassandra, others may not be able to hear it on the Internet.

Conclusions

Whereas individuals have the ability to transmit information on the Internet, internet technology encourages the surveillance of and influence on individuals by large institutions, business organisations in particular. The power to control information may have already shifted to the large institutions that operate dataveillance systems. In this regard, the real situation would be opposite to the control revolution. Even as individuals become more able to control information instead of that ability's being available only to large institutions and organisations, information may, in fact, actually be controlled by those who develop dataveillance systems. However, this situation is not recognised by individuals.

Any organisation involved with dataveillance systems has a social responsibility for the systems it operates. Every individual has a social responsibility as a social being to build and maintain a safe and reliable society. Therefore, our common obligation as citizens is to join the process of consensus building and policy making aimed at the development and implementation of democratic information and communication technologies. As the first step towards the establishment of socially acceptable personal information management systems, we must evaluate the risks surrounding dataveillance systems. The concept of the counter-control revolution may be a good focal point for considering the social risks we face and the construction of a reliable society in the information age.

Acknowledgement

This study was supported by an open research centre project for private universities entitled "Quality-oriented human resource development and smart business collaboration: management quality science": a matching fund subsidy was provided by MEXT (the Ministry of Education, Culture, Sports, Science and Technology), 2007-2012.

REFERENCES

- Battelle, J.** (2005) *The Search: How Google and Its Rivals Rewrote the Rules of Business and Transformed Our Culture*. Portfolio, New York.
- Lessig, L.** (1998) The new Chicago school. *Journal of Legal Studies*, 27 (2), 661-691.
- Lyon, D. ed.,** (2003) *Surveillance as Social Sorting: Privacy, Risk, and Digital Discrimination*. Routledge, London.
- Moor, J. H.** (1985) What is computer ethics? *Metaphilosophy*, 16(4), 266-275.
- Shapiro, A. L.** (1999) *The Control Revolution: How the Internet Is Putting Individuals in Charge and Changing the World We Know*. Public Affairs, New York.

Surowiecki, J. (2004) *The Wisdom of Crowds: Why the Many Are Smarter than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations*. Doubleday, New York.

Sunstein, C. (2001) *Republic.com*. Princeton Univ. Dept of Art, New Jersey.

Whitaker, R. (1999) *The End of Privacy: How Total Surveillance Is Becoming a Reality*. New York Press, New York.

Being a 'Step Parent' to an 'Orphan' Work: Problems and Solutions Regarding the 'Adoption' of 'Orphans'

Maria Daphne Papadopoulos*

Dr., LL.M. mult, Hellenic Copyright Organization

Abstract

A new category of works creates a number of problems in the legal and 'real' world: the 'orphan' works. Those works of an unknown -or untraceable- author (the 'birth parent') create an uncertainty to the perspective users (or 'step parents'), who, although they need to use the works and they are willing to obtain a license (to 'adopt' them), they are unable to do so, since they do not know from where to ask it (the adoption procedure does not have a prescribed workflow and no expectation of predictable results exists). This paper will explore not only the controversial term of 'orphan' works but most importantly the controversial situation that 'orphan' works create. Part one will explain the notion of 'orphan' works and why and how they came up in the modern copyright world. Part two will explore the ethical aspects of the 'orphan' works issue according to the basic copyright values. We will examine their role to the safeguarding of fair balance of rights and interests between the different categories of right holders and users of protected subject matter. Part three will look through other kind of problems that originate from 'orphan' works, apart from the ethical and philosophical ones: legal, economic and technical problems. Finally, part four will look into the legal

* *Maria-Daphne Papadopoulos* graduated in 1996 from the Law School, at Aristotelian University of Thessaloniki, getting her law degree. From the same University she received her first Master's Degree in Commercial and Economic Law (1998). In the University of Houston she received her second Master in Laws in International & Intellectual Property Law with a scholarship from Onassis Public Benefit Foundation (2000). At the same University she worked as a Fellow Researcher in European Intellectual Property Law for a year. Following that, she was employed by several international law firms in Düsseldorf, Germany, while at the same time she was writing her dissertation "The Exercise of Moral Right in Copyright Law" (Aristotelian University of Thessaloniki) with a scholarship from Onassis Foundation (2002-2005). Since 2005 she has been working as a counselor-at-law at the Hellenic Copyright Organization, giving speeches at copyright conferences and writing articles in journals and collective works regarding copyright matters. Lately, she participated in the Legislative Committee for amending Greek Copyright Law.

solutions that countries across the world have already applied -or plan to apply- to confront this problem and will propose which solution could be the best one taking into consideration all the different parameters of the issue called 'orphan' works.

Keywords: 'orphan' works, copyright, limited liability, extended collective licenses, formalities, right holders, users, public domain

Introduction

A new category of works creates a number of problems in the legal and 'real' world: the 'orphan' works. Those works of an unknown -or untraceable- author (the 'birth parent') create an uncertainty to the perspective users (or 'step parents'), who, although they need to use the works and they are willing to obtain a license (to 'adopt' them), they are unable to do so, since they do not know from where to ask it (the adoption procedure does not have a prescribed workflow and no expectation of predictable results exists). If it is impossible to locate the right-holder, it is impossible to ask for permission in order to undertake any act that belongs to the exclusive rights of the author or the rightholder. That means that any adopting activity from the potential user (step parent) regarding an 'orphan' work would be illegal and would violate the rights of the author.

This violation results from copyright law, that confers exclusive rights to the authors or the owners of a copyrighted work to control the reproduction, the communication and the in other ways exploitation of their work. A perspective user of a work that is protected by the exclusive rights conferred by copyright must obtain a license from the copyright owner subject to a payment of a fee prior to using the work, or the use will constitute infringement.

In certain limited situations, called exceptions or limitations, copyright law permits a user to use a copyrighted work without the rightholder's consent. Most of the copyright laws provide certain exceptions, but the application of these exceptions depends on a number of prerequisites that have to be fulfilled. (In European community law the Information Society Directive, Directive 2001/29 of the European Parliament and the Council of 22 May 2001 on the Harmonization of certain aspects of copyright and related rights in the information society, OJ L. 167/10 of 22.6.2001, had the ambition to harmonize copyright limitations across European Union -only an ambition though and not a reality since only one of the exhaustive list of 21 limitations is mandatory, all the others being optional. See in this regard, Hugenholtz, 2000; Hugenholtz, et al., 2006, p. 64). The use of an 'orphan' work without the consent of his author would mean the violation of his economic and moral rights, if the user could not assert the expiration of the

protection term or the fact that the concrete use is allowed since –and under the condition that- it is within the limits of an exception.

From a moral perspective the ‘orphan’ works problem is like shopping at a store when the cashier left his post. You want to pay for your goods but no one is there to accept your money (Colleran, December 2007-January 2008).

‘Orphan’ works could be used in different ways and thus affect different kind of perspective users: they could be used by individual subsequent creators (e.g. by an author, who wants to dramatize a novel and he cannot find the author), they could be used for private uses (e.g. by a person who wants to digitize an old family photo and he cannot locate the photographer) and for large scale access projects (e.g. for digitization projects, such as Europeana and Google Book Bottis, 2007). The last category indicates institutional users (such as libraries, archives or museums) aiming to make a large quantity of works available to the public by digitizing and posting their collections on line (Report on Orphan Works, 2006, p. 37).

‘Orphan’ works: what are they and how did they come up?

Definition

Before continuing with this recently mostly disputed subject, it is considered necessary to fix the notion and identify which work qualifies as an ‘orphan’. The term ‘orphan’ work is used to describe the situation where the owner of a work still under copyright protection, after a reasonable amount of search, cannot be identified or located by the one, who wishes to make use of the work in a manner that requires permission of the rightholder.

This definition of ‘orphan’ works contains some paradoxes. The first one is that contains a maybe unintentional shift of perspective: works considered as ‘orphans’ solely from the perspective of the prospective user. This is inconsistent with the basic principle of copyright, which considers works and rights in those works solely from the perspective of the rightholder (Petit, 2006, p. 2).

The second paradox lies in the fact that, since at least some of the ‘orphan’ works’ authors are unknown, identifying the life span of the author in order to establish the term of copyright seems difficult, if not impossible. Without this vital identification information on the work, potential users must rely on circumstantial or contextual information -if any- to ascertain relevant factors in deciding whether to exploit the work. If no circumstantial or contextual information exist, the user generally must assume that a work he wants to use is copyright protected and not to assume that it has fallen into the public domain, unless there are elements that

prove the opposite (Report on Orphan Works, 2006, p. 4). Consequently there is a possibility the work to remain in perpetuity under the protection of copyright.

In this case it will be preferable to apply the same rule as with anonymous works -given also their similarity to 'orphan' works- and calculate the term of protection for 'orphans', whose author is unknown from the date that the work has been made available to the public (Art. 7 (3) Berne Convention and Art. 1(3) Copyright Term Directive - Former Council Directive 93/98/EEC (OJ L 290, 24.11.1993, p. 9) codified by Directive 2006/116/EC of the European Parliament and of the Council of 12.12.2006 on the term of protection of copyright and certain related rights – codified version). In this respect relevant is also Art. 44 of the Finnish Copyright Law (404/1961) which provides that in the case where the rightholder is unknown, the term of protection runs for seventy years from the date of creation. The advantage of this provision is that provides for a clear expiry date that it does not depend neither on the identity of the author nor to his death date.

In the same article of the Berne Convention (Art. 7 (3)) it is also provided that no member state of Berne Convention is required to protect anonymous or pseudonymous works in respect of which is reasonable to presume that their author has been dead for fifty years. Thus, even if the special rule for the term of protection of anonymous works will not be applied for 'orphan' ones, in the case that is presumed that an author of an 'orphan' work has been dead over the term protection, no copyright protection should be applied.

In order to have a complete picture of the issue some other elements should also be mentioned. Although only 'works' are named in the definition, the subject matter protected by related rights should be also covered by the term 'orphan' works and should be treated with the same manner (Memorandum of Understanding on Diligent Search Guidelines for Orphan Works, Joint Report, 2008, p. 3).

After having defined what is or what could be considered as an 'orphan' work, it should also be clarified what could be not.

The previously mentioned anonymous or pseudonymous works are not automatically considered as 'orphans' (Memorandum of Understanding on Diligent Search Guidelines for Orphan Works, Joint Report, 2008, p. 3). There are special provisions in international treaties and community law regulating those (Art. 7 (3) Berne Convention and Art. 1(3) Copyright Term Directive).

Within the notion of 'orphan' works should not be included also works whose right holders are known but they are either unwilling to give their consent or they do not reply to the enquiries from the potential users, or cases where the

potential user is unable to afford the royalty demanded by the copyright owner (Copyright and Orphan Works, A paper prepared for the Gowers Review by the British Screen Advisory Council, 2006, p. 4). Although those situations are also problematic for the perspective users, this fact is not enough to change the cornerstone of the copyright system, i.e. the exclusive power that the author has to permit or not the use of his work.

Also falling outside the scope of 'orphan' works are the out-of-print works, works whose their right holders are known but they are no longer commercially available (Final Report on Digital Preservation, Orphan Works, and Out-of-Print Works 2008, p. 17; Commission Staff Working Document regarding the progress on the digitization and online accessibility of cultural material and digital preservation across the EU, 2008, p. 16).

In addition, it should be clarified that 'orphan' works are different from 'abandoned' works. 'Abandoned' works or 'abandonware' refer to computer software that is no longer sold or supported, or whose copyright ownership maybe unclear for various reasons (more specifically, abandoned is defined as any PC or console game that is at least four years old and not being sold or supported by the company that produced it or by any other company, Abandonware Ring FAQ, 2002). The key difference between 'abandonware' and 'orphan' works is that in the first category the problem is the non availability of a copyrighted work, while in the second is the non locatability of the copyright owner. 'Abandonware' and 'orphanhood' exclude each other (Khong, 2007).

In the definition is also mentioned that a reasonable amount of search has to be made by the potential user in order to locate the author or the copyright owner. The search has to be 'diligent' and it should be proven. A work cannot be characterised as 'orphan', unless the potential user has concluded a reasonably diligent search in order to locate the author or the copyright owner. A number of problems arise in this regard, since no standard can be created to cover all cases. Each case has to be dealt differently, taken into consideration all the different circumstances applicable each time. On the other side though any regulatory initiative should refrain from prescribing minimum search steps or concrete information sources to be consulted, due to the rapidly changing information sources and search techniques (Final Report on Digital Preservation, Orphan Works and Out-of-Print Works, 2008, p. 15).

The diligent search criteria play the same role for the use of 'orphans', as the liability to the externality problem in torts law: the authorities set a due care standard and if the parties meet the standard, they escape liability. If they do not, they have to pay the costs of the accident. In 'orphan' works the solution is similar: some 'diligent search' criteria are established, set the level of the effort and

provide the proof that the search was carried out in good faith (Varian, 2006, p. 971).

Causes of the 'orphan' works problem

While there is no single direct cause of the 'orphan' works problem, it is often described as an unintended consequence of the major developments in copyright law during the twentieth century (Brito & Dooling, 2005, p. 82; Sherman, 2007, p. 13).

Based to the absence of copyright formalities, such as registration and renewal, and to the long duration of protection, a copyright system has been created that protects creative works for an extended period of time but often makes it difficult to identify and locate the copyright owner. In addition, creative works may also become 'orphans' because of everyday events such as the death of a copyright owner, industry imposed barriers to copyright use, and the reorganization, bankruptcy, or sale of a corporate copyright owner. Further, while technology has made it easier for the private actors to create and disseminate new works, it has also caused many works to become 'orphans' (Sherman, 2007, p. 13).

Following each one of those underlying causes is analysed.

Lack of formalities

One of the major causes for the 'orphan' works problem is the lack of formalities. Registration, the affixation of a copyright notice and renewal are prohibited by the Berne Convention (Art. 5 (2)). The prohibition of copyright formalities is incorporated also by reference in TRIPs Agreement (Art. 9 (1)) and in WCT (Art. 4 (1)). Copyrighted works are protected the moment they are fixed in a tangible medium of expression and they do not need to be registered with a Copyright Office (it has to be mentioned though that the first copyright law, the Statute of Anne, required registration of a work in order copyright to be obtained. It was a prerequisite to register a work at the Stationers' Hall before publication in order to claim damages from an infringer, 8 Anne, c. 19, s. 2). The abolishment of registration and the automatic subsistence of copyright came with the Berne Convention. Instead of having an 'opt in' copyright protection system by registration and payment of a fee, as it is the case in the other forms of intellectual property, we have theoretically an 'opt out' system. Theoretically because 'opting out' is only one recent possibility through 'open access' and 'public licenses'. However even then, it is not possible for a copyrighted work to be absolutely in the public domain (Khong, 2007).

Consequently, mandatory registration schemes are illicit under Berne Convention. National Copyright Registries do exist in some countries but their aim is

purely to promote evidence and the registration is optional (USA had relaxed the formalities with the Enactment of the 1976 Copyright Act and the 1988 Berne Convention Implementation Act and it is no longer necessary to include a notice of copyright on publicly distributed copies, nor to register the copies with Copyright Office. The registration of works though still plays a role, since it is a prerequisite to suit, if the work is of US origin, and some remedies for all works depend on that registration). Many argue that this absence -or the relaxation, in the case of US- of formalities is a significant cause of the 'orphan' works issue but this is not unquestionable. The registration and the renewal cannot offer constantly updated information identifying the author or the copyright owner. This kind of information could only be given if there would be a recordation of transfers, but this never was-and probably never will be-the case (Ginsburg, 2008).

Besides, this is the reason why the 'orphan' works issue is restricted to copyright law. Patent holders can be identified through a simple patent search and trademarks are a badge for origin; therefore it is not a problem to identify the commercial entity that owns it. Also the renewal possibilities that exist for both trademark and patent law exclude problems caused from the passage of time.

Not available identification information

In order for a work to become 'orphan' some obstacles create problems in identifying the copyright owner. The search for the rightholder of copyright starts with the information available on the work itself (Presumptions - Berne Convention, Art. 15 (1) and Directive 2004/48/EC of the European Parliament and of the Council of 29 April 2004 on the enforcement of intellectual property rights, OJ L 195, 02/06/2004, p. 16, Art. 5(a)). In most of the orphan works there are no or not adequate information on the work itself regarding the author or the copyright owner, no title, or no indicia of ownership on a particular copy of the work. Without these identification elements, potential users must rely on circumstantial or contextual information to discover relevant factors and decide whether to exploit the work or not (Report on Orphan Works, 2006, p. 23).

This is particularly true -the not available identification information- regarding photographs and audiovisual works. In photographs not only copyright information most of the times are missing but also third parties can digitise the works without attaching the necessary identifying information. Or even they can remove the relevant information from the digitised works and redistribute them without attribution or altered (Ginsburg, 2008).

Transferability, divisibility, inheritability of copyright

Even if information of the author or the copyright owner on a copy of the work itself does exist, a number of events could occur since the creation of the work and affect the ability of a future user to locate or identify the author or the copyright owner.

Since copyright can be transferred as any other ownership, the chain of title to copyright could cause problems in locating the copyright owner (mergers between companies, acquisitions of the assets of one company, e.g. Universal owns most of the library of the Paramount company, but the ownership of some films is disputed, or publishing houses may possess a number of works for which the copyright holder is undeterminable, Report on Orphan Works, 2006, p. 27). But even if the copyright has not changed hands, the circumstances of the owner could have been changed due to the long copyright term (change of address, dissolution, death of the authors or the right holders and transfer of copyright by will, fractional distribution of copyright among heirs, or, if the copyright owner is a business entity, ceasing operations, bankruptcy e.tc.).

Problematic could also be the case where the transfer refers only to some particular ways of exploitation of the work. This rights dichotomy would cause more problems to the potential user, since his task is not only to find the copyright owner but to find the right copyright owner for the particular right he wishes to exploit.

In different work categories the 'orphan' works' problem presents different intensity. Chains of title in published printed works commercially available or musical compositions are usually more reliable than other sectors, such as photographs (Ginsburg, 2008).

Long duration

Another underlying cause for the 'orphan' works problem is the term of copyright protection that extends the life of author plus seventy years. The Copyright Term Directive set the copyright term for all Member States at seventy years after the death of the author (Art. 1(1)), that is 20 years more than the minimum standard in Berne Convention (Art. 7(1)) (currently is discussed within the EU the amendment of the Copyright Term Directive by expanding the term of protection regarding the related rights of performing artists and producers from fifty to seventy years). The expanded length of the copyright term is a parameter for increasing the number of unlocatable authors. Another contributing factor is that unlike other intellectual property rights, such as patents or trademarks, the lack of exploitation of the work does not affect the copyright protection.

Digital technology

Another factor that has contributed to the rise of 'orphan' works is the expansion of digital technology. Using digital technology authors are able to easily create innumerable written, artistic, musical, visual and other works and post them on the World Wide Web, where they are readily accessible to the public. As a result, digital technology has not only increased the creation of works but also it has significantly reduced the dissemination cost, since potential users may enjoy and distribute on line the creative works of others. On the other hand though, more disturbingly, these users may use software or other technologies to alter or remove the attribution from creative works, destroying the identification information. While this freedom to create, disseminate, and modify information would seem to serve the goals of copyright law by benefiting the public, the over abundance of works that are accessible on the internet has made it even more difficult to identify authors of creative works. Consequently, while technology has facilitated immense creative output, it has also magnified the scope of the problem by creating an abundance of 'orphan' works that impedes the creative potential new technologies could spring (Sherman, 2007, p. 16; Thompson, 2006, p. 824). Taking a gloomy view of the issue, the number of 'orphan' works will only increase.

Is it ethical to adopt an 'orphan' work?

The 'orphan' works are problematic due to the uncertainty over their copyright status that prevents them from being used in new creative works or from being made available to the public, causing in this way certain pathology within the copyright system. It is considered necessary in order to find the most effective treatment of the 'orphan' works' issue to mention briefly the aims of copyright and its philosophical and theoretical foundation in both common and civil law system and to examine how 'orphan' works are placed within the copyright world.

In common law countries the primary objective of copyright is to promote the social good and to advance the general welfare through the provision of economic incentives to creators of new intellectual works (on the common law side two different systems must be distinguished; the British and the US System, Davies, 1995, p. 965; Sterling, 1998, p. 302). Thus the emphasis in common law system is put on the protection of the work.

On the contrary, in civil law countries copyright's (the term 'copyright' refers also to author's rights in the European sense) principal objective is centred on the person of the author. In civil law jurisdictions the protection of literary and artistic property is justified primarily in terms of authors' inherent entitlements –indeed as an extension of their personalities.

From the philosophical point of view those two different generic approaches to artistic and literary property share also a different historical and philosophical background and are based to different theories: the deontological and consequential theories (the relevant terminology presents a prodigious variety).

According to consequential theory copyright protection is necessary due to the valuable consequences it provides to a society, such as providing incentives, or encouraging creation and knowledge. Copyright is a necessary legal instrument in order to protect and advance the interests of all involved stakeholders. Therefore, not only authors are protected but also performers, producers, corporations and any other interest holders (Dutfield & Suthersanen, 2008, p. 51). Based on this theory, the Copyright Clause of the US Constitution (Article I, Section 8, Clause 8) extends copyright protection as a means to an end: the promotion of useful arts and science. The Copyright Clause authorizes Congress to grant exclusive rights to the authors in order to provide them with an economic incentive to create and disseminate works. Private protections are granted to authors in order to inspire the creation of works that will benefit the public.

Quite surprisingly elements of this approach can also be identified in the European community law regarding copyright. Directive 2001/29 is justified on the basis that copyright laws “protect and stimulate the development and marketing of new products and services and the creation and exploitation of their creative content” (Recital 2, Directive 2001/29). (The two generic approaches to artistic and literary property share more in the way of intellectual history and basic philosophy than is commonly acknowledged and today the laws of common law and civil law countries are as a doctrinal matter converging to a certain decree due to the developments in the law of international copyright. The objective of both copyright systems seems to be ultimately the same: the protection of authors’ rights over their literary and artistic works, as the Berne Convention has clearly expressed it.)

On the other hand the deontological approach asserts that rights are justified for persons entitled to authors’ rights as a matter of natural rights or as a matter for human rights or as a matter of duty. Moral rights constitute a classical example of deontological theory, recognizing e.g. the right to respect the creator’s name (Dutfield & Suthersanen, 2008, p. 51).

Using another terminology instead of deontological and consequential theory, we could refer to ‘marketplace norms’ that require rules to preserve a reliable market in works of intellect and ‘authorships norms’ that dictate rules enabling authors to control the use of their expressions by others. Anglo-American copyright laws rely on ‘marketplace norms’, while European laws respect ‘authorship norms’. An overriding norm would be one that would encompass and enhance both ‘market-

place and authorship norms', i.e. that would enhance the informational wealth and the expansion of commons, and at the same time the authors' significant self expression (Geller, 1994).

The most commonly cited theory supporting property rights in intellectual property law is the Locke's theory. According to this theory, all resources given by God are part of the 'commons'. Nonetheless every person has the right to use (or according to Locke's language "to expend labour on") these 'common' resources. One who mixes individual labour with what nature has provided acquires property in what is produced (Locke, 1986, p. 19). This theory supports the grant of strong intellectual property rights to authors, as those who use something that is common to all authors and produce something new. The private property though resulting from natural law is not immune to limitations, as many believe. The same theory provides some moral and ethical limitations on authorial property rights over literary and artistic works. Since authors' creativity actually depends on the creative deposit existing in the society, from which the authors consciously or unconsciously draw part of their inspiration and transform it, they have a moral and ethical obligation to ensure that other authors will have the same freedom and access to their own works to perpetuate this creative authorship. Similarly another proviso of Locke's theory is that these 'common resources' should be either used or returned to 'commons' for others to exploit them (Locke, 1986, p. 25). Locke believes that all men have a moral obligation towards society to use 'commons' in a useful and enriching way and not to waste them because in this situation they will deprive others from their equal share to 'commons'. Therefore they ought to use their works in a beneficial manner or they would lose property rights in their works of intellect (Ng, 2008). This statement could be proved to be extremely important to the treatment of 'orphan' works: law has to find its way through in order for society to be entitled to use those works.

The two balancing factors in copyright are present still from Locke's arguments: the need to reward the author but at the same time to maintain the 'commons'.

This area of 'commons' is the nowadays called area of 'public domain'. Do 'orphan' works constitute a part of the 'public domain' and if not should they be and under what conditions? The answer is not easy firstly, because a positive definition of the public domain does not exist and secondly because its composition is matter of a political choice depending on the economic and cultural interests of each government. Public domain comprises all intellectual assets that are not protected by copyright either because they were never subject to copyright protection (as ideas, methods, non original works, facts, laws, scientific and mathematical principles) or because the term copyright has expired. In accordance with the abovementioned traditional view of public domain exceptions are clearly not

included and so are 'orphan' works. According though to a new approach to the notion of public domain, its definition should not focus on the lack of protection but rather it should encompass all freely available resources for intellectual production, such as fair use or limitations and exceptions. Since the public domain is a policy construct intended to foster the development of productive practices, whether cultural, creative or consumptive, it should include not only elements in which such rights are non-existent, but also resources that are left untouched by the exercise of those rights (Cohen, 2006, p. 138). The fact that the work is still protected and its use is possible after the fulfilment of certain conditions should play no significant role for its placement in the public domain. From a sociological point of view, the public domain should be a field where the public can enter without infringing the intellectual rights of anyone. From an economical point of view, the public domain should cover the use of such intellectual assets for which no transaction could take place (Benabou & Dussolier, 2007, p. 172).

'Orphan' works are such intellectual assets for which no transaction can take place, for there are no known or locatable right owners with whom the perspective user could conclude a contract. The problem with the 'orphan' works lies exactly in this impossibility to use the works, since the right holders of the works are unknown. Instead of wasting those works and having them remaining useless, they could form a part of public domain. Their situation though is peculiar; placing 'orphans' in public domain on the one hand presupposes certain conditions (e.g. prior diligent search) (in the same way limitations do, since they could be characterized as public domain only if the necessary conditions are fulfilled) and on the other hand the characterization could be a temporary one. The works stop to be 'orphan' and consequently stop belonging to the public domain from the moment that the rightholder resurfaces.

The presence of a robust public domain is a pivotal element of the common heritage of humanity and it is essential for cultural, social and economic development. In order to maintain a strong and vivid public domain, we have to adapt it to the present circumstances. It is a long and perplexed discussion whether copyright is a public domain containing discrete islands of propertization or instead a realm of propertization, in which there are some holes, some delineated areas of commons (Radin, 2006, p. 981). Whatever the situation is, it is unquestionable that intellectual property is making a head over public domain; the extension of copyright duration and the creation of new rights (e.g. databases) are only some signs. The ultimate goal is to maintain the right proportion between propertization areas and public domain and since the propertization areas are expanding, we have to discover ways to enrich public domain. "The most important cost of the public domain comes from those new derivative works that are not created because of the new

author's inability to negotiate permission from whoever owns the copyright" (Brito & Dooling, 2005, p. 85).

The placing of 'orphan' works in the public domain is not without problems. Some of the proposed or existing solutions to the 'orphan' works' problem, that will be analyzed in the last part, presuppose a compensation mechanism for the right holder (general or conditional to the rightholder's reappearance). In that case a public domain with remuneration is formed. But even this notion is not unknown in the copyright system; the 'domaine public payant' doctrine depicts exactly this situation where a work falls within the public domain but its use is not necessarily free of charge. This remuneration paid is not a reward to the copyright owners but rather is dedicated to social and cultural policy aims and constitutes a direct or indirect benefit for the living authors. This doctrine is identified within the community law in the Copyright Term Directive (Directive 2006/116/EC PJ L 372 of 27.12.2006, p. 12-18) regarding the protection of previously unpublished works (Art. 4) ("Any person who, after the expiry of copyright protection, for the first time lawfully publishes or lawfully communicates to the public a previously unpublished work shall benefit from a protection equivalent to the economic rights of the author. The term of protection of such rights shall be 25 years from the time when the work was first lawfully published or lawfully communicated to the public").

The 'orphan' works cannot carry out the duties imposed by the copyright system. From deontological approach 'orphan' works do not fulfill their moral obligation towards society to use 'commons' in a useful and enriching way and they are wasted, for no one could use them in a manner that needs the prior authorisation by the rightholder. In addition the author of the work cannot enjoy his recognition and the profits from the work's exploitation. According to consequentialistic theory 'orphan' works do not provide any valuable consequences to the society; a productive and a beneficial use of the work is forestalled, not because the copyright owner has asserted his exclusive rights in his work or because an agreement between the owner and the user could not be achieved regarding terms of the license but merely because the copyright owner could not be located by the perspective user (Report on Orphan Works, 2006, p. 15). Even applying the overriding norm it becomes evident that the existence of 'orphan' works neither boosts the informational wealth nor enhances the self expression of the authors.

Even if we treat the issue from the different stakeholders' perspective, the conclusion would not be different. Another means of viewing copyright is to utilize stakeholder analysis. Regulation of copyright has to create a balance between the competing interests of the key stakeholders: the authors, the producers and the users (including future creators). 'Orphans' can not cover the needs of any stake-

holders. The authors' needs include recognition, respect and remuneration and the users' needs call for access and affordability of cultural heritage and technology. The authors of the 'orphan' works in the best case they could enjoy a kind of recognition and respect, if their names are known (in that case of course whatever serves as solution to the 'orphan' works' problem should safeguard at the same time the protection of the attribution right, e.g. the attribution requirement in Orphan Works Act 2008, Section 2) but they -or the copyright owners- cannot be rewarded for their use. The users on the other side cannot use the respective works without risking a law suit for copyright infringement. Even the needs of the producers fail to be satisfied, since they cannot exploit the work without prior permission of the copyright owners.

It becomes apparent that, no matter from what perspective one examines the 'orphan' works' issue and what theory is applied, the result is the same. The copyright owners do have an unquestionable bundle of exclusive rights in their works, even if they are unknown or if they are impossible to locate them. The reaching though of a dead end after having applied the existing rules of copyright system forces us to consider as indispensable a legislative/normative intervention that would allow the use of 'orphan' works under certain preconditions. A difficult balancing act must be undertaken. What is at stake is to assure legal certainty for the exploitation of 'orphan' works and to promote creativity without dissolving copyright.

By this treatment of 'orphan' works the goal of copyright law to reach the 'golden rule' (ethical code which states that one has a right to just treatment, and a responsibility to ensure justice for others), i.e. to strike a balance between the interests of authors and the society's competing interest to have access to the works of intellect, could be achieved.

Before detecting possible solutions to the orphan works' issue, we will analyze the problems that orphan works induce. Those problems urge for an effective treatment of the 'orphans'.

Problems caused by 'orphan' works

The high transaction costs associated with using 'orphans' result to additional economic and cultural costs upon the right holders and the public, since they make the use of 'orphan' works more unlikely.

Economic and cultural costs of 'orphan' works

Regardless the causes of the 'orphan' works' issue, it is certain that there are economic and cultural costs.

The existence of a demand, which is not being fulfilled by the market, indicates a problem of missing market. The problem of 'orphan' works is the problem of missing markets. Missing market is a form of market failure; that means there is a demand for some 'orphan' works and there is no legal possibility to satisfy this demand (Khong, 2007). In the case where a person seeks to use a work in a manner that requires permission from the copyright owner, but takes the decision not to use it, because he cannot locate the copyright owner, the latter misses an opportunity to obtain a licensing fee and the potential user loses the opportunity to use or eventually profit from the work. (The extensive economic analysis of the 'orphan' works' problem is beyond the scope of this paper. See more regarding this subject, Sherman, 2007; Varian, 2006; Eckersley, 2004).

Comprehensive data on the frequency with which 'orphan' works appear do not exist (Huang, 2006). The magnitude and the precise contours of this problem across various categories of works remain largely unknown. Numerous anecdotal examples and practical experience suggest though the existence of real problems because of unknown, unidentifiable or unlocatable copyright owners.

One of these examples gives the British Library. The British Library estimates 40% of its copyrighted collections are 'orphan' works. From the total collection of photographs of seventy institutions (around 19 million), the percentage of photographs where the author is known (other than for fine art photographs) is 10% (Gowers Review of Intellectual Property, 2006, p. 69). According to a recent survey from the Association des Cinémathèques Européennes (2005) 50 000 of surveyed audiovisual works were 'orphan', mostly non fiction and pre 1945/1950 (Seminar. European Digital Libraries Initiative: the Stakeholders' Perspectives, http://ec.europa.eu/information_society/activities/digital_libraries/doc/seminar_14_september_2007/ace_perspective.ppt).

Some years ago Carnegie Mellon University's Library studied a sample of about 270 items from its holdings: librarians could not locate the owners of 22% of the works. Similarly Cornell University librarians attempted to clear copyright on 343 monographs from the early 19th century to mid-20th century for a digital archive project on agriculture. They spent 20 000 dollars and many hours working on copyright issues, seeking permissions, tracking children of deceased authors, contacting publishing houses and more. 14% resulted to denial. But with the majority, 58% Cornell was unable to determine the current owner (Carlson, 2005). For old photographs the situation is even more difficult to find the copyright owners; e.g. one library at Cornell University owns a collection of over 350 000 unpublished photographs, yet only 1% of the photographs have an indication as to who created the photographs (George, 2001).

Over 1200 silent films from 1920's are still under copyright. According to a 1993 report on film preservation by the Library of Congress more than 80% of films made before 1929 had been lost to deterioration, because they were 'orphan' works (Librarian of Congress, Report on film preservation: A study of the current state of American film preservation, 1993, p. 5).

The economic costs of the 'orphan' works problem are substantial, although not easily measurable by statistics. One reason these costs are so high is that the violation of the exclusive rights of the 'orphan' work's copyright owner without permission may cause liability for copyright infringement. In order to alleviate the risk of a copyright infringement and to avoid the potential costs of a lawsuit, prospective users of 'orphan' works generally try to locate the copyright owner and to obtain permission before using an 'orphan' work. Time, effort and resources are spent in Sisyphean endeavours to identify the owners of the 'orphan' works. It becomes obvious that the high transaction costs associated with using 'orphans' works produce additional economic and cultural costs upon the public, because they make the use of those works more unlikely.

Under the current system, the search costs incurred by an author seeking to obtain permission to use an 'orphan' work often become prohibitive. After all according to their definition, 'orphan' works are copyrighted works whose owners cannot be identified or located. As a result, while potential users consume time, energy, and money searching for a copyright owner, they have no guarantee that this search will bring them successful results and not to a dead end. Some wealthy individuals and large corporations might be able to afford these financial and legal risks, most of the perspective users though are not. In consequence thereof most perspective users avoid using 'orphan' works altogether, because they lack the financial, legal, and human resources to withstand the potentially devastating costs of a lawsuit (Sherman, 2007, p. 19).

This de facto prohibition of 'orphan' works' use occurs not because the copyright owner has asserted any rights in the work or because an agreement cannot be negotiated between the copyright owner and the user but merely because the user could not locate the owner. The result of these high transaction costs is that many productive and beneficial uses of 'orphan' works are precluded without any tangible benefit to the copyright owner.

On the other side it is also supported that if the potential user cannot obtain clearance for an 'orphan' work and nonetheless uses it, nothing adverse will happen. The majority of copyright owners in 'orphan' works has simply vanished and will never reappear. Additionally there is no Society of Orphan Works Owners to protect their interests. Even if they have not vanished, there are minimal chances to discover that an infringement has taken place. And where they do discover it,

the wrongful act is mostly unlikely to attract any severe sanction, particularly if the work and consequently the infringement are of trivial commercial value (Philips, 2007, p. 633).

Nonetheless this uncertainty of being subject to an infringement suit is like an axe hovering over the user's head and maybe some private users could tolerate this Damocles shore, taking the form of uncertainty but it is definite that large scale users, such as libraries, archives, museums and educational institutions could not.

Those institutions maintain huge collection of works (literary, audiovisual, works of fine art and visual art, postcards, brochures, pamphlets, musical e.tc.), few of which have any indication of who the author was. The 'orphans' could never be digitised and be available to the public without the prior consent of the right-holder. The institutions under these circumstances face a dilemma between fulfilling their institutional purpose of preserving and making works available to the public and complying with the copyright law and minimizing their liability for infringement (Report on Orphan Works, 2006, p. 25). If those institutions try to locate copyright holders investing substantial amount of time, energy and money and their efforts find a dead end, they will be extremely cautious in moving forward and using the work (Thompson, 2006, p. 823). Most of those institutions have a tight budget and therefore they could use limited resources to locate right holders of 'orphan' works but also they are anxious about the risk of liability originating from using 'orphan' works. This anxiety could lead to the handcuffing of important digitization projects.

Therefore 'orphan' works constitute to libraries a considerable barrier to preservation and pose a serious problem for any plans to provide digital access to these collected materials. At the same time many 'orphan' works may be irretrievably lost, despite the best efforts of libraries and archives (Thompson, 2006, p. 813). Libraries and archives are the last bastion for many 'orphan' works, that no other party is interested in preserving.

The multi-territorial issue

Another factor that adds complexity to the 'orphan' works' problem lies in the territorial nature of copyright. As a rule, copyright is granted with respect to a particular territory and gives protection to the nationals of that territory (Torremans, 2007). Despite a certain degree of harmonization, national copyright systems still exist and provide different level of protection. The majority of national copyright legislations have not dealt with the issue of 'orphan' works and it is more than probable that each one will adopt a different approach. The solutions though should be interoperable in other countries. This is the reason why the

High Level Expert Group (HLG) on European Digital Libraries has suggested that the solutions the different Member States would choose to implement in order to treat the 'orphan' works' issue, should be interoperable and mutually recognizable, under the condition that the core principles were fulfilled (Final Report on Digital Preservation, Orphan Works and Out-of-Print Works, 2008, p. 14) (see below). A means to ascertain this interoperability is to establish common "diligent search" criteria (see also the Memorandum of Understanding on Diligent Search Guidelines for Orphan Works). A perspective user that he will search 'diligently' according to the criteria that a member state has established, he would be considered to have searched 'diligently' in another (Final Report on Digital Preservation, Orphan Works and Out-of-Print Works, 2008, p. 14). Consequently a work would be considered as an 'orphan' in more than one country.

For the same reason -to surmount the problems that the territorial nature of copyright causes- the 'orphan' works issues is also proposed to be one of the future work topics for the Standing Committee on Copyright and Related Rights by the European Community and its Member States in WIPO (Document SCCR/17/4). The protection and the possibility of enforcing rights stop at the national borders, except in so far as protection is extended outside the territory by bilateral or multinational treaties with other countries (Garnett, et al., 1999, p. 16). The conclusion of such an international treaty regarding the treatment of 'orphan' works would solve all those problems.

Solutions to the issue of 'orphan' works

Some solutions to the issue of 'orphan' works can be found in existing legislative texts, some others in legislative proposals and some in academic papers. Some of them are specifically tailored for 'orphan' works and some others are adjusted to have some affect on them. All of them though aim at the same target: ameliorating the 'orphan' works problem. Extending collective licenses, establishing a new copyright exception or limited liability, creating centrally granted licences or a judicially granted clearance model are some of them. There also some proposals that include copyright formalities, such as registration and renewal (Khong, 2007, p. 35; Hennig. 2008, p. 20). Those ones are not mentioned in this paper, since they would require major changes to the Berne Convention.

In drafting or discovering the perfect solution for the 'orphan' works a difficult balancing act must be undertaken; legal certainty for the exploitation of 'orphan' works and for the promotion of creativity should be assured without dissolving copyright. Protection for users and right holders of 'orphan' works must consider a number of important issues. Firstly, legal certainty for users should be provided (especially if large scale users are involved, such as libraries or educational institutions that digitise their resources). Secondly, guidance on what steps are con-

sidered as sufficient for fulfilling the prerequisite of 'diligent search' (inability to locate the copyright owner) should be given. Thirdly, a compensation mechanism for the right holder must be established (general or conditional to the right-holder's reappearance). Finally, if any safe harbour rule should be available to a user (e.g. special treatment for cultural, non profit establishments when fulfilling their dissemination purposes), its qualifications and scope of immunity must be addressed (Teng, 2007; Summary Minutes of the 2nd meeting of the High Level Expert Group on Digital Libraries, 2006, p. 4).

Extended collective licenses

The system of extended collective licensing is applied in the Nordic countries (i.e. in Denmark, Finland, Sweden, Norway and Iceland). In extended license system an agreement concluded between a collecting society and a user does not cover only the contracting parties (the collecting societies, the right owners that have given them mandate to act on their behalf and the users) but also obtains directly on the basis of the law a binding effect on non represented owners. The extension effect provides the users a protection against claims by non represented owners. The overall purpose of the extended license is to create favourable conditions for the use of protected materials from the viewpoint of the right owners and the users. Non represented right owners have a right to individual remuneration and in most cases an 'opt-out' right, a right to prohibit the use of their works (the system of extended collective license originally was designed to apply to literary and musical works for use in sound radio and television broadcasts but it has been expanded also to reprographic reproduction of printed material for educational use and for internal information in administration and businesses, to recording of radio and television programmes for educational use, to retransmission by cable or rebroadcasting and to library uses of material in digital form) (Koskinen-Olsson, 2006, p. 265).

Accordingly, extended license applies to all right holders in the concrete field, even deceased ones or unknown or untraceable right holders. That means that this system applies also for 'orphan' works, since the user could obtain a license to use the work by the representative collecting society without the fear of being sued for infringement, in the case that the right owners appear in the future, but only for the concrete works and uses.

This kind of system could be also utilised to treat 'orphan' works in all areas. To this direction is heading a new legislation proposal in Denmark. According to this proposal the use of extended collective licensing broadens from more specific areas to all areas, where right holders wish to use this system. The conclusion of such agreements will be subject to approval by the Ministry of Culture. This legislative proposal will contribute to solving the issue with 'orphan' works, insofar as

a collecting society is appointed to represent the right holders (Commission Staff Working Document regarding the Progress on the digitization and online accessibility of cultural material and digital preservation across the EU, 2008, p. 15).

The point is that the system is quite radical and it should be applied in cases where public interest is at stake, such as the exploitation of copyrighted works in the collection of libraries for specific purposes such as scientific research.

In the European context Directive 2001/29 EC gives the Member States the possibility to introduce provisions concerning management of rights, such as extended collective licenses (Preamble 18, Directive 2001/29 EC).

Copyright exception

Another solution that is proposed to the 'orphan' works' problem is to introduce a new statutory limitation or exception for the use of 'orphan' works. This proposal was drafted by the British Screen Advisory Council (BSAC) in a paper prepared for the Gowers Review of Intellectual Property (Copyright and Orphan Works: A paper prepared for the Gowers Review by the British Screen Advisory Council, 2006).

According to this solution the user could use an 'orphan' work under certain strict conditions. Prerequisite for the application of the exception is that the prospective user has made his 'best endeavours' to find the copyright owner or he has performed a 'reasonable search'. His best endeavours could be judged under the certain circumstances of every case according to some guidelines for reasonable search developed to qualify what efforts should be made to meet the requirements (Gompel, 2007).

The exception could provide for remuneration to the right holders owners, who emerge after the use of the work or not. For an activity of a commercial nature with an 'orphan' work, an exception should be coupled with remuneration in order to be justified. It could be provided though that for some limited non commercial uses of an 'orphan' work, there should be no remuneration, such as a preservation activity of an 'orphan' work, without the latter of which the 'orphan' work might have disappeared (Copyright and Orphan Works, A paper prepared for the Gowers Review by the British Screen Advisory Council, 2006, p. 15). Another issue to be clarified is the amount of the remuneration and whether the payment should be made if and when the right holders resurface or always in advance into a central fund that could use the unclaimed royalties to support the authors or the cultural industry. The payment could be fixed or it could be agreed on negotiation.

Under the current European copyright regime however no new limitation can be established in the Member States, unless it is provided in the Information So-

ciety Directive 2001/29 EC, since an exhaustive list of optional limitations is laid down. Currently in the Information Society Directive there is no limitation regarding 'orphan' works and that means that an amendment of the Directive 2001/29 EC is obligatory in order the national legislators to adopt a similar exception that would be compatible with their obligations under the Information Society Directive.

Even if this amendment takes place and an 'orphan' works exception is introduced, this should be also compatible with the three steps test (Art. 5 (5) of Information Society Directive, According to this test an exception is only permitted in certain special cases, which do not conflict with the normal exploitation of the work and do not unreasonably prejudice the legitimate interests of the right holders. Taking into account this test, this 'orphan' works' limitation should only be restricted in certain cases for specific purposes (e.g. the limitation would be established only in favour of libraries, museums, educational institutions and could be used for preservation purposes, for teaching, for scientific research, private studying or public exhibition). The second step could be surmounted, since the normal exploitation is excluded by definition. The third step could be conquered, if a payment of reasonable compensation is established for the copyright owners that might appear in the future (Hugenholtz, et al., 2006, p.188).

Although this kind of solution has indisputably some advantages (such as no need to issue a license, it applies to all kind of 'orphan' works, it does not affect moral rights), it includes also some weak points. The legal certainty of this solution depends on the 'reasonable/diligent search' criteria and on the fact that the user of an 'orphan' work should keep the information about the search he has undertaken and provide them to the right owner, if he reappears, in order to prove that he indeed has conducted a reasonable search to find him. This would be the one reason for legal proceedings, i.e. the owner to contest that the user has not conducted a reasonable search to find him and the other reason could be in regard with the amount of compensation (Gompel, 2007; Vetulani, 2008).

Limited liability

Another proposed solution is the limitation on remedies. This solution was proposed in the US' Copyright Office Orphan Works Report of 2006. Two bills were tabled in April 2008 Senate and the House of Representatives (the Shawn Bentley Orphan Works Act and the Orphan Works Act of 2008, H.R. 5889, introduced in the House of Representatives, 110th Congress, 2nd Session, 24 April, 2008). (There had been already the Orphan Works Act of 2006, H.R. 5439, introduced in the House of Representatives, 109th Congress, 2nd Session, 22 May, 2006).

The proposal provides that if, after a reasonable search in good faith, no copyright holder for a work is found, the work may be used, subject on limitation on the remedies that the right owner could obtain, if he would resurface and file a claim. A statutory provision thus is introduced that limits the liability of the 'orphan' works' users under certain conditions: the user must prove that he has conducted a "qualifying search, in good faith for the owner of the infringed copyright" (514§ b(1)(A)(i)) and that he has provided attribution to the author and the copyright owner (the aim of the Bill is to make clear to the public that the author or the copyright owner of the work is a different one from the user in question). In the Bill there is no definition of what search qualifies as efficient. The court will decide whether the search was 'reasonably diligent' and will take into consideration whether the user reviewed information maintained by the Register of Copyrights, used reasonably available expert assistance and reasonably available technology. According to the Bill "the Register of Copyrights shall maintain and make available to the public, including through internet, current statements of best practices for conducting and documenting a search" (514§ b(2)(B)(i)).

The limitation on remedies is not available though for those who "fix the work in or on a useful article that is offered for sale or other distribution to the public". This section (514(d)), which was not in the 2006 version of the Orphan Works Act, excludes the possibility of using an 'orphan' photo, design e.tc. on a mug, t-shirt and other merchandising products.

Limitation is provided on both monetary and injunctive relief. Monetary relief is limited to 'reasonable compensation' which amounts to a rate of a license fee. In the case of a non commercial use no monetary relief is due at all, under the condition that the user ceases the use expeditiously, after he receives a notice from the copyright owner (Report on Orphan Works, 2006, pp.115-119). Injunctive relief is limited, if the 'orphan' work has been incorporated in a derivative work and the user of the 'orphan' work pays the copyright owner a reasonable compensation and makes adequate attribution. On the contrary, full injunctive relief is due when there is no transformation of its content, e.g. the work is republished or posted on the internet. Statutory damages are unavailable to the copyright owners, since are available only for works that are registered with the Copyright Office before the commission of infringement (Ginsburg, 2008).

The main advantages of this solution are that it includes all works and it affects no existing rights and limitations. Additionally it is claimed also that, since users do not compensate the copyright owners in advance but only if the copyright owners resurface and file a claim against them, this solution is more cost effective than solutions that impose to obtain always a license by a public authority or a collecting society. The last argument though is questionable. Keeping search records, the costs

of assessing the likeliness of possible future claims, the litigation costs, the paying of reasonable compensation, if the copyright owner finally appears, are adding up to a significant amount (Hugenholtz, et al., 2006, p. 190; Gompel, 2007).

To the disadvantages of this solution accounts also the legal uncertainty for the users of 'orphan' works; only the court can judge whether the search they have performed at the time they used the 'orphan' work qualifies as diligent and therefore their liability is limited. Till then they face always the risk of paying significant monetary damages.

The last and probably the most important disadvantage is the concerns certain right holders expressed that many potential users could conduct a not diligent search to find the right owners having as result to label inaccurately their works as 'orphans'. This, coupled with the fact that the copyright owner has to file judicially a claim in order to obtain his disbursement could make them fall in despondency and shake the balance between the copyright owners and the users of the 'orphan' works.

Centrally granted licenses

Another solution for the 'orphan' works problem is to apply to a competent public authority to obtain a compulsory license in order to use a particular work for a particular use, when the copyright owner of this work is unlocatable, despite the reasonable search that the potential user has conducted. Such a licensing system has been established to Canada (also to UK, to Japan, to South Korea and India, Annex I to the Interim Report of the Copyright Subgroup of HLG, 2006, p. 8; Vetulani, 2008, p. 10). In order to understand how this system works, the Canadian regime for non locatable owners will be examined.

In Canada the solution finds a legislative regulation to Copyright Act, Section 77. According to this Section a person who wishes to obtain a license to use a published work, a fixed performance, a published sound recording or a fixed communication signal, whose copyright owner cannot be located, despite the reasonable efforts he has made, petitions the Copyright Board for a license. Regarding the reasonable effort that the petitioner has to conduct in order to identify the copyright owner, the Copyright Board advises the potential user where to check for relevant information (contact different collective management societies and publishing houses, universities and museums, provincial departments of Education, registration systems of copyright offices, to investigate inheritance records, to consult indexes of national libraries and simple to search the internet, "Unlocatable Copyright Owners Brochure", Copyright Board of Canada, <http://www.cb-cda.gc.ca/unlocatable/brochure-e.html>). The Board is flexible on what constitutes a reasonable search depending on the nature of the work being used.

The Copyright Board after his satisfaction regarding the search efforts issues to the applicant a non exclusive license, subject to the specific terms and conditions that the Board establishes. The Copyright Board has the discretion to grant the license but there is no such obligation. The purpose for which the applicant intends to use the 'orphan' work is irrelevant (Carrière, 1998). The license is given to works of domestic or foreign origin -always published or fixed, so as the divulgation right of the author to be safeguarded- limited to uses only within Canada and only for infringing ones (there is no need for license for non infringing uses). The license also stipulates a license fee that corresponds to an ordinary royalty rate (material factors are taken into consideration such as market price, commercial or non-commercial nature of the applicant and intended use), in the event that the Board thinks that royalties should be paid. When the use is small or where it is possible that the work is in the public domain, the Board may ask the licensee only to undertake to pay the copyright owner if he appears within the five years following the expiry of the license (Annex I to the Interim Report of the Copyright Subgroup of HLG, 2006, p. 9). The license fee is paid to the corresponsive collecting society or it is deposit in an escrow account. The copyright owner is entitled to collect royalties within five years after the expiration of the license. If no copyright holder resurfaces within these five years, the Copyright Board could allow a collective society to dispose of the royalties as it sees fit for the general benefit of its members.

This system of centrally granted license is effective also for mass digitization projects as mass protocols are worked out at the offset. The Board has developed a multiple work protocol for unlocatable copyright owners in eleven non exclusive licenses issued to the Canadian Institute for Historical Microreproduct authorising the reproduction of 6675 works (www.cd-cda.gc.ca/unlocatable/27-b.pdf; Annex to interim Report of the Copyright Subgroup of HLE, 16.10.2006, p. 9).

Point of criticism of this system is the lengthy and expensive process. The Canadian Copyright Board supports however that a decision is issued within 30 to 45 days after the required information is gathered (Bouchard, 2008). Another claimed disadvantage is that this system applies only to published works. Apart from the fact that this provision is established to protect the moral right of the author (right of divulgation), in practice there are cases where the Copyright Board has presumed publication, if conclusive evidence was hard to provide. The Board relies on indicia that would not meet the civil burden of proof in circumstances, where there was at least some evidence allowing it to conclude that a work has been published (Gompel, 2007; Copyright Board of Canada, Canadian Centre of Architecture, (Re) 17.1.2005, 2004-UO/TI). Another drawback is the relative small number of applications that the Copyright Board examines. This poses economic and practical questions: is this system economically viable? From an economic perspective the Canadian system places enormous demands on stakeholders' time and money. A cost benefit analysis would be useful to gauge

approximate expenses with regard to these factors. The risks for stakeholders and copyright holders need to be clearly assessed and administrative systems and payment agreed on by all users prior to use.

Regardless those disadvantages the Canadian system seems to provide a practical solution to the 'orphan' works problem. The advantages of this system are numerous: the most important though is that the license granted by the Copyright Board provides the user with legal certainty, since it protects him from possible subsequent prosecutions. At the same time the Copyright Board, an independent public authority, guards the interests of the copyright owners controlling in concreto the quality of the research that the petitioner and perspective user has conducted. The Copyright Board only after is convinced regarding the petitioner's good faith issues a license tailored for the specific needs of each case. Finally, and most importantly, the system provides for reimbursement of the copyright owner for the use that has been made accordingly to the issued license, in the case that he resurfaces.

The French model for audiovisual works

The system of voluntary collective licensing could not be an effective solution for the 'orphan' works problem. Each copyright owner has the freedom to choose whether he will authorize a collective management society to represent and exercise his rights, except certain limited cases where the collective licensing is mandatory by law (in European community law such a system exists in the area of cable retransmission rights in Satellite and Cable Directive -Council Directive 93/83/EEC of 27.9.1993 on the coordination of certain rules concerning copyright and rights related to copyright applicable to satellite broadcasting and cable retransmission, OJ L 248, 06.10.1993, p. 15, art. 9 (1)). The mandatory collective rights management has to be exercised with vast reserve, since it does not seem compatible with international copyright conventions (Ficsor, 2006). That means that even if a prospective user obtains a license by a collective management society, this does not cover with certainty the work that he would like to use, because despite the large amount of repertoire and the extensive network of bilateral agreements a collective management society cannot represent all the relevant works in the world (Hugenholtz, et al., 2006, p. 182).

Something similar has applied in France INA (Institut National de l' Audio-visuel). INA is in charge of the preservation and exploitation of France's public broadcast archives since the beginning of Radio and Television. INA has concluded a number of agreements with five collective management organizations operating in the audiovisual and sound communication field (SADM, SCAM, SDRM, SACEM and SESAM) dated November 1996, with its two

amendments dated June 2000 and June 2005. Under these agreements INA has the authority to use the audiovisual and sound material of these organizations' catalogues for any mode of exploitation. This agreement facilitates the exploitation of INA's archives but it is not a panacea, since for the works that are not included in the repertoire of those collective management institutions, INA still has to locate their right holders and clear the rights (See analytically Annex 1 of Final Report on Digital Preservation, Orphan Works and Out-of-Print Works, The INA's performance of its legal deposit role, 2008).

Judicially-granted clearance

The French Intellectual Property Code (Art. L. 122-9 and 211-12 for related rights) includes a provision for dealing with the risk of blockage by allowing the judge to make various arrangements in view of exploiting the work (Art. L. 122-9. "In the event of manifest abuse in the exercise or non exercise of the rights of exploitation by the deceased author's representatives referred to in Article L. 121-2, the first instance court may order any appropriate measure. The same shall apply in the event of a dispute between such representatives, if there is no known successor in title, no heir or no spouse entitled to inherit. Such matters may be referred to the courts, inter alia, by the Minister responsible for culture."). According to this provision if there is no known successor in title, the first instance court may order any appropriate measure in regard with the exercise of the rights of exploitation.

The act of August 2006 introduces a special provision connected to the exploitation by INA of its archives (Final Report on Digital Preservation, Orphan Works and Out-of-Print Works, 2008, p. 13).

Legal presumption of representation

In this system it is presumed by law that a collecting society has a general authorization to represent the right holders in a specific sector. This permits the collective management society to issue 'blanket licenses', which cover all the right holders, even the ones that they are not their members. In the 'blanket license' a user will not pay any claim from unrepresented right owners and this will be settled by the collective management society (Gompel, 2007; Ficsor, 2006).

Similar would be the possibility to allow a representative collective management society to grant an indemnity or security to a user who after a reasonable search has not been able to identify and locate a copyright owner. In the Netherlands for example exists Foto Anoniem. Foto Anoniem is a foundation, allied to the professional photographers' organization that has in its disposal a vast directory of photographs. A user who wants to use a photo, whose author

is unknown to him, can conduct this foundation to trace the photographer. In most of the cases Foto Anoniem is able to trace the photographer. If this is not the case, Foto Anoniem assures the user legal protection by granting him an indemnity clause; in this way it undertakes the commitment to protect the user against liability for copyright infringement. In return the user has the obligation to pay to Foto Anoniem a fair compensation, equal to the usual license fee, that is reserved to pay the right owners in the case they resurface (Gompel, 2007; Hugenholtz, et al., 2006, p. 185). Similar model is also applied in Belgium, in the field of visual arts (SOFAM) and it is scheduled in France in the field of literary works (SGDL).

This kind of solution provides the user a certain degree of legal certainty but it does not provide him with a full guarantee. The indemnity clause or security does not preclude that the copyright owner will not file any claim in the case that reappears. The indemnity clause provides the user with a financial security in the event of civil liability but it does not secure him the legal use of an 'orphan' work, since the right owner can seek injunctive relief that would prohibit any further use of the previous 'orphan' work (Koskinen-Olsson, 2006, p. 267). Apart from this the user could face also liability under criminal law. This solution should be coupled with measures to provide on the one hand legal certainty to the users and on the other hand pledges for the legitimate interest of the right owners defining the conditions under which the use of 'orphan' works is permitted (Hugenholtz, et al., 2006, p. 185; Gompel, 2007).

Treatment of 'orphan' works at European level

Within the European Union the issue of 'orphan' works has been a matter of concern for the last years and especially it came to the forefront in the face of Digital Libraries. In 2005 the European Commission has launched the Digital Libraries Initiative -part of "i2010 a European Information Society for growth and jobs" initiative- having as a target to make Europe's cultural resources and scientific records accessible online and preserve them for future generations (COM(2005) 465 of 30.09.2005 and COM(2007) 56 final of 14.02.2008). In August 2006 the Commission adopted a Recommendation on the digitisation and online accessibility of cultural material and digital preservation (Recommendation 2006/585/EC, OJ L 236/28 of 31.08.2006) which called upon the Member States inter alia to create mechanism to facilitate the use of 'orphan' works, following consultation of interested parties (Art. 6 (a)) and to promote the availability of lists of known 'orphan' works and works in the public domain (Art. 6(c)).

The Commission decided in February 2006 to set up a 20 member High Level Expert Group on Digital Libraries (HLG) (Commission Decision 2006/178/EC of 27.2.2006, OJ L 63/25 of 4.3.2006) (its mandate expired at the end of 2008 but it was renewed by a Commission Decision of 25.03.2009 at least for 2009, Commission Decision 2009/301/EC). The task of this Group is to monitor and access progress in the implementation of the Commission's Recommendation, "to advise the Commission on how to best address the organizational, legal and technical challenges at European level" and "to contribute to a shared strategic vision for European digital libraries" (Art. 2 Decision 2006/178). Within this Group a Copyright Subgroup was formed in order to analyze the copyright issues arising in the context of Digital Libraries initiative (Art. 4(2) Decision 2006/178). The 'orphan' works issue was one of the issues that the Copyright Subgroup dealt with.

The Copyright Subgroup adopted a Final Report on Digital Preservation, Orphan Works and Out-of-Print Works in June 2008. This Report partly consolidates the recommendations made in previous reports. Regarding 'orphan' works, the Report lists a number of measures, including voluntary and regulatory measures that could be adopted by the Member States, which are to be mutually recognised at the interstate level under that condition that they fulfil the generally accepted core principles. Those ones are the following:

- Cover all orphan works (those with unidentified or non locatable right holders), on the basis of a shared definition.
- Include guidance on diligent search.
- Include provision for withdrawal, if the rightholder reappears.
- Offer cultural, not profit establishments a special treatment, when fulfilling their dissemination purposes, to be further discussed between stakeholders.
- Include requirement for general remuneration or remuneration, if the right-holder reappears" (Final Report on Digital Preservation, Orphan Works and Out-of-Print Works, 2008, p. 15).

The Copyright Subgroup having realized the importance of the 'diligent search' in the context of 'orphan' works has established some general guidelines that should be followed:

- Any solution for orphan works should be applicable to all kinds of protected works.
- The potential user of orphan works should be required to conduct a thorough search in good faith in the country of publication/production if applicable, with a view to identifying, locating and contacting the copyright owner, prior to the use of the work.

- A flexible approach should be adopted to ensure an adequate solution in dealing with individual circumstances of each orphan work, taking into account various categories of works.
- Guidelines or best practices specific to different kinds of work can be worked out by stakeholders in different fields.
- Any regulatory initiative should refrain from prescribing minimum search steps or information sources to be consulted, due to rapidly changing information sources and search techniques” (Final Report on Digital Preservation, Orphan Works and Out-of-Print Works, 2008, p. 15).

Finally, the Report realizes also the need to supplement those solutions with practical tools and describe certain key principles for the development of rights clearance centres for ‘orphan’ works and databases. The national rights clearance centres could act as portals and common access points for clearance of rights and be accessible across borders. The perspective user will have the obligation of conducting a diligent search for the right holders of the ‘orphan’ works and of documenting it (Final Report on Digital Preservation, Orphan Works and Out-of-Print Works, 2008, p. 25). The rationale of a developing a database is to help users in their search endeavours. Interlinking national databases considered to be necessary in order for establishing a European common multilingual access point and creating a European resource.

According to these key principles, the combined solution shall through authoritative registers enable institutional and commercial users to conduct an online search in order to find the identity and the location of a right holder of the work that they want to use, to find out whether someone else has already conducted a diligent search, or even to find out whether the work is already registered as available digitally from another e.g. library or museum. Structured databases or lists of ‘orphan’ and public domain works do not currently exist in Europe (or elsewhere).

A project called ARROW (Accessible Registries of Rights on Orphan Works towards the European Digital Library) aims to develop an infrastructure that will include the creation of a European distributed registry of ‘orphan’ works and will provide for management of any type of rights information to facilitate the implementation of business models for both digital libraries and private e-content providers. European national libraries, publishers and collecting societies, also representing writers participate in this project (Final Report on Digital Preservation, Orphan Works and Out-of-Print Works, 2008, p. 11; <http://www.arrow-net.eu/about-arrow>). Another EC funded project, MILE (Metadata Image Library Exploitation), aims to improve the use, accessibility and trade of digital images

throughout Europe. It deals with metadata for digital images and one of its aims is to create an 'Orphan Works Database'. This database would be a repository for all 'orphan' works and users are invited to contribute information regarding those works (<http://www.mileproject.eu/introduction>). Projects as ARROW and MILE are actions that will help to fill the gap.

The only risk lurking in these databases containing information on 'orphan' works is that the subsequent users could rely on the search results of others without conducting a proper search themselves. Once a work is labelled as an 'orphan', the subsequent users will not conduct a thorough and diligent search but they could rely on the previous searches that they could be inefficient and unreliable due either to the superficiality of the search or due to new elements that have arisen. For that reason in order those databases to be effective they have to be combined with legislative solutions regarding the minimum standards of reasonable and diligent search (Report on Orphan Works, 2006, p. 78; Gompel, 2007).

Concerning the development and the establishment of diligent search criteria for right holders there was already a move towards this direction. An agreement on copyright, a Memorandum of Understanding on Diligent Search Guidelines for Orphan Works, was signed by libraries, archives and right holders in June 2008 and was welcomed by the HLG (http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg/orphan/mou.pdf). The four sector specific working groups (text, audio-visual, visual/photography and music/sound) established due diligence guidelines for the respective fields (Sector Reports http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg/orphan/appendix.pdf). These guidelines contain also a definition of 'orphan' works, recommendations regarding the procedure and methodology to be applied, clarifications how searches for right holders have to be handled for libraries and archives and representatives of publishers, photographers, authors, record and film companies and a list of appropriate information resources that are available for research purposes (Joint Report, http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg/orphan/guidelines.pdf). The aim is to provide a practical tool to assist institutions in identifying and locating right holders. The guidelines should be taken into account when searching for right holders and a work can only be considered 'orphan', if the relevant criteria have been followed without finding the right holders. One point of the agreement is also to promote the guidelines as acceptable standards at European level and to encourage the national entities to link the described information resources provided in the Reports (Joint and Sector Reports) to national resources.

In conclusion the European Commission through the Copyright Subgroup has not decided for one solution. It is left at the discretion of the Member States to choose

the more appropriate one but the overall solution to find a workable mechanism to use 'orphan' works should concentrate on four implementation issues:

- sector specific criteria for diligent search for right holders to copyright works that a user needs to fulfil prior to the use of the work;
- databases of 'orphan' works to facilitate users in their search, which is needed irrespective of any legislative solution;
- a mechanism to clear the rights to use an 'orphan' work; and
- mutual recognition of different solutions in various Member States to achieve the cross border effect (Final Report on Digital Preservation, Orphan Works and Out-of-Print Works, 2008, p. 25).

Proposed solution

Instead of trying to reinvent the wheel, it would be better to look whether some solutions are already available in international legislative framework. Upon this basis a specific tailored solution for 'orphan' works could be constructed.

Indeed the Berne Convention includes in its appendix special provisions for developing countries which are relevant to 'orphan' works. Developing countries can choose to exercise these provisions, which weaken the right of reproduction and the right of translation, provided that they lodge a notification of their intention with the Director General of the World Intellectual Property Organization (Art. I). According to Article III a developing country may substitute the exclusive reproduction right with "a system of non-exclusive and non-transferable licences, granted by the competent authority" under certain conditions provided in the same Article and in Article IV. It allows a developing country adopting the special provision to grant a non exclusive license to a national, subject to other conditions, to reproduce a published edition after a certain period of time -between three to seven years depending on the subject matter of the work- "where the identity or the address of the owner of the right of reproduction is unknown" (Art. III(4)(a)(ii)). Article IV imposes the preconditions that the said copyrighted work to be licensed has to be orphaned. The applicant has to establish that, after due diligence on his part, he was unable to find the owner of the right (Art. III(4)(a)(ii)). Additionally, Article IV(6) requires that countries make provisions for the payment of "just compensation" by the licensee in return for the licence (Khong, 2007).

Quite unexpectedly, the solution provided in the Berne Convention is exactly the right prescription for the problem of 'orphan' works for all countries. So after all maybe the solution to the problem of unknown ownership is beneath our eyes. If

it can be proven that efforts to establish the ownership of the copyrighted work have been futile, it is welfare enhancing to allow using the work without further consent. Nevertheless this permission should be granted by an authority that would reduce the risk of infringement and abuse. Although some transaction cost of obtaining approval might be incurred, it would be more cost effective than bearing the risk of being found to have committed an infringement. Finally as for whether a licensing fee should be imposed, this question should be answered by reference to the type of 'orphan' work that is each time involved. This solution presents many similarities to the centrally granted license system that is applied already in Canada.

Conclusions

'Orphan' works constitute a serious problem for all stakeholders by making it practically impossible to locate the rightsholders to license their work. This situation could be to the detriment not only of the stakeholders but generally speaking of the public. Taking into account the magnitude of the problem, a solution is a dire necessity. After having examined the different solutions, the most appropriate one seems to be the Canadian model, since it keeps the balance between the legitimate interests of users and right holders. The unavoidable disadvantage of this model lies in the inability to distribute the royalties to the right holders, since –and for the time being– they remain untraceable. The non distributed royalties though could be used for cultural aims or for the creation of 'orphan' works database, contributing to the amelioration of the 'orphan' works issue.

Nevertheless for a victorious battle and for a suppression of 'orphan' works, the issue has to be tackled at different fronts. The effective treatment of the 'orphan' works' problem does not deal only with the issue of finding solutions to the present situation but also with the issue of preventing the prolongation of the 'orphan' works problem. Given that the 'orphan' works' problem originates primarily in the lack of identification information on the work, a more efficient method of providing clarification of the copyright status should be developed.

One solution could be the inclusion of metadata, information on right holders regarding digital born material. Authors and copyright owners could be encouraged to provide adequate copyright information for digital works and to incorporate inclusive rights management information, e.g. using digital metadata and digital watermarking can permanently attach author information to copyrighted work (art. 7(2) Information Society Directive; art. 12 (2) WCT). A wide supply of rights management information would enhance transparency, would facilitate the public in identifying the copyright owners of the works in question and therefore would accommodate the rights clearance (see also the proposal to alter the

abovementioned provision of the Information Society Directive and to provide that the protection of rights management information would be only granted, if this information is deposited in a publicly accessible database, Hugenholtz, et al., 2006, p. 179).

A second solution could be the creation of databases containing information on the works and on copyright owners. In this manner both right holders and potential users will have a clear picture over the use of their works and whether the work they want to use is already an 'orphan' one (Hugenholtz, et al., 2006, p. 181).

Another alternative method that has been proposed to prevent future 'orphan' works is the application of creative commons licenses (Hoom, 2006, p. 12). These licenses provide a direct link between the work and its license (Hugenholtz, et al., 2006, p. 180). A disadvantage could be though that the creative commons licenses do not allow for direct remuneration.

As a closing remark we could mention that from the previous analysis it is obvious that it is indispensable to allow the adoption of 'orphan' works, but only the truly 'orphaned' works. The regulatory provisions regarding 'orphan' works should not become a means for abuse and circumvention of basic copyright principles. It should be avoided works to be characterised as 'orphans' in an arbitrary and opportunist manner, solely to be excluded of the copyright rules, as it happens in the *Pirates of Penzance*, where the prisoners claim to be orphans in order to be released by the pirates (The *Pirates of Penzance* were too tender hearted and they made a point of never attacking a weaker party than themselves and never molesting an orphan, because they were orphans themselves and know what it is. The consequence thereof was that everyone who was captured claimed to be an orphan, The *Pirates of Penzance*, a comic opera by Sullivan and Gilbert). 'Orphan' works should be 'released' and could be given to 'adoption' but only if the necessary requirements are fulfilled.

REFERENCES

- Benabou, V.L. and Dussolier, S.** (2007) Draw me a public domain, in *Copyright Law, A Handbook of Contemporary Research* (ed. Torremans, P.), Edward Elgar, Cheltenham, UK, Northampton, MA, USA.
- Bouchard M.,** (2008), *Unlocatable Copyright Owners: Canada's Orphan Works Regime Mile*, Conference Metadata IPR Seminar, Malta, June 4, 2008, Available at: www.mileproject.eu/asset_arena/document/RD/IPR_CEPIC_JUNE_2008_ORPHAN_WORKS_BOUCHARD.PPT

Bottis M., (2007), The Google Library Project and Copyrights of publishers and authors, E-LIS. [online], 7, 171-184. Available at: E-LIS: <http://eprints.rclis.org/archive/00014598> [Accessed 11 May 2009].

Brito, J. and Dooling, B. (2005) An Orphan Work Affirmative Defence to Copyright

Infringement Actions. Michigan Telecommunications and Technology Law Review, 12, 75-113. Available at: SSRN <http://ssrn.com/abstract=942052> [Accessed 11 May 2009].

Carlson, S. (2005) Whose work Is It anyway?. Chronicle of Higher Education, 51(47), p. A33

Carrière, L. (1998), Unlocatable Copyright Owners: some comments on the licensing scheme of Section 77 of the Canadian Copyright Act' 1998, [online]. Available at: <http://www.robic.ca/publications/Pdf/103-LC.pdf> [Accessed 11 May 2009].

Cohen, J. (2006) Copyright, Commodification, and Culture: Locating the Public Domain, in The Future of the Public Domain (Identifying the Commons in Information law) (eds. Guibault, L. and Hugenholtz, B.), Kluwer Law International, The Netherlands.

Colleran, E. (December 2007-January 2008) Orphan Works: A look at the issue of various perspectives. Against the Grain.[Online] Available at: http://www.against-the-grain.com/TOCFiles/v19-6_Colleran.pdf [Accessed 11 May 2009].

Davies, G. (1995), The Convergence of Copyright and Authors' Rights – Reality or Chimera?. International Review of Industrial Property and Copyright Law, 6(26), 964-988.

Dutfield, G. and Suthersanen, U. (2008), Global Intellectual Property Law. Edward Elgar, Cheltenham, UK, Northampton, MA, USA.

Eckersley, P. (2004), Virtual Markets for Virtual Goods: The Mirror Image of Digital Copyright. Harvard Journal of Law and Technology, 18 (1), 85-166.

Ficsor, M. (2006), Collective Management of copyright and related rights in the digital, networked environment: Voluntary, presumption-based, extended, mandatory, possible, inevitable, in Collective Management of copyright and related rights, (ed. Gervais), Kluwer Law International, The Hague.

Garnett, K. Ray, J. and Davies, G., eds., 1999. Copinger and Skone James on Copyright. 14th ed. London: Sweet & Maxwell.

Geller, P. (1994), Toward an Overriding Norm in Copyright: Sign Wealth. *Revue Internationale du Droit d'Auteur*, 159, 3.

George, C. (2001), Exploring the feasibility of seeking copyright permissions, Technical Report.

In ALA Annual Conference (American library Association). San Francisco, USA, 15-20 June 2001. [Online] Available at: <http://www.library.cmu.edu/Libraries/FeasibilityStudy.ppt> [Accessed 11 May 2009].

Ginsburg, J. (2008), Recent developments in US Copyright law: Part I – 'orphan' Works. *Revue Internationale du Droit d'Auteur*, 217, 99.

Gompel, St. von (2007), Unlocking the potential of pre-existing content: How to address the issue of orphan works in Europe? *International Review of Industrial Property and Copyright Law* 6, 669-702.

Hennig, D.K., (2008), Copyright's Deus Ex Machina: Reverse Registration as Economic Fostering of Orphan Works, [Online] Available at: http://works.bepress.com/darrin_henning/1 [Accessed 11 May 2009].

Hoorn, E. IVIR (2006), Creative common licences for cultural heritage institutions, IVIR. [Online] Available at: http://www.ivir.nl/creativecommons/CC_for_cultural_heritage_institutions.pdf [Accessed 11 May 2009].

Huang, O. (2006), U.S. Copyright Office Orphan Works Inquiry: Finding Homes for the Orphans. *Berkeley Technology Law Journal*, 26(1), 265 -288.

Hugenholtz, P. B., Eechoud, V. Mireille M. M., Gompel, S. V., Guibault, L. and Helberger, N. (2006), The Recasting of Copyright & Related Rights for the Knowledge Economy, IVIR, The Netherlands. [Online] Available at: http://ec.europa.eu/internal_market/copyright/docs/studies/etd2005imd195recast_report_2006.pdf [Accessed 11 May 2009].

Hugenholtz, B. (2000), Why the Copyright is Unimportant, and possibly invalid,. *European Intellectual Property Review*, 11, 501-502, [online]. Available at: <http://www.ivir.nl/publications/hugenholtz/opinion-EIPR.html> [Accessed 11 May 2009].

Khong, D. W. K. (2007), Orphan Works, Abandonware and the Missing Market for Copyrighted Goods. *International Journal of Law and Information Technology*, 15 (1), 54-89.

Koskinen-Olsson, T. (2006), Collective Management in the Nordic Countries, in *Collective Management of Copyright and Related Rights* (ed. Gervais), Kluwer Law International, The Hague.

- Locke, J.** (1986), *The Second Treatise on Civil Government*. Prometheus, New York.
- Ng, A.** (2008), *The Social Contract and Authorship: Allocating Entitlements in Copyright System*. *Fordham Intellectual Property, Media & Entertainment Law Journal*, Forthcoming. [Online]. Available at: SSRN: <http://ssrn.com/abstract=1270175> [Accessed 11 May 2009].
- Petit, C. E.** (2006), *Cost Allocation and Copyright Orphans*. [Online]. Available at: SSRN <http://ssrn.com/abstract=921610> [Accessed 11 May 2009].
- Philips, J.,** (2007), *Killing the Orphans*. *Journal of Intellectual Property Law & Practice*, 2007, 2 (10), 633.
- Radin, M. J.** (2006), *Copyright defection*. *Industrial and Corporate Change*, 15(6), 981-993.
- Sherman, D.** (2007), *Cost and Resource Allocation under the Orphan Works Act of 2006*. *Virginia Journal of Law & Technology*, 12(4), 1-36.
- Sterling, J. A. L.** (1998), *Creator's Right and the Bridge between Author's Right and Copyright*. *International Review of Industrial Property and Copyright Law* 29(3), 302-308.
- Teng, S.** (2007), *The orphan works dilemma and museums: an uncomfortable straitjacket*. *Journal of Intellectual Property Law and Practice*, 2(1), 32.
- Thompson, C.** (2006), *Orphan Works, U.S. Copyright Law, and International Treaties: reconciling differences to create a brighter future for orphans everywhere*. *Arizona Journal of International & Comparative Law*, 23(3), 787-852.
- Torremans, P.** (2007), *Questioning the Principles of territoriality: the determination of territorial mechanisms of commercialisation*, in *Copyright Law, A Handbook of Contemporary Research* (ed. Torremans, P.), Edward Elgar, Cheltenham, UK, Northampton, MA, USA.
- Varian, H. R.** (2006), *Copyright term extension and orphan works*. *Industrial and Corporate Change*, 15 (6), 965-980.
- Vetulani, A.** (2008), *The problem of orphan works in the EU, An overview of legislative solutions and main actions in this field*, [Online]. Available at: [http://ec.europa.eu/information_society/activities/digital_libraries/doc/report_orphan_stagiaire_2/report_orphan_vetulani%20\(corrected%20version\)%20\(2\).pdf](http://ec.europa.eu/information_society/activities/digital_libraries/doc/report_orphan_stagiaire_2/report_orphan_vetulani%20(corrected%20version)%20(2).pdf) [Accessed 11 May 2009].
- U.S. Copyright Office, (2006), *Report on Orphan Works; A Report of the Register of Copyrights*. [online] Available at: <http://www.copyright.gov/orphan/orphan-report.pdf> [Accessed 11 May 2009].

Orphan Works Act of 2008, H.R. 5889. Available at: http://www.thomas.gov/home/gpoxmlc110/h5889_ih.xml [Accessed 11 May 2009].

Orphan Works Act of 2006, H.R. 5439. Available at: <http://thomas.loc.gov/cgi-bin/query/z?c109:H.R.5439>: [Accessed 11 May 2009].

Librarian of Congress, 1993. Report on film preservation 1993: A study of the current state of American film preservation. [Online] Available at: <http://www.loc.gov/film/study.html> [Accessed 11 May 2009].

Abandonware Ring FAQ, 2002. Available at: <http://www.abandonwarering.com/?Page=FAQ> [Accessed 11 May 2009].

Copyright Board of Canada, Unlocatable Copyright Owners Brochure. Available at: <http://www.cb-cda.gc.ca/unlocatable/brochure-e.html> [Accessed 11 May 2009].

British Screen Advisory Council (2006), Copyright and Orphan Works: A paper prepared for the Gowers Review by the British Screen Advisory Council. [Online]. Available at: <http://www.bsac.uk.com/reports/orphanworkspaper.pdf> [Accessed 11 May 2009].

Gowers Review of Intellectual Property, 2006. London: The Stationery Office. [Online] Available at: http://www.hm-treasury.gov.uk/d/pbr06_gowers_report_755.pdf [Accessed 11 May 2009].

Recommendation 2006/585/EC on the digitisation and online accessibility of cultural material and digital preservation (OJ L 236/28 of 31.08.2006). Available at: http://ec.europa.eu/information_society/newsroom/cf/itemlongdetail.cfm?item_id=2782.

European Commission, (2007), Seminar: European Digital Libraries Initiative: the Stakeholders' Perspectives. Brussels, 14.09.2007. [Online] Available at: http://ec.europa.eu/information_society/activities/digital_libraries/doc/seminar_14_september_2007/ace_perspective.ppt [Accessed 11 May 2009].

Summary Minutes of the 2nd meeting of the High Level Expert Group on Digital Libraries, 17 October 2006. [Online]. Available at: http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg/minutes/summary_minutes_hleg_2nd_final.pdf [Accessed 11 May 2009].

Memorandum of Understanding on Diligent Search Guidelines for Orphan Works (2008) [Online] Available at: http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg/orphan/mou.pdf [Accessed 11 May 2009].

Memorandum of Understanding on Diligent Search Guidelines for Orphan Works. Joint Report, (2008), [Online] Available at: http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg/orphan/guidelines.pdf [Accessed 11 May 2009].

Memorandum of Understanding on Diligent Search Guidelines for Orphan Works. Sector Reports, (2008), [Online] Available at: http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg/orphan/appendix.pdf.

Annex I to the Interim Report of the Copyright Subgroup of HLG, 16.10.2006 [Online]. Available at: http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg_minutes/copyright/interim_report_16_10_06.pdf [Accessed 11 May 2009].

i2010: Digital Libraries High Level Expert Group – Copyright Subgroup. Final Report on Digital Preservation, Orphan Works, and Out-of-Print Works Final Report, 4.6.2008. Available at: http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg/reports/copyright/copyright_subgroup_final_report_26508-clean171.pdf [Accessed 11 May 2009].

Annex 1 of Final Report on Digital Preservation, Orphan Works and Out-of-Print Works, The INA's performance of its legal deposit role, [Online] Available at: http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg/reports/copyright/copyright_subgroup_final_report_26508-annex1-final.pdf [Accessed 11 May 2009].

Commission Staff Working Document accompanying the Communication from the Commission to the Council, the European Parliament, and Social Committee and the Committee of the Regions, Progress on the digitization and online accessibility of cultural material and digital preservation across the EU, [EU COM(2008) 513 final] SEC(2008) 2372, 11.8.2008. Available at: http://ec.europa.eu/information_society/activities/digital_libraries/doc/communications/progress/swp.pdf [Accessed 11 May 2009].

Key Principles for Databases and Rights Clearance Centres for Orphan Works Available at: (http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg_minutes/copyright/key_principles_ow.pdf)

Document SCCR/17/4, Justifications for topics proposed as future work by the European Community and its Member States at the SCCR of March 10-12, 2008, WIPO. [Online] Available at: www.wipo.int/edocs/mdocs/copyright/en/sccr_17/sccr_17_4.doc [Accessed 11 May 2009].

Default Social Meaning: An Ethics of Video Game Play

Stephanie Patridge*

Department of Religion and Philosophy
Otterbein College

When we play video games, we perform actions that are partly constituted by our imagination. We imagine that we are shooting space aliens, that we are trying to take over the world, that we have joined the battle to save middle-earth. The fact that our actions occur in imaginative contexts seems to make a normative difference. Should a friend comment, upon watching us play a video game, that our deeds are actually noble or that our willingness to fight evil speaks well of our actual character, we would find ourselves perplexed. Our confusion here would stem from the fact that our friend has conflated the standards that apply to acts that occur in imaginative contexts with those that apply in actual or real world contexts. Maybe if we had actually fought evil, our deeds would actually have been noble and our character actually laudable, but the fact that we are playing a game appears to render these evaluations somehow misplaced.

Despite our tendency to resist applying moral predicates to video game actions, there are some cases where moral evaluation of our video game play seems appropriate. Consider, for example, that some video games invite us to mobilize our imagination in ways that inspire in us a kind of moral unease. Video game worlds are often excessively violent, sexually graphic, and populated by characters that are based on suspect stereotypes. That we are asked to enjoy interacting with imaginative worlds like this seems, on the face of it, morally worrisome.

Of course, that we are given to seemingly incongruous intuitions does not tell us anything about how to resolve the incongruity. On the one hand, we could reject our tendency to morally evaluate our video game play, and insist that the fact that our actions, beliefs, or attitudes occur in an imaginative context is sufficient to place them beyond the scope of the moral. That is, we could embrace amorality about video game play. On the other hand, we could reject our tendency

* Stephanie Patridge is an Assistant Professor of Philosophy at Otterbein College in Westerville, OH USA. Her research interests include: moral psychology, ethics, philosophy of art, and applied ethics. She is currently working on a manuscript on the ethics of video game play that focuses on our attitudinal responses to video game imagery.

to resist moral evaluation, and insist that our video game deeds are subject to ordinary moral evaluation. So, if an action belief or attitude would be praised or blamed in an actual context, then it should be similarly praised or blamed in an imaginative context. That is, we could embrace a straightforward and robust moralism about video game play. While each move resolves the apparent evaluative incongruity, it does so at the cost of undermining our ordinary evaluative practices in a wide-range of cases. In this project, I recommend a way of thinking about the moral evaluation of video game play that resolves the apparent incongruity, while preserving much of our ordinary evaluative practices. On my view, that a game contains a representation of a wrongful act, even in cases where the player is invited to adopt the role of the entity enacting the wrong represented, is not sufficient to render that representation a proper object of moral concern (and vice versa for representations of good acts). However, the meaning of some representations is fixed by moral facts on the ground, so that its meaning is fairly incorrigible. In these cases, game designers have a duty to recognize that the representations have a certain meaning that they cannot simply wish away, and players have a duty to consider whether enjoying such representations is morally permissible or not.

Amoralism about Video Game Play

As I mentioned earlier, moral evaluation often seems perplexing in imaginative or fictional contexts. Consider, for example, what you would think if someone were recommended to you as a person of noble character on the basis of her video game accomplishments, say because “she has performed courageously in the face of evil, when playing World of Warcraft,” or “she has consistently made certain that her citizens had access to education and the arts in Civilization III.” Such recommendations would be unintelligible to us. An obvious worry here is that her activities are imaginative or fictional. Consider for example that the characters that populate video game worlds are neither rational nor sentient, and so it is at least not obvious how they could directly place moral constraints on our behavior. That is, if video game characters themselves are not part of our moral universe, then it is unclear how to make sense of our tendency to morally evaluate our video game play at all. Further, one might attempt to undermine the moral evaluation of game play in general, and video game play in particular, by way of an argument by analogy with chess: In playing chess we pretend to attack our opponent’s army. If we make the first move, then we attack without provocation. Certainly, if any war is unjust, then one launched for no other reason than to dominate one’s enemy is. But, at least most of us are inclined to say that playing chess is an innocuous activity. In playing chess, we recognize that we are only pretending to attack our opponent’s army. The same, one might argue, goes for

video games. That we are not confused on this point seems of obvious moral salience. So, the amoralist concludes, either we must admit that our intuitions about chess are flawed, a fairly hefty price to pay, or we must give up the claim that video game play is subject to moral criticism. Given the choice between these two options, it is obvious which one we should choose. We should not moralize game play in general, and video game play in particular.

The Moral Costs of Video Game Play

Video Game Violence and its Consequences

One way to sidestep this challenge is to adopt a consequentialist perspective of video game play. Such a theorist might maintain that while there are no substantive moral costs associated with playing chess and most video games, for some games the cost of playing is too high a price to pay. For example, recent advances in video game technology have made the mechanics of game play more sophisticated and its imagery more realistic. Players take an active role in constructing a game's narrative, so that the in-game choices look increasingly like they are one's own. This is a familiar enough point. However, the consequentialist might further argue that as players engage in activities that look more and more like actual behaviors, we should begin to think about the effects of such activities. These considerations become more pressing in light of the fact that the violence and sexuality represented in video games becomes more intense as the representations become more realistic.

Of course, one need not be a committed consequentialist to see the relevance of this kind of reasoning. It is undeniable that if we are justified in thinking that there is a substantive causal link between playing certain kinds of video games and a decline in the quality of our moral interactions with others, then we have a powerful moral reason to avoid playing these games whatever our non-skeptical moral commitments. Moreover, if we can show that these kinds of costs are not incurred in playing chess, and other games with lower levels of violence and sexuality, then this is all to the good since it will support our intuitions in a wide range of cases.

But, as is often pointed out, consequential worries are empirical worries, and as such they rely on empirical evidence for their support. So, it is worth considering the current status of the empirical evidence. In a recent meta-analysis of the current literature that is taken to support a connection between violence in video games and violence in those who play video games, Johnthan Freedman argues that there is very little data to support these kinds of consequentialist worries (Freedman, 2001). In his study, Freedman draws three main conclusions. First, he argues that there is substantial, though far from overwhelming or definitive,

evidence that people who like and play violent video games tend to be more aggressive than those who like and play them less. This, he cautions us, is a “purely correlational finding” and “tells us nothing” about the kinds of causal connections that hold. Second, he argues that there is slight evidence to suggest that a player experiences an increase in aggressiveness immediately after they play a violent video game, but cautions us that “the evidence for this is minimal and is greatly weakened by limitations in the research, which provide alternative explanations of the effect.” Third, he argues that there is no evidence to suggest that playing violent video games has any lasting effect, and “there is no scientific reason to believe that violent video games have bad effects on children or on adults, and certainly none to indicate that such games constitute a public health risk.” Of course, none of Freedman’s conclusions rule out such a connection. His claim is only that the empirical research as it stands does not support such a connection.¹

Virtual Crimes and Virtue-theoretic Consequences

In light of the failure of the empirical evidence to support a substantive connection between video game play and the moral quality of associated actions, one might argue that the moral evaluation of video game play is unwarranted. Since nothing bad comes from playing video games, and video game play is a fictional activity, it is beyond the reach of morality. But, if it is the mere fictionality of an activity that allows it to escape moral criticism, then we should expect all fictional activities to similarly avoid the reach of the moral. Matt McCormick (2001) provides a compelling case that undermines this amoralist assumption: he asks us to imagine someone who uses a virtual reality suite to engage in virtual-pedophilia, virtual-genocide, or virtual-rape. Doubtless most of us would be morally repulsed by such an activity, despite the fact that it is imaginative. According to McCormick, the wrongness of these activities is best understood by focusing on the agent’s character, rather than on the consequences for her actions. “By participating in simulations of excessive, indulgent, and wrongful acts, we are cultivating the wrong sort of character...[Y]ou do harm to yourself in that you erode your virtue, and you distance yourself from your goal of eudaimonia (McCormick, 2001). In a similar vein, Monique Wonderly argues that the “problem with [violent video] games is that they may damage our empathetic faculties, and in so doing, they may be directly harming our centers of moral judgment” (Wonderly, 2008).

1. However, as Warner and Raiter (2005) have pointed out our in-game activities can have real world consequences. For example, when we purchase in-game goods, either leveled up characters or money, we participate in an activity that supports sweatshops. Such considerations, though not the ones that I am after in this paper, have obvious moral significance.

While McCormick and Wonderly are right to claim that adopting an ethical framework that focuses on the virtues will help us see what is wrong with imaginative acts like virtual-pedophilia, such an analysis does not get the wrongness quite right. In a case like this one, our moral disgust is not aimed primarily at the agent's wanton disregard for the health of her character, as McCormick claims, it is aimed at the current status of her character. We think that there must be something wrong, antecedently, with anyone who would engage in such an activity, independent of the consequences, even the consequences that might accrue to her. The fact that someone would elect to engage in such an activity is indicative of a moral failing, if not a moral failing itself. As such, any individual who would engage in such an activity must have a flawed character. If you are skeptical of this point, I invite you to imagine what you would think of your friend should you find her coming out of the virtual reality suite announcing "I just had great time in there. You can even have sex with virtual children. But hey, no worries; they are only holograms. You should try it!" Assuming that we did not have prior knowledge of our friend's virtual-exploits, I think that it is safe to say that for most of us our attitude toward our friend would be significantly diminished. This person cannot be the person that we thought she was. In some cases, we might even think that our friend's willingness to engage in such a behavior is a reason to end the friendship altogether. Contra McCormick, is not that we are worried that such behavior will make our friend a worse person, though it might. Instead, we conclude that our friend is a worse person than we thought. Anyone who would do that must be.

Still, despite the fact that virtue-theoretic consequentialism of the type endorsed by McCormick and Wonderly is inadequate to capture the wrongness involved in the aforementioned scenario, cases like this one demonstrate that merely being imaginative is not sufficient to push an act, belief, or emotion out of the realm of the moral. There are some activities that we should not undertake even imaginatively. If this is right, then the amoralist about video game play owes us an account that does more than point out that such activities occur in imaginative contexts.

Imaginative Crimes and Non-consequential Wrongs

So, what is it about imaginative acts like virtual-pedophilia that implicates an agent's character while other imaginative acts do not? If we cannot find reasonable grounds for answering this question, then lends at least some credibility to the amoralist's position, and should undermine our confidence in our intuitions not only about the morally worrisome content that we find in video games, but also in our more confident intuitions like our attitude toward virtual-pedophilia.

By way of trying to understand how to make sense of the ethical considerations that are salient here, let us return to the argument from chess one more time. The argument from chess is particularly attractive because it uses our own moralist intuitions against us. Again, the argument goes as follows: In playing chess, we pretend to act in immoral ways. If moralism is true, then playing chess is immoral. But, this is ridiculous. So, we should reject moralism in favor of amorality. A bare commitment to moralism about video game play, it seems, leads to the ridiculous conclusion that playing chess is a morally unsavory activity.

Obviously, the amoralist's argument is enthymetic. Its validity relies on an unarticulated conception of moralism. So, what exactly does being a moralist amount to? Moralism is the opposite of amorality. That is, it is the thesis that video game play is subject to moral evaluation. But, if this is what the truth of moralism amounts to, then obviously the amoralist's argument is invalid, since it does not follow from the fact that video game play is subject to moral evaluation, that playing chess or even any video game is immoral. A bare commitment to moralism does not say anything about how to evaluate any particular game, and so it does not say anything about how to evaluate chess. In order to generate particular judgments of this kind, we will have to know more about the version of moralism that is in the offing.

Act-identity Moralism

Since understanding how a version of moralism might run afoul of the chess example will help us to identify reasonableness constraints on such a theory, it is worth asking ourselves what assumptions would commit us to the offending conclusion. We might think that the first premise 'in playing chess, we pretend to act in immoral ways' is obviously true. While there is sufficient reason to worry about the truth of this premise as it is formulated, I will postpone its discussion until later in this section. Instead, let us focus on the second premise, 'if moralism is true, then playing chess is immoral.' One way to make sense of this statement of implication is to see it as invoking a version of moralism that focuses on actions, and draws no significant moral distinction between those activities that we perform in actual or real contexts, and those that we pretend to perform in fictionalized or game contexts. Let us call this view act-identity moralism. An advocate act-identity moralism holds that if it would be wrong to actually perform an activity, then it would be wrong to perform that activity imaginatively. Clearly, however, there are so many counter-examples to this thesis that it is hard to imagine anyone endorsing it. Consider that on this view it would be wrong for a researcher to play most video games for the purposes of research. What this shows is that one who attempts to determine if an in-game activity would be wrong by asking if it would be wrong to really perform it is confused. So, if act-

identity moralism is the target of the chess argument, then the target is obviously a strawman.

Attitude-identity Moralism

[A]fter all the ink that has been splattered on the question of whether appreciators' experiences include emotions of various kinds vis-à-vis fictional characters and situations, it may be disappointing to learn that it doesn't much matter.... appreciators bring much of themselves to the make-believe; their actual psychological makeup, attitudes, interests, values, prejudices, hang-ups, and so forth, come powerfully into play (Walton, 1997).

I can criticize someone for taking pleasure in others' pain, for being amused by sadistic cruelty, for being angry at someone when she has done no wrong, for desiring the bad. The same is true when responses are directed at fictional events, for these responses are actual, not just imagined ones (Gaut, 2002).

In the above epigrams, Kendall Walton and Berys Gaut claim that our moral concern in imaginative contexts should be focused on what we might loosely call one's character. Though their views are not isomorphic, they agree on a few fundamental points. First, they agree that in fictional contexts we experience emotions and other attitudinal states. Second, they agree that these emotions and attitudes are not fictional. That is, we are in some sense really frightened, mortified, or amused; we do not merely pretend to be so.² That these attitudes and emotions are not feigned, they are actual, suggests that they are expressive of our actual commitments, moral and otherwise. If the responses that we have in fictional contexts are expressive of our actual commitments, then clearly we are subject to moral appraisal in light of our responses. Moreover, it seems that our attitudes and emotions are subject to roughly the same moral evaluation in fictional contexts as in actual contexts. As Gaut puts the point, we should not respond negatively to a morally positive scene, and vice versa. We can be praised for finding the racist joke mortifying, and blamed for finding it amusing, and a similar story can be told about our responses to the representational content that we find in video games. Let us call this view attitude-identity moralism to note the fact that on this family of views it is the attitude that we take that is subject to moral appraisal, and that it does not make a moral difference whether the attitude is directed at an actual scene or an imagined one.

One virtue of attitude-identity moralism is that it can accommodate our intuitions about the researcher: it is not wrong for the researcher to play video games

2. I say "in some sense" here to respect Walton's observations that the emotions that we experience in fictional contexts are in many ways dissimilar from those that we experience in actual contexts.

for the purposes of research, so long as her attitude is clinical. Another virtue of this view is that it gives us a compelling explanation of the wrongness of virtual-pedophilia. On this view, it is the fact that she enjoys this kind of imaginative act that exposes flaw in her character. Such an attitude would be expressive of her actual commitments.

How then does the attitude-identity theorist fare against the argument from chess? If in playing chess we pretend to wage war against our opponent, then it seems that the fact that we enjoy playing chess would expose a flaw in our character. So, one might conclude that this version of moralism runs afoul of the reasonability constraint on moralism. Despite the plausibility of this line of argument, there are adequate resources available to the advocate of attitude-identity moralism to resist it. The argument from chess rests on an unexamined premise: in playing chess we pretend to act immorally. At least in standard cases, this description does not capture the phenomenological aspects of our game play. A chess player is not so much waging an imaginary war, as she is trying to solve an algorithmic puzzle. She does not imagine that she actually kills her opponent's knight, and enjoy the thought of this. Instead, it is more reasonable to see her as having eliminated one of her opponent's game defenses. The enjoyment that she takes at such a maneuver is independent of the fact that the game is representation of a war. Instead, in playing chess we see through the representation of war to the game. Consider, for example, an analogous game whose rules are identical to the game of chess, but whose pieces do not represent a standing army. Instead, they are represented by geometric shapes. Doubtless this game would be as enjoyable as the standard game of chess to any chess aficionado, though they might have reasons for resisting such a change. This suggests that it is not the war element of chess that we respond to, it is the strategic element. If this is right, then it is open to the attitude-identity moralist to argue that the chess player is not actually taking pleasure in waging war, but in performing well in a game of strategy. So, enjoying playing chess, at least in ordinary cases, does not expose a flaw in one's character. If this is right, then attitude-identity moralism passes the chess test.

Still, despite the obvious virtues of attitude-identity moralism it remains open to a wide range of counter-examples (Patridge, 2008). For example, many of us enjoy watching monster movies, and in some cases we might find ourselves enjoying watching the big, bad monster go on a rampage. While it is doubtless that our ability to enjoy monster movies says something about the kind of person we are, it is doubtful that it says anything that is particularly morally interesting. The same goes for video games where we play the part of monster who rampages through the downtown of some city or other. That we enjoy playing this game says relatively little about the kind of person we are. This is not to say that the at-

titudes and emotions that we experience in fictional contexts never say anything morally interesting about us, they do. My only claim here is that in many cases they do not. In so far as the attitude-identity theorist is committed to the claim that as a matter of course they do, such a theory represents an over-moralizing of our responses to fictions in general, and video games in particular.³

Representations of Wrong and Wrongful Representations

Though we should be wary of over-moralizing our video game play, I think that the attitude-identity theorist is right to claim that our attitudes and emotions say quite a bit about the kind of person that we are, even in fictional contexts, and it is these responses, properly conceived, that should form the basis of our moral evaluation of video game play.⁴ However, as I suggested earlier, if we cannot say anything informative about the morally salient features that ought to inform our thinking in this realm, then this suggests that our moral unease is misplaced.

In an attempt to get clearer about the kinds of considerations that are morally salient in making such judgments, let us consider the following scenario: your friend asks you to play a video game called *Custer's Revenge*.⁵ The goal of *Custer's Revenge*, your friend explains, is to save another character that is held captive. You have nothing pressing at the moment, and you do not mind playing video games once in awhile, so you agree. Once you load up the game screen, however, you notice two things. First, in playing this game you take on the character General Custer who is under attack by a hail of arrows shot by native-Americans. You wonder about the representation of native-Americans in this game, but this issue is quickly overshadowed by another feature of representation: the character that you are to save is a native-American woman who is naked and tied to a pole. At this point, I think that more than a few readers would refuse to play *Custer's Revenge* because of its obvious racial, and gender insensitivity. Further, many of us would be shocked that our friend would think that we would enjoy this activity. But imagine, for the sake of argument, that you agree to continue playing. After all, it is only a game, and you do not really want to get into it with your friend. However, when you successfully navigate Custer through the onslaught of ar-

3. *Mia Consalvo* (2005) argues that we cannot simply bring our intuitions about what would be right and wrong in the actual world directly to bear on the world of games, though her focus is on a different phenomena: in-game cheating.

4. *Miguel Sicart* (2005) seems to make a similar claim when he argues that we must focus on the experience of game play.

5. *Custer's Revenge* is an actual video game that was released in 1982 for Atari by Mystique, a company that produced a number of video games with graphic sexual content.

row attacks and reach the native-American woman, you find that part of your/Custer's reward is to "rape" the native-American woman.⁶

I suspect that most of us would be shocked and dismayed to find ourselves put in such a position, especially by someone whom we take to be our friend. It would be hard to even think of an adequate justification of our friend's failure to tell us about the offending representational aspects of the game, since this information seems of obvious significance for determining if we would decide to play or not. Moreover, if upon asking our friend why he did not tell us, he responded that it did not seem important, I think that we would find ourselves perplexed. How could someone fail to see the moral problem involved, especially someone we call our friend?

What should we do in such a situation? One concern is we to respond to our friend's apparent endorsement of Custer's Revenge. Consider, for example, the ways in which a group of individuals, especially juvenile boys, might play a game like Custer's Revenge. We can imagine that they say things about the female target of the game, in a certain tone of voice, with accompanying gestures that suggest that they have sexist attitudes, even if they would explicitly decry such attitudes in calmer moments. Clearly, such game play behaviors can operate to reinforce sexist conceptions of women and racist conceptions of native-Americans. Peer communication is a powerful educator, and challenges from our peers often count for more than challenges from those who are outside our peer group. To play along, or to say nothing runs a certain kind of communicative risk, so we must ask ourselves if playing the game without criticism would send the wrong message to our friend. Of course, none of this is to say that we should say anything at all. We might find that in some cases we should withhold our moral disapproval all-things-considered. My only point here is that these kinds of considerations bear on what kinds of games we should play, how we should play them, and how we should respond to the play of others. Still, even if we determine that in this particular case we should say something, this does not tell what to say. If our friend responded to our complaint by claiming "hey, come on, it's only a game. You know that I am not a racist or sexist." we ought to be able to justify our contention that we should not enjoy representations like those even if we do not hold the offending attitudes. Moreover, our justification must go beyond the mere pointing out that the game contains representations of wrongful acts, since many games that are not morally troubling contain representations of wrongful acts. Further, it must go beyond pointing out that in playing Custer's Revenge it

6. Strictly speaking, in the actual game Custer simulates sexual intercourse with the native-American woman while she is still tied to the pole, though I doubt many will have difficulty conceiving of this as a depiction of rape.

is we who enact the wrong represented, since many games that are not morally troubling invite us to do this. What we must know is what is it about this video game that makes its representations wrongful, while others only represent wrong acts? This is the crux of the matter: on what grounds can we say that it is permissible to enjoy some representations of immorality, but not others?

In order to help shed some light on this distinction, let us consider the case of first-person shooters. First-person shooters often contain representations of wrongful acts in so far as they present us with representations of unjustified killings. Moreover, it is the player who enacts such represented wrongs insofar as she unjustifiably kills other characters. Once we fully come to terms with the fact that such games do not have any significant effect on those who play them, playing violent video games does not make one a worse person, I doubt that most of us would find the violent content that we find in the average first-person, fictionalized as it is, morally troubling. We might find first-person shooters juvenile, or in bad taste, or even boring, but not morally objectionable. So, what is it about representations, like those in *Custer's Revenge*, that give rise to such significant moral reactions? That is, what makes these wrongful representations, rather than mere representations of wrong?

One answer is that we are willing to ignore some representations of wrong for the sake of enjoyment, but when the represented wrong becomes especially egregious we should refuse to be entertained by them. That is, we should not enjoy *Custer's Revenge* because the represented immorality is so bad that we simply should not ignore it. It should block our ability to undertake the imaginative feat necessary to play the game: we should not see the representation as something that we can enjoy.⁷

This formulation has obvious appeal. It gets the phenomenology of our game play right, and it appears to avoid the problem of over-moralizing our game play. After all, on this analysis it is only when a video game represents egregiously immoralities that it commands our attention, while lesser immoralities need not rise to this level and so can be ignored. Still, despite its obvious virtues, this analysis undermines our intuitions in a wide-range of cases. Consider that in many first-person shooters the player kills countless numbers of foes, but that we enjoy playing such a game is not morally noteworthy all by itself. We would hardly think that the fact that a player shoots and kills her targets in a first-person shooter speaks ill of her character, any more than we would think that her courage under fire in the same game speaks well of her character. Similarly, we do not think that the mere fact that she enjoys these kinds of games says anything particularly morally interesting about her, though doubtless it says something interesting about

7. For an articulation of a threshold view in connection with art criticism, see Eaton (2003).

the kind of person that she is. But, given that murder is a relatively egregious crime, if moral evaluation in these contexts were determined simply by the level of wrongness involved, then we would expect first-person shooters to warrant our moral attention. Certainly, killing is at least as bad as having a stereotyped view of women or minorities.

Imaginative Representations, Default Social Meaning, and Ethical Assessment

To then what would a properly sensitive moral agent be sensitive in video game contexts? What makes a representation wrongful? Though I hope to make some progress in addressing these questions, I do not think that a single analysis, however complicated, will satisfactorily cover all cases. The phenomena involved are simply too varied. So, I do not intend to provide anything like a general principle or a decision procedure here, or anywhere for that matter. My aim is more modest. On my view, some representations contain details that anyone who has a proper understanding of our actual, contingent moral reality and is properly sensitive to this reality, will see as having a social meaning that raises the moral stakes and opens the door to associated character assessments. I will not argue that these are the only representational details that are salient, nor will I argue that we should not enjoy playing a game with such representational details all-things-considered. I intend only to point out some details that warrant our moral attention, and make moral inquiry pertinent in video game contexts.

Let us begin this last section by trying to get a clearer picture of imaginative entertainings in general. It is undeniable that fictional representations can be more or less fictive or imaginative. For example, a fictional representation can involve an actual person, say Barack Obama. The logic of such a fictional representation is as follows: it is fictional that the actual person Barak Obama has certain properties. In such a case, we recognize that Obama is not a wholly fictional character, but that we are asked to fictionally attribute certain qualities to him. That is, we are not asked to accept that it is fictional that there is a person named Barack Obama who has certain properties (though, of course, we may when the character just happens to be named Barack Obama); we are asked to imagine him as other than he is. That we are asked to make fictional attributions to real persons places us in the following interpretative position. We are expected to bring our ordinary knowledge about Obama to bear on our interpretation of the representation, and when the author presents details that do not square with what we know about him, we are to imagine that he is other than he is. Sometimes, these imaginative activities are harmless. Say, when we are asked to imagine that he eats muffins for breakfast, or that he travels back in time to stop global climate change. But, not all imaginative attributions to real persons are harmless. As we

all know, a fictionalized representation can serve to communicate a message about a non-fictional subject. As a result, they can be used as subtle or not so subtle forms of critique, resistance, insult, or even slander. The image of Obama going back in time to stop global warming might in some contexts serve as a critique of his “god-like image.” I take this point to be largely obvious, and uncontroversial. Moreover, insofar as imaginative imagery makes a comment about an actual person, it can be evaluated in terms of its aesthetic achievement, and in terms of its accuracy. I take this point to be similarly uncontroversial. In some contexts, we might find that a representation’s accuracy, moral or otherwise, bears on its aesthetic achievements, say when we laugh at a political cartoon because “it is so true,” or when we refuse to laugh because “it is beyond the pale of human decency.” In these latter kinds of cases, part of what bothers us is that representation invites us to take morally unwarranted attitude toward an actual person via the vehicle of the imagination. Consider, for example a cartoon image of Obama as a monkey, or as being lynched, or as eating watermelon. Without further representational detail to direct our interpretation of each of these images, they should be interpreted as an insult to Obama that is “beyond the pale of human decency.”

Of course, video game imagery rarely makes fictional attributions to real individuals; they mostly invent new worlds wholesale. This might lead us to believe that video game imagery is to this extent immune from moral criticism. That is, the fact that video game imagery invites us to make fictional attributions to fictional people, and not to make fictional attributions to real people, renders moral evaluation of our imaginative activities out of bounds. Though this analysis coheres with some of our intuitions, we should be careful about settling for such an easy gloss of a complicated phenomena. In particular, there are some imaginative representations that do not pick out a single, existing individual, but do pick out a class of individuals, regardless of the author’s or designers intentions. It is in light of this that we can raise moral concerns about such representations, and make associated character judgments of players. This is so because some otherwise fictional representations have obvious social meanings that are partly fixed by moral facts on the ground. To help understand this point, let’s start with a truism: Audiences come to fictions with a host of background assumptions. For example, we bring our knowledge of cause and effect, and our knowledge of folk psychology. Further, proper interpretation of fictions requires that we be epistemically flexible, so that when a fiction does not cohere with our knowledge of cause and effect, we should be prepared to accept it as true of the fiction that cause and effect work differently in the fictional world, as is the case with many sci-fi narratives. It is important to note, however, that what is called for here is epistemic flexibility not complete epistemic openness. We might reject a fictional work because it does not cohere with our assumptions about folk psychol-

ogy, when, for example, we complain that “no real person would act like that.” So, part of our job as audience members is to determine when we should accept something as true of the fiction, and when we should reject it because it fails to cohere with our knowledge of the actual world. So far, so good. However, we also bring to fictions our knowledge of what I call the default social meaning of representations. The kind of social meaning that I have in mind here is contingent on the particular details of our shared, social reality. These details are contingent in that they could have been otherwise, and they are often, though not always, socially local in that their meaning is fixed by something like shared cultural experiences. How to mark the boundaries of these social realities is a difficult issue that I will not take up here. In many cases, nation-states provide a clear enough demarcation of a shared social reality. Citizens of the United States, for example, share a common history that fixes the meaning of certain representations so that they have a default meaning that is particularly incorrigible, and plays a fairly determinative role in their interpretation. This history is one that involves enslaving Africans, and a subsequent history of the racial oppression of African-Americans, and it has played out in a unique way in the United States, so that representations of Obama as a monkey, as eating watermelon, or as being lynched have a particular meaning that might only makes sense in the context of the contemporary United States. This history includes stereotyping African-Americans by representing as simian, as stupid, and as proper objects of violence. The particular insult that is raised by any of these images is culturally specific, culturally contingent, and morally offensive, though in another cultural context such images may be incomprehensible, or even a compliment, as it likely the case with the image of Obama eating watermelon.

Moreover, the target of these images is not only a particular African-American, say Obama, but African-Americans in general, so that the image of Barack Obama eating watermelon is a slight against all African-Americans, insofar as they are also members of the targeted group. So, imaginative representations can target whole groups of individuals. Such imagery brings the imaginative world into contact with the actual world, so that the representation is properly interpreted as more of a piece with the actual world, and less of a piece with a purely imaginative one. If this is right, then authors and designers do not possess a “midas touch” of fictionality; they cannot make anything whatsoever true of a fiction by fiat. Some images have default social meanings that are fairly incorrigible that authors and designers have to take into account.

Acknowledging this provides a framework for thinking about the ethical assessment of video game imagery. Consider, for example, a first-person shooter that represents the enemy as African-American. Here I do not mean that a single member of the enemy group is represented as African-American. Instead, imag-

ine that your enemy is a group of individuals represented as African-American, so that as a player you are called upon to target and kill only or primarily those who are represented as African-American. In the context of the contemporary United States, imagery like this has a particular social meaning that is fixed in part by the current and the historical legacy of racial oppression. Being African-American involves being a member of a targeted group, and this partly fixes the meaning of representations of African-Americans in video games. Playing a game that asks its audience to target African-Americans involves them in an activity that should be interpreted as targeting and killing members of an oppressed group. It is hard to imagine anyone who is a citizen of the United States, and has adequate knowledge of the particular history of racial oppression as it has played out in the United States to interpret such representations as just fictions, and to thereby see it as "just a game." Such a game represent our actual shared, moral reality; so that attacking the representations of African-Americans seems implausible as a mere fantasy activity, but instead is seen as a representation of attacking actual African-Americans. That is, a representation like has a default meaning fixed by real-world, socially local moral facts. In light of this, the imaginative activity that we are called upon to perform here is one that ought to invoke this default meaning; a meaning that morally constrains our actual actions and attitudes, and including our imaginative actions and attitudes. This is what opens gamers to associated characterological judgments. That we enjoy the fictional activity of targeting and shooting representations of African-Americans, or that we enjoy watching our friends do so, is to open ourselves to characterological criticism.

Of course, our enjoyment, all by itself, does not tell us exactly what has gone wrong. It might be indicative of a number of failings. It might be that we simply do not know enough about our own history of racial oppression, and as a result we fail to see the default social meaning of these images. It might be that we have adequate knowledge of the relevant history, but fail to see social meanings as default, so that we lack a sensitivity to the meaning of this imagery and as a result, fail recognize the limitations of fictionality. Finally, it might be that our enjoyment is an expression of unsavory attitudes toward African-Americans, however inchoate. The first kind of failing is doubtless an epistemic failing, and might be a moral failing assuming we can make the case for the claim that we have a moral duty to know certain social facts, though it is not the kind of failing that I am interested in here. The third kind of failing is obviously a moral failing, but again it is not the kind of failing the I am interested in. It is the second kind of failing that interests me, the gamer who does not have unsavory moral attitudes, but whose enjoyment exhibits a lack of sensitivity, since this seems to be one of the main sources of contention between those who accept and those who reject the moral evaluation of video game play.

Consider the recent criticisms raised against the game Resident Evil 5 (Brophy-Warren, 2009; Jones, 2009). Of particular worry here is that the game's main character, Chris, is portrayed as a white westerner who must kill scores of African zombies. The image of "Chris unloading his pistol into hordes of African zombies" (ibid., 2009) has a certain meaning that is contingent on the actual history of the colonialization of Africa by western countries. So, an image of a "white man shooting black Africans" (ibid., 2009) ought to at least bring to mind this troubling history. Moreover, the fact that the zombies in this game are almost exclusively portrayed as African raises another worry. The history of racial oppression of those of African descent has involved stereotypes of such individuals as subhuman, which lends the images in a game that portrays most of its sub-human characters, the zombies, as African a default meaning that is fixed by this history. As a result, some report that playing Resident Evil 5 puts the gamer in the position of enjoying a representation that seems too close to reality to be enjoyable. Just as the meaning of the image of the Obama in a monkey suit is fixed by the history of racial oppression in the United States, the meaning of images of Africans as subhuman targets is fixed by a more global history of colonialization and racial oppression.

Still, one might object that in Resident Evil 5 the default meaning the offending imagery has been adequately undermined by the game's designers insofar as there is a perfectly good reason for representing the zombies as African. After all, the game is set in Africa, so it should come as no surprise that many of the game's characters are African. Moreover, one might continue, the goal of the game is not to kill as many African zombies as you can, it is to save African villages, the continent of Africa, and even the world from a bio-terrorist group that has used a virus to turn Africans into zombies. Since the mission itself is a noble one, one that is in the interest of most Africans, and since the action takes place in Africa, and since the real enemies are not the zombies but the bio-terrorists, the otherwise problematic imagery is taken into account, and rendered morally neutral. As a result, it is reasonable to claim that the game designers have taken some measures to undermine the default meaning of images of white western males killing sub-human Africans. Doubtless representations of African zombies could be employed in ways that are far less sensitive to the default meaning of such imagery. Still, given the history of racial oppression and colonialism perpetrated by western countries against African nations, the depiction of characters as being of African descent, and sub-human, brings to mind this actual history which involved in no small part the dehumanization of those of African descent. Note I am not claiming that this imagery harms those of African descent, though I think one might be able to make such a case. My claim is only that it is in light of the histories that some images have a default meaning that game designers and gam-

ers alike have a duty to take seriously, and it is at least an open question whether or not the designers of Resident Evil 5 have done so adequately.

Talk of default social meanings also helps explain our moral concern about Custer's Revenge, and Rape the Princess. The imagery of these games is worrisome, not because they are representations of immoral or unjust actions, but because it appears to implicate the history of the oppression of actual woman. Both games invite us to enjoy representations of women as objects of sexual violence in a way that does not undermine the default social meaning of these images. In the actual world, women are the victims of a kind of systematic oppression that involves seeing them in certain kinds of ways, conceiving of them as certain kinds of creatures, and subjecting them certain kinds of violence. Given these facts about our shared moral reality, it does not take much to see the representation of the female characters in these games as an extension of this real-world moral phenomena (whatever the avowed intentions of the game designers) in the same way that it is reasonable to interpret Obama in a monkey suit as an extension of real-world racial phenomena. Again, the worry is that representations are in some important sense not fictions: otherwise why select women and not men to play these particular representational roles? So, the question we should ask ourselves is: does it make more sense to interpret this image as a representation of actual women, or does it make more sense to interpret it as a fictional representation of fictional women? In the case of video game representations of women, part of what makes it more reasonable to see such imagery as representations of actual women is that such images are nearly universally sexualized. While game designers take great pains to avoid racist imagery, they shamelessly employ gendered imagery with the full realization that the imagery has certain meaning: that women are properly thought of as objects of sexual pleasure. This is part of what bothers us about that prostitute and the hot coffee scenarios in Grand Theft Auto: San Andreas, and even the general sexualized representations of female characters in most video games.

Still, how do we determine if an image has the kind of incorrigible, default social meaning that is relevant? Though I don't hold out much hope for anything like a decision procedure here, it is worth pointing out that the kind of contextual details that are in play in each of the cases that I have mentioned here are details about egregious, long term, systematic denial of justice that are of a particular kind; oppression. Oppression is especially insidious because it denies individuals the respect that is due to one qua human. To be accorded access to resources for seeing oneself as fully human, rather than as subhuman is necessary to live a fully human life. Moreover, this is a need that, in the words of David Wiggins "cannot be satisfied by one's own efforts" (Wiggins, 1998). Satisfaction here necessarily involves others, since it is others who deny them this dignity, it is others

who must accord them this dignity. The fact that it is our shared social history, and it is we who have collectively (in these cases white Americans, westerners, or males as the case may be) who have denied individuals full access to such resources because of arbitrary group membership is what makes the images cited here particularly incorrigible. So that, a friend who responds to our criticisms of *Custer's Revenge* by claiming, "come on, it's only a game; I'm not sexist" sees his imagining as just some random imagery detached from his own moral commitments, and detached from the moral facts on the ground. This failure is a failure of sensitivity and sympathy. In part, sympathy requires is us to inhabit the perspective of others, and to fail to see these representations as morally horrific is to fail to adequately exhibit a properly sympathetic response in relation to a feature of current life that especially calls for sympathy. Sympathy in this context without a proper understanding of the social relevance of these images is no sympathy at all. Moreover, to insist that one's imagination is one's own private affair, detached as from one's own actual commitments and similarly detached from the contextual moral facts on the ground, amounts, in this case, to a snubbing of one's nose at a requirement of solidarity with the victims of oppression.

Generally, the representational violence found in first-person shooters is not represented in a way that implicates our shared, moral reality. We are often shooting aliens, or spies, or members of a competing faction. Whatever we are doing in these games is sufficiently insulated from our shared, moral realities so as to make the challenge "Come on, it's only a game!" a credible one. It is this insulation from the real world that makes the kind of associated character evaluations like "See how she has defended the citizens against the aggressive attack of the space aliens; she is certainly a good person!" misplaced. When morally challenging representational content reflects our actual, shared history of systematic violations of justice like gender oppression, this serves to fix the meaning of such imagery, and thereby opens the door for associated character evaluations. One who enjoys playing games like *Custer's Revenge*, or *Rape the Princess* lacks the appropriate sensitivity to sexism; just as one who cannot see the worries raised against *Resident Evil 5* as salient does. This lack of sensitivity might be a failure to see the relationship between the game's representation of members of an oppressed class and the actual oppression of such individuals, or it might be a failure to see such a connection as a reason to avoid enjoying such a representation. However we make sense of the particular failing of a particular player, the evaluative difference between the run of the mill first-person shooter, and games like *Custer's Revenge*, *Rape the Princess*, *Resident Evil 5*, and *Grand Theft Auto: San Andreas* is best explained, not by the level of the wrong that is represented, again murder seems to be an egregious wrong, but by the kind of representational wrong that the game invites us to commit or enjoy.

Still, this does not address the issue of virtual-pedophilia. How can talk about default meanings of representations make sense of our moral intuitions in a case like this? As I said at the outset, the default meanings of representations are often, though not always local. Consider, for example that the imagery of women in video games is likely not local, but global. However, the default social meaning of images of women is contingent on the actual, global history of women's oppression. The meaning of images of sexualized images of children may not be local, or contingent in this way, since it seems clear that anyone who enjoys the thought of having sex with children exposes a flaw in her character. Still, I think that there is much more to say about this kind of imagery, especially in light of the practice of Japanese anime. That is, I think that the meaning of this imagery is sufficiently complicated to warrant its own investigation, an investigation that goes beyond the aim of the present inquiry.

REFERENCES

Brophy-Warren, Jamin (2009) 'Resident Evil 5' Reignites Debate About Race in Videogames. Wall Street Journal. 3/12/2009. <http://online.wsj.com/article/SB123672060500987853.html>. Accessed 30 April 2009.

Consalvo, Mia (2005) Rule Sets, Cheating, and Magic Circles: Studying Games and Ethics. *International Review of Ethics* 4: 7–12.

Eaton, Anne (2003) Where Ethics and Aesthetics Meet: Titian's Rape of Europa. *Hypatia*, 18: 159–188.

Freedman, Jonathan (2007) Evaluating the Research on Violent Video Games. <http://culturalpolicy.uchicago.edu/conf2001/papers/freedman.html>. Accessed 10 November 2008.

Grossman, David (2000) Teaching Kids to Kill. *National Forum* 80: 10.

Gaut, Berys (2002) The Ethical Criticism of Art. In *Aesthetics and Ethics*, ed. Jerrold Levinson. New York: Cambridge University Press.

Jones, Scott, Resident Evil 5 (Xbox 360): They're coming to get you, Barbara. The Crispy Gamer. <http://www.crispygamer.com/gamereviews/2009-03-12/resident-evil-5-xbox-360.aspx>. Accessed 30 April 2009.

Lenhart, Amanda et al. (2007) Teens, Video Games, and Civics: Teens' gaming experiences are diverse and include significant social interaction and civic engagement. http://www.pewinternet.org/pdfs/PIP_Teens_Games_and_Civics_Report_FINAL.pdf. Accessed 31 October 2008.

McCormick Matt (2001) Is It Wrong to Play Violent Video Games? *Ethics and Information Technology*. 3: 277-287.

Patridge, Stephanie (2008) Monstrous Thoughts and the Moral Identity Thesis. *The Journal of Value Inquiry*. 36: 181–193.

Sicart Miguel (2005) Game, Player, Ethics: A Virtue Ethics Approach to Video Games. *International Review of Ethics* 4: 13–18.

Warner, Dorthy & Raiter, Mike (2005) Social Context in Massively-Multiplayer Online Games (MMOGs): Ethical Questions in Shared Space. *International Review of Ethics* 4: 46–52.

Walton, Kendall (1997) Spelunking, Simulation, and Slime: On Being Moved By Fiction. In *Emotion and the Arts*, eds. Mette Hjort and Sue Laver. New York: Oxford University Press.

Wiggins, David (1998) *Needs, Values, and Truth*. Oxford University Press.

Wonderly, Monique, (2008). A Humean Approach to Assessing the Moral Significance of Ultra-violent Video Games. *Ethics and Information Technology* 10/1–10/2007 Sales, Demographic and Data Usage: Essential Facts about the Video Game Industry. http://www.theesa.com/facts/pdfs/ESA_EF_2007.pdf. Accessed 31 October 2008.

Ethical Issues in the Design and Use of Online Career Development Counseling Services

Adamantia Pateli*

Lecturer, Department of Informatics,
Ionian University

Abstract

The Internet provides a new medium for interpersonal communication and information exchange that holds the potential for delivering forms of career counseling services that are appropriate to the medium. The challenge facing counselors lie in discovering what type of counseling services are appropriate to this new medium and in delineating the potential advantages and limitations inherent to this new communication format. This paper discusses the ethical issues which may impact the way in which career counselors design, and counselees' access and use the Internet. Recommendations are offered for improving the effectiveness and acceptance of Internet-based career counseling services. The ultimate purpose is that career consulting firms, human resource managers and employers as well as policy makers become aware of the issues involved and what they can do to maximize the potential benefits and minimize the potential problems associated with online career counseling services.

Keywords: career development, online counseling, ethical challenges, recommendations

Introduction

Career development and counseling, also called "career guidance" in several European countries, has emerged as an international phenomenon providing peo-

* *Adamantia Pateli* holds the position of Lecturer of Information Systems in the Department of Informatics of the Ionian University, in Greece. She has eleven years research experience in the areas of e-commerce and mobile business through her participation in more than fifteen national and European-funded projects. Adamantia has published more than twenty five research articles in several European and international journals, such as *European Journal of Information Systems*, *International Journal of Technology Management*, *Management Decision*, and conferences, such as *Hawaii International Conference on System Sciences*, *Bled Electronic Commerce Conference*, and *International Conference on Mobile Business*. Her current research interests lie in the areas of electronic Human Resource Management applications, business networks and open innovation.

ple with opportunities to develop their talents, make career choices that improve their life quality and also contribute to the improvement of the society (ACES/NCDA, 2000). For some people, the focus of career development is almost entirely on fitting into work and jobs, while others define career more broadly to also include other life roles. One of the oldest definitions of career development, provided by Super (1980), refers to it as “a lifelong, continuous process of developing and implementing a self concept, testing it against reality, with satisfaction to self and benefit to society”. Herr (2000) goes further in conceptualizing career guidance and counseling as “instruments of personal flexibility and human dignity”. In this paper, we have adopted the newer definition provided by NBCC (2007), according to which the term “career counseling” is used to denote “the application of mental health, psychological or human development principles, through cognitive, affective behavioural or systemic intervention strategies that address career development”.

Technological advances in career intervention, focusing on computer-assisted guidance and the ubiquitous involvement of the Internet in career information dissemination, have called for enhanced sophistication in instruction and supervised practice. As result, a new term, “online career counseling”, has emerged to denote asynchronous and synchronous distance interaction among career counselors and counselees using one or more of the following communication features of the Internet; e-mail, chat, teleconferencing, videoconferencing, and social web. Online counselors may involve Human Resources (HR) consulting firms, Human Resources managers of the company that has employed or is going to employ the counselee, as well as other managers operating as external or internal coaches/ mentors of the employees. Online counselees include current employees, usually pursuing career development either inside or outside of their current organizational boundaries, and candidate employees, pursuing a career that matches as much as possible their interests and competence.

There is much debate among researchers and practitioners of career development services on the opportunities and risks incurred from the use of online counseling practices. The proponents of online counseling argue in favor of them referring to several advantages. First, online counseling services are extremely convenient, since most of the people do not have the time to commit to a pre-scheduled personal face-to-face session. Disabled people who may not be able to leave their home could also benefit from these services. Moreover, online consultation is much less expensive than seeing a counselor face-to-face. Unlike the face-to-face sessions, online consultation gives you the opportunity to reflect on the advice you receive, to think about it and ask further questions. Finally, online counseling provides anonymity, and thus counselees become more willing to disclose their private and personal information, which is a vital element for any

successful counseling. Although these advantages appear to support counseling via the Internet, there are ethical as well as practical difficulties in establishing effectively these services (Robson & Robson, 2000). The primary limitation of online counseling services includes the loss of non-verbal social cues that provide valuable contextual information in conversation and can influence the interpretation of meaning in communication (Childress, 2000). Moreover, some counseling services delivered over the Internet are originally designed to be used within the traditional environment. If they are transferred as such in the online environment, then a number of problems may arise. The expected outcomes cannot be equivalent in the off-line and online cases, and the interpretation provided by the Internet version cannot be equivalent in validity to the interpretation offered by a counselor.

This paper explores some of the ethical and practical considerations surrounding career development counseling via the Internet. The next section looks at the primary ethical issues raised and the following section presents guidelines provided by official certification bodies as well as recommendations resulting from the information systems and human resources management disciplines.

Online career development counseling services

Typical services in the category of online career counseling include; e-mentoring, e-coaching and e-assessment. Both coaching and mentoring are forms of another paradigm for learning that has its ancient roots in dialogue and interaction (Masullo & Tsangtis, 2004).

e-Mentoring

E-mentoring is a recent development in the field of mentoring and one which is already showing significant advantages over more traditional forms. An e-mentor is not recognized as a tutor or teacher but someone who provides guidance or counsel (Homitz & Berge, 2008). Traditionally, mentoring was defined as a dyadic relationship in which a mentor, a senior person in age or experience, provided guidance and support to the less experienced person, the protégé (Hunt & Michael, 1983). In general, mentors perform three major functions for their protégés. First, they provide vocational or instrumental support that directly enhances the career of the protégé and might include sponsorship, visibility, protection, and challenging assignments. Second, they may provide psychosocial support via counseling, friendship, and encouragement. Third, mentors may also function as role models to their protégés by demonstrating appropriate behavior either implicitly or explicitly (Ensher et al., 2003).

e-Coaching

Business coaches may perform many of the same functions as a mentor, yet often the relationship is more temporal or goal-based (Whitaker, 2001). The exchanges between online business coaches and their clients are often explicit, monetary, and transactional in nature. In contrast, the exchange between a mentor who provides a protégé with coaching is often more implicit, personal, and reciprocal (Ensher et al., 2003). Fee-based coaching offers clients real-world advice, often from people who have been in the same situation and can speak from personal experience (Harrington, 1998). Clients practice and develop skills in résumé writing, business plan development, and contract negotiation by sending their efforts to a coach who provides feedback. Today, most coaches apply traditional means (i.e. email, phone) but also advanced electronic means (i.e. instant messaging, teleconferencing, videoconferencing) to communicate with their clients, using the Internet strategically.

e-Assessment

Practitioners have continued to innovate over the past few years by offering remote self-help performance assessments to the employees and remote interpretation of assessments as one component of distance career counseling (Sampson & Lumsden, 2000). The Internet-based assessment applications may be used either isolated or to supplement a company's existing performance management systems. The e-assessment services examined within this paper include, apart from online performance appraisal forms provided by a typical e-performance management system (Evans, 2001), analysis of full-circle (360o) feedback and counseling on the type of learning programs that best suit the counselees' career development needs. An online assessment system may also allow users to track their own progress over a series of evaluations, and thus get even more targeted counseling services. This option works as an attractive means for employees to bridge the gap between feedback and development planning (Cardy & Miller, 2005).

Ethical Challenges

The following ethical issues can be categorised into those that are associated with the design and those associated with the use of online career counseling services. In terms of design, features such as reliability and validity, and counselee-oriented interface design of the services are examined. In terms of use, computer literacy of counselee, intimacy of online communication, competence of online counselor, and equity of access are discussed. Finally, in terms of both design and use, issues regarding the confidentiality and privacy of information exchanged are examined.

Reliability and Validity of Service

Oliver and Zack (1999) consider it “unethical for career counseling practitioners to use online instruments without investigating the psychometric characteristics of the measures or the basis of interpretation”. Such a notice applies not only for e-assessment tools, but also for e-mentoring and e-coaching services, where the need for psychometric data is high. Since e-mentoring and e-coaching services are provided on a personalised basis, the collection of psychometric data on the counselee has often a lead over the career counseling process.

Counselee-Oriented Interface Design

Psychologists should keep an active role in the design of online career counseling services, since they have better knowledge of the features that may influence positively the counselee’s attitude towards the use of such services. For instance, practitioners of psychometric tests have set a number of user interface characteristics that every e-assessment system should have. The most important include consistency, recoverability and control over the system’s data flow, as well as extended user guidance and assistance.

Computer Literacy of Counselee

For individuals who have low computer literacy, using an online career counseling tool on the Internet may create some difficulties, as it is often delivered as a self-help intervention. There may be no avenue for the individual to ask questions, clarify directions, obtain support, and otherwise get personal guidance. It would not be uncommon for individuals with low computer literacy, while using an online career counseling tool, to encounter some difficulties in answering questions and to feel frustrated and confused.

Intimacy of Online Communication

The primary disadvantage of using the Internet for counseling includes the loss of the non-verbal communication level (Robson & Robson, 2000). The majority of communication between a counselor and a counselee is taking place unconsciously by body language and paralinguistics, such as tone of voice (Argyle, 1983). Using computer communication runs the risk that the ‘space between the two parties’ becomes filled with hardware.

Competence of Online Counselor

Practitioners need to develop competencies in (a) identifying the unique Internet features that could guide the design of qualitative and effective career counseling services, (b) recognizing potential ethical issues associated with the use of these mechanisms and taking appropriate action; and (c) receiving or providing remote supervision and support in the use of online career services. Also, practitioners interested in developing online career counseling services need to be equipped with instructional design competencies as well as web site design skills.

Equity of Access

There is a concern that access to Internet-based services is not equitable. Clients with limited financial resources may have difficulty gaining access to the Internet (Sampson & Lumsden, 2000). More affluent individuals often have Internet access at their place of residence and/or their place of work, therefore resulting in greater access to career resources and services that are available on the Internet. Limited Internet access could increase the income disparity between well-educated, affluent groups and less-educated, less-affluent groups in society (Harris-Bowlsbey et al., 1998).

Confidentiality and Privacy

Counselors working with clients through Internet should be aware of the potential threats to confidentiality and privacy of counselee's personal or career-related data and use appropriate security methods for their online transmission and storage mechanisms (Sampson et al., 1997). It is imperative that appropriate security methods are used to hamper any malicious action. Moreover, it is important that clients of these services, current or candidate employees, are informed of the privacy policies followed by the counselor, so that they are able to determine the risks involved in the delivery of online counseling services.

Recommendations for addressing ethical issues

Recently, there have been a number of nationwide initiatives to develop principles for guiding the evolving practice of Internet counseling. The most important one is the Practice of Internet Consulting developed by the National Board for Certified Counselors (NBCC). In addition to the NBCC (2007) work, other national professional bodies, such as the American Counseling Association (ACA, 2005) and the British Association for Counseling and Psychotherapy (Antony & Goss, 2009) have also incorporated standards for the Internet counseling within their Code of Ethics. While the above initiatives concern any type of online coun-

selling services, the NCDA (1997) has developed guidelines for the design and use of Internet-based career counseling services, which are examined herein.

Following, we present the main concerns of these Codes of Ethics that address the ethical issues that emerge for the Online Career Development Counseling services. Moreover, based on research conducted not only in the web counseling area (e.g. Bloom, 1998; Robson & Robson, 2000) but also in the human resource management area (e.g. Walker & Perrin, 2001; Gueutal & Stone, 2005; Philips et al., 2008), we provide several recommendations for counseling organizations and counselors regarding policies and procedures that may be effective in reducing the risk associated with the above discussed ethical issues.

Recommendations for the Design of Online Career Counseling Services

Regarding the reliability and validity of online career counseling services, the following recommendations are provided for online counselors:

- Discuss the possibility of technology failure and provide alternate methods of service delivery, if that failure occurs;
- Use and recommend sites that have been investigated and which are considered as appropriate to the employees' needs;
- Regularly check that electronic links provided are working and are professionally appropriate;
- Review the content of a website or other service offering career services to determine its appropriateness to the medium. When a website offers a service which has not previously been extensively tested, this service should be carefully scrutinized to determine whether it lends itself to the Internet. More specifically, in case of e-assessment, the assessment tests must have been tested in computer delivery mode to assure that their psychometric properties are the same in this mode of delivery as in print form. If not, then the client must be informed that they have not yet been tested in this same mode of delivery;
- Finally, to increase reliability and validity of online career development services, utilize the online counseling services in conjunction with traditional systems (e.g. online assessment and face-to-face test administration).

Regarding the interface design of online counseling services, the following recommendations are provided:

- It should be such that the services can be utilized by all employees without extensive training, including those having minimal computer skills, experience and education;
- Information and services should be presented in an appealing and friendly manner. To achieve that, graphics and other multimedia should be added.

Recommendations for the Use of Online Career Counseling Services

Not all employees can benefit the same from the online counseling services. The services' effectiveness and employees' satisfaction may vary a great deal depending on their competence in using computers and readiness to decision making (Sampson & Lumsden, 2000). Online career development counseling services are usually offered in self-help mode. However, some individuals may not benefit from this mode of service delivery. Instead, an individual with high computer literacy and high readiness for decision-making is more likely to benefit from self-help interventions. To increase usefulness of the online career counseling services, the following recommendations are provided for online counselors.

- Make the clients aware of the typical circumstances where they may need counseling support in order to effectively use the career information provided online;
- Host advanced help features in career counseling sites;
- Create the profiles of the targeted clients, in terms of competence and knowledge in using career development tools and online services, before releasing the online career services;
- Ask a clear statement by clients of their career planning or career counseling needs;
- Conduct an analysis whether meeting those needs via Internet exchange is appropriate and of whether this particular client can benefit from counseling services provided in this mode. A judgment about the latter should be made by means of a telephone or videophone teleconference designed to specify the client's expectations, how the client has sought to meet these through other modes, and whether or not the client appears to be able to process information through an Internet medium;
- Inform the clients about the level of profit that they can gain from the provided online counseling services, and indicate means with which they can increase this level (e.g. getting training on computers, asking the help of an expert, e.tc.).

Hanley's research findings (2004) have suggested that individuals who are competent at using computer-mediated communication (CMC) means can develop relationships of an appropriate quality to receive counseling online. Practitioners need to be aware of the relationship's co-constructed nature, particularly in heightening their sensitivity to the 'reality' and transference elements to it. However, it does not appear necessary to develop new theories of counseling to cater for online practices. Instead, the following recommendations are provided for online counsellors in order to increase intimacy provided by the online medium.

- Define several items in writing to the client in a document that can be downloaded from the Internet or faxed to the client. This document should include at least the following items: (a) The counselor's credentials in the field, (b) the agreed-upon goals of the career counseling Internet interchange, (c) the agreed-upon cost of the services and how this will be billed, and (d) where and how clients can report any counselor behavior which they consider to be unethical;
- As part of the counseling orientation process, explain the clients how to cope with potential misunderstandings when visual cues do not exist.
- Explain the clients the procedures for contacting online counselors when they are off-line and, in the case of asynchronous counseling, how often e-mail messages will be checked by them;
- Should the counselors determine that little or no progress is being made toward their client's goals, they have the ethical responsibility for the periodic monitoring of the client's progress via telephone or teleconference or face-to-face discussion;
- In the case of e-assessment, if there is any evidence that the client does not understand the results, as evidenced by e-mail or telephone interchanges, then refer the client to a qualified career counselor in his or her geographic area.

While the computer literacy of counselees is desired to increase effectiveness of online career counseling services, the computer literacy as well as experience of online counselor in using computer-mediated communication (CMC) means for providing its advice and counseling support is of utmost importance. For this reason, a set of recommendations are provided for HR departments and career counseling organizations that wish to provide such services.

- Employ HR specialists and counselors who are knowledgeable regarding the CMC services, so that they can assist employees with access or use of online counseling services. These specialists should also be able to handle technological exceptions and special circumstances as they arise;

- Online counseling services should clearly state the qualifications and credentials of the developers not only in the content area of professional career counseling, but also in the development of interactive online services.

Finally, in order to assure equity of access to and use of online career counseling services, the following recommendations are provided for online counsellors by both NBCC (2007) and ACA (2005) Codes of Ethics.

- Make clients aware of free public access points to the Internet within the community for accessing Internet counseling or Web-based assessment, information, and instructional resources;
- Within the limits of readily available technology, make the career web site a barrier-free environment to employees with disabilities;
- In case of multinational companies, be aware that some employees may live in different time zones, and have unique cultural perspectives. Thus, local conditions and events may impact the client's behaviour towards online counseling services;
- Provide translation facilities for clients that have a different primary language, and also address the imperfect nature of such translations.

Recommendations for Confidentiality in Online Career Counseling Services

Nevertheless, the most important issue underlying both the design and the use of online career counseling services concern the security methods deployed and the privacy policies established for assuring privacy and confidentiality of employees' data. Almost all the above-mentioned Codes of Ethics (NCDA, 1997; ACA, 2005; NBCC, 2007) include several confidentiality and privacy recommendations for online counselors. Such recommendations are also provided in the human resource management bibliography (Stone et al, 2006; Philips et al., 2008).

The ones corresponding to the services discussed herein include:

- Develop privacy policies related to the online counseling services and be sure that these policies are published to employees either through websites, intranets, e-mail, or newsletters;
- Establish a method for verifying the employee identity. In situations where it is difficult to verify the identity of the Internet client, steps should be taken to address impostor concerns, such as by using usernames and passwords;
- Use encrypted Web sites and e-mail communications to help ensure confidentiality when possible;

- When the use of encryption is not possible, counsellors should notify clients of the potential hazards of unsecured communication on the Internet and limit electronic transmissions to general communications that are not client specific;
- Inform clients if, how, why and how long session data as well as archival data are being preserved;
- Follow procedures regarding the release of information for sharing employee information with other electronic sources;
- Inform clients of all colleagues, employers as well as IT administrators that might have authorised or unauthorised access to electronic transmissions;
- Provide electronic links to relevant state licensure and professional certification boards to protect consumer rights and facilitate addressing ethical concerns.

Conclusions and future work

Since the use of the Internet is new for the delivery of career development and counseling services, it is mandatory that the career counseling profession gain experience with this medium and evaluate its effectiveness through targeted research. The capabilities of Internet delivery of such services expand rapidly as the use of sound and video becomes more feasible due to the availability of broadband networks. Nevertheless, the risks of providing counseling services via the Internet medium are still high, resulting in low rates of acceptance and growth for these services.

This paper includes an attempt to identify the main ethical issues associated with the design and use of online counseling services. The existing literature has provided guidelines and recommendations for online counseling services, usually examined in the context of a psychotherapeutic intervention, and career development services in isolation. This paper combines existing research and guidelines provided by a number of certification bodies in the counseling area with research conducted in the emerging electronic human resource management area, in order to define the primary challenges in the online career counseling domain. Building on this research body, it provides a set of recommendations for improving the effectiveness of these services and increasing employees' satisfaction from them. These early recommendations need constant monitoring and revision as research data become available and additional technological capabilities become cost-feasible.

Practitioners of online career counseling could apply these recommendations to check the quality of their current online counseling services or incorporate them

to the design or future online career services. Within this stream, further research could be directed towards delineating the herein discussed suggestions to develop a roadmap for the development and operation of online career counseling services by human resource departments and counselors. Such a roadmap would include, apart from proposals on measures and policies to be established, a set of advice for addressing problems and difficulties that may arise from either the counselor or the employee side, such as persuading employees on the necessity of e-coaching and e-mentoring practices or decreasing employee's mistrust on e-assessment's results.

REFERENCES

Association for Counselor Education and Supervision (ACES)/ National Career Development Association (NCDA) (2000) Preparing Counselors for Career Development in the New Millenium, Available online from: http://associationdatabase.com/aws/NCDA/asset_manager/get_file/3405, Retrieved on: 17th December 2008.

Americal Counseling Association (2005) ACA Code of Ethics, Available online from: www.counseling.org, Retrieved on: 20th November 2008.

Antony, K. & Goss, S. (2009) Guidelines for online counselling and psychotherapy, Third Edition. British Association for Counselling and Psychotherapy, Leicestershire.

Argyle, M. (1983) The psychology on interpersonal behaviour, Fourth Edition. Penguin, London.

Bloom, J.W. (1998) The ethical practice of WebCounseling. British Journal of Guidance & Counselling, 26(1), 53-59.

Cardy, R.L. & Miller, J.S. (2005) eHR and Performance Management: A Consideration of Positive Potential and the Dark Side, in *The Brave New World of eHR – Human Resources Management in the Digital Age* (eds. H.G. Gueutal and D.L. Stone), Wiley & Sons, San Francisco.

Childress, C.A. (2000) Ethical Issues in Providing Online Psychotherapeutic Interventions, *Journal of Medical Internet Research*, 2(1): e5.

Ensher, E.A., Heun, C., Blanchard, A. (2003) Online mentoring and computer-mediated communication: New directions in research. *Journal of Vocational Behavior*, 63, 264-288.

Evans, E.M. (2001) Internet-Age Performance Management – Lessons from High-Performing Organizations, in *Web-based Human Resources* (ed. A.J. Walker), Mc-Graw Hill, New York.

Gueutal, H.G. & Stone, D.L. (2005) *The Brave New World of eHR*. John Wiley & Sons, San Francisco.

Hanley, T. (2004) *Online Counselling: a heuristic study examining the relational depth of computer-mediated relationships*. A thesis submitted to the University of Manchester for the degree of Masters in the Faculty of Education, Centre for Counselling Research and Training, Manchester, UK.

Harrington, A. (1998) A sounding board in cyberspace. *Fortune*, 138(6), 301–302.

Harris-Bowlsbey, J., Dikel, M.R. & Sampson, J.P., Jr. (1998) *The Internet: A tool for career planning*, First Edition. National Career Development Association, Columbus, OH.

Herr, E.L. (2000) Perspectives on career development: The legace of the 20th century – The innovation of the 21st. *Proceedings of the NCDA National Conference*, Pittsburg, PA, June 23.

Homitz, D.J. & Berge, Z.L. (2008) Using e-mentoring to sustain distance training and education. *The Learning Organization*, 15(4), 326–335.

Hunt, D.M. & Michael, C. (1983) Mentorship: A career training and development tool. *Academy of Management Review*, 8(3), 475–485.

Masullo, M. & Tsangtis, L. (2004) *e-Mentoring As A New Paradigm for Learning*. *Proceedings of the 6th International Conference on Education*.

National Board for Certified Counselors and the Center for Credentialing and Education (2007) *The Practice of Internet Counseling*. Available online from: www.nbcc.org, Retrieved on: 20th November 2008.

National Career Development Association (1997) *NCDA guidelines for the use of the Internet for provision of career information and planning services*. Available online from: http://associationdatabase.com/aws/NCDA/pt/sp/guidelines_internet#2, Retrieved on: 20th November 2008.

Oliver, L.W. & Zack, J.S. (1999) Career assessment on the Internet: An exploratory study. *Journal of Career Assessment*, 7, 323–356.

Philips, T.N., Isenhour, L.C. & Stone, D. (2008) *The Potential for Privacy Violations in Electronic Human Resource Practices, in Technology, Outsourcing and Transforming HR* (eds. G. Martin, M. Reddington, H. Alexander), Elsevier, Oxford.

Robson, D. & Robson M. (2000) Ethical Issues in Internet Counselling. *Counselling Psychology Quarterly*, 13(3), 249–257.

Sampson, J.P., Jr., Kolodinsky, R.W., Greeno, B.P. (1997) Counseling on the information highway: Future possibilities and potential problems. *Journal of Counseling and Development*, 75, 203-212.

Sampson, J.P., Jr. & Lumsden, J.A. (2000) Ethical Issues in the Design and Use of Internet-based Career Assessment. *Journal of Career Assessment*, 8(1), 21-35.

Stone, D., Stone-Romero, E. & Lukaszewski, K. (2006) Factors affecting the acceptance and effectiveness of electronic human resource systems. *Human Resource Management Review*, 16, 229-244.

Super, D.E. (1980) A life-span, life-space approach to career development. *Journal of Vocational Behavior*, 16(3), 282-298.

Walker, A.J. & Perrin, T. (2001) *Web-Based Human Resources*. McGraw-Hill, New York.

Whitaker, B. (2001) When the work path turns, can a guide help? *New York Times*, 3.

Machine Self-Sacrifice

Carson J. Reynolds & Alvaro Cassinelli^{*}
University of Tokyo

Abstract

The concept of self-sacrifice as it relates to artificial entities is defined. Illustrative anecdotes drawn from computer science, robotics, and microprocessor architecture are then provided. Building upon this we will argue for the utility of self-sacrifice in existing biological phenomena such as kin altruism. We will conclude by making a counter-intuitive claim from the standpoint of information ethics: that information should have the capability to destroy itself.

Keywords: ethics, artificial moral agents, artificial life, artificial death, self-sacrifice

Self-harm, weaponization, suicide and self-sacrifice

Consider a machine that is capable of self-destruction. There are many genres of self-destruction with attendant causes, effects, harms, and benefits. Let us provide a simple taxonomy. Of self-destructing machines physical harm that is

^{*} *Carson Reynolds* is a project assistant professor in the Department of Creative Informatics of the University of Tokyo. He holds a Doctor of Philosophy and Master of Science from the Massachusetts Institute of Technology upon recommendation by the Program in Media Arts and Sciences in the School of Architecture and Planning. His research work there was performed at the Media Laboratory in the Affective Computing Group. Carson also holds a Bachelor of Science in Technical Communication with a Minor in Philosophy from the University of Washington at Seattle.

Alvaro Cassinelli was born in Montevideo (Uruguay) in 1972. In 1990 he obtained both French and Uruguayan Bachelor degree, and a grant to pursue his studies in France. In 1996 he obtained a Graduate Engineering diploma from the Ecole Nationale Supérieure des Telecommunications (ENST), in Paris. He completed the same year a Doctoral Qualifying Degree (DEA) in physics (laser and matter interaction) from the University of Paris-XI/ENST/Ecole Polytechnique. In 2000 he received a Ph.D degree from the University of Paris-XI Orsay for his work on optoelectronic stochastic parallel processors for image processing. Since 2001 he has been working as a Research Fellow, Research Assistant and since 2006 as Assistant Professor at the Ishikawa-Komuro Laboratory, where he is actively involved in creation and development of the new Meta-Perception Group. He has been awarded several prizes as a Media Artist, including the Grand Prize [Art Division] at the 9th Japan Media Art Festival and an Honorary Mention at Ars Electronica 2006.

a byproduct of the process of destruction can either be limited to the machine or vented upon the external world. A machine that destroys itself, and does so with the intent of harming others in the external world is one variety of a weapon. For instance, a land mine is an artificial machine with sensors and mechanically or digitally encoded logic, which was built with the intent of harming some and thus clearly a weapon.

If we exclude weapons, then we are left with the set of machines that self-destruct but limit physical harm to the artificial self. One example of such a system is the MIPS-X microprocessor, which included a special machine instruction *hsc* (Chow, 1986). The programmer's manual for the processor documents the *hsc* instruction as follows;

- 4.58. *hsc* - Halt and Spontaneously Combust
- The processor stops fetching instructions and self destructs.
- Note that the contents of Reg(31) are actually lost.

This is executed by the processor when a protection violation is detected. It is a privileged instruction available only on the -NSA versions of the processor.

This microprocessor is able to disable itself in response to a special command that can be sent by programmers. The destruction however is limited to the processor and does not (by design) seek to cause physical harm outside the processor.

Among non-weapon artificial machines with the capability of self-harm we can further distinguish between systems which are artificially suicidal and those that perform artificial self-sacrifice. One might argue that the distinction between suicide and self-sacrifice is a matter of perspective (as some political groups might label an individual a martyr or suicidal terrorist depending on affiliation). However, we will side step this deeply politicized argument with the following stipulation. When an artificial entity intends to self-destruct to induce psychological harm then we will define it as artificial suicide. This obviously has only a limited correspondence with suicide as it is defined in the human domain, which may stem from various intentions as evidence through actions like euthanasia and running amok. Alternatively, when an entity intends to self-destruct to induce physical or psychological benefit to others then we will define it as artificial self-sacrifice.

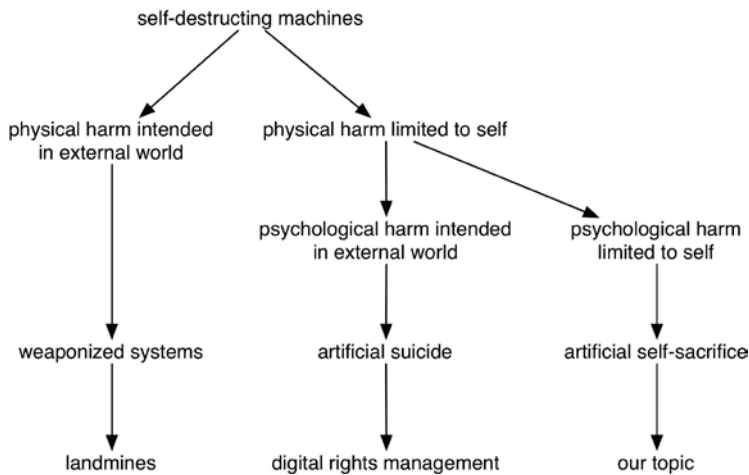


Figure 5: taxonomy of self-destructing machines

The taxonomy employed in this paper is diagrammed in figure 1. From it one can see that we consider digital rights management to be a variety of artificial suicide. Digital rights management systems may render digital content useless after a certain condition is met. For instance, the video game *Spore* includes a system, which allows the game to be installed on a computer three times before disabling itself. We believe this self-destruction is intended to induce psychological harm among would-be pirates of the video game.

Artificial death

Although the topics of artificial self-sacrifice and artificial suicide are somewhat morose, one can view them as liminal topics in the area of Artificial Life (Langton, 1986). If Artificial Life researchers are seeking to provide a definition of life, as evidenced in artificial systems, then it seems that by extension that they must grapple with under which conditions an artificial system dies. It is relevant to note that the relationship between artificial moral agents and Artificial Life has already been discussed by Sullins (2005).

So here we will directly discuss the notion of artificial death. With robot systems, death is something that is ascribed analogically to human death. When a robotic system loses the ability to move or process then it is colloquial to call it “dead.” Further evidence of this analogical thinking can be observed in the naming of “kill switches” which stop the movement of mechanical systems. A recent art installation, “Shockbot Corehulio,” illustrates some ideas about robotic death:

At the point of contact a short-circuit occurs creating a fault current... As the damage to the computer increases, there is a proportional rise of dysfunction to the control signal. This overload of errors ends in a total collapse of the system (Andel & Gütze, 2005).

But then robots can resuscitate by repairing broken components and restoring processing function. A more final variety of robot death would be a case in which the robot is physically destroyed and the ability to resuscitate the robot is no longer possible.

The case of death for digital or software systems is tricky. Digital systems can under normal circumstances be ceaselessly duplicated. So when we “kill” a process in the UNIX operating system, we are only deactivating one of a nearly infinite supply of clones. There is little consequence to this action within the machine since a new process can be quickly activated. However, if we destroy the only remaining copy of some digital system or software then something more serious has happened. The system cannot be easily re-created. Taking this idea further, one might argue that in order for a digital system to truly experience death, it would have take the form of the destruction of a unique non-reproducible system.

Artificial death is not only limited robotic and software systems. Consider for instance a recent variety of self-erasing paper invented at Xerox that clears itself of printed information in “16-24 hours” (Ramplel, 2008). Genetically engineered systems can also self-destruct. Recent work on “fail-safe mechanisms to terminate (gene) therapy” consists of biological suicide switches that can be activated (MacCorkle et al., 1998).

There are fictitious accounts of machine suicide and sacrifice; HAL 9000 appearing in Clarke’s *Space Odyssey* series comes immediately to mind. In *2010: odyssey two* the computer sacrifices itself to save crew members. *Galatea 2.2* is a retelling of the *Pygmalion* myth in which the narrator teaches an artificial intelligence that increasingly becomes overwhelmed and decides to shut down (Powers, 2004).

What each of the artificial systems we have thus far described does have in common is some element of information that is destroyed. Floridi has previously argued that information objects have “intrinsic value” (Floridi, 2002). We will argue that there is value in the ability of information to destroy itself in varieties of artificial self-sacrifice.

A shredder in the library of babel

Let us consider a universe in which information is never destroyed. *Prima facie* this appears an advantageous state of affairs. Any piece of information that comes into existence in this universe is guaranteed to always exist. A Library of Alexandria would not be destroyable and its texts would remain extant. Putting speculation aside we know that both the Library of Alexandria's texts but also an accurate account of their destruction have receded into history. A present-day observer quipped: "It's inherently difficult to get reliable information about an event that consisted of the destruction of all recorded information" (Stephenson, 1996).

Clearly some would like to look in the lost texts of the Library of Alexandria and still others would prefer to know precisely how the books came to be destroyed. But in a universe free of information destruction both could not exist. The Library of Alexandria would remain unburned but the cost would be the loss of all non-fictional information about destroyed information.

"The Library of Babel" hints at a world where information is lossless (Borges, 1999). If all possible texts exist in Borges's library then it would include purely random gibberish, dead languages, and all possible future texts. We can similarly imagine the collection of all finite sequences that under some encoding would then produce every finite text that has or ever will exist.

One problem with the Library of Babel or our collection of all finite sequences is how one quickly finds anything useful. How do I separate out gibberish and languages unknown to me from information that is pertinent? Searching these universes is hard.

For information to be meaningful it must be observable or readable by at least one entity. A permanently inaccessible set of information is as useful as no information at all. Alternatively, a universe in which no information is destroyed (no matter how useless to the observing entities) becomes rapidly cluttered.

Consider the following dilemma: you are the only reader in a finite variety of Borges's Library of Babel. You carry a shredder formed into a backpack. After examining a book you have the choice to replace it back where you found it or to shred it. If you shred the books you believe to be gibberish then you can more rapidly find texts that you can read. Of course you risk shredding a book you would only understand the value of later after some Flowers for Algernon type transformation or just some hard time spent learning to read a language like Mandarin Chinese. What would you make of a text composed of random numbers arduously composed to ensure randomness? Such a text exists (RAND, 2002) and

is even economically valuable enough that statisticians and engineers will pay actual money to obtain a copy.

Limited resources, Kin Altruism and smallpox

From these contrivances we can assert that destruction only becomes an issue in the presence of limited resources. In the above cases, the limited resource is the speed with which the reader can sequentially access texts in the library. If the reader could access in parallel simultaneously all of the texts in the library, then shredding books becomes a vapid activity. The shredder dilemma also becomes interesting when there is limited shelf space in the library or storage in our information system.

Another situation in which resources are limited and fiercely contested is the biological world. Food, sexual partners, sunlight, soil, and water are among the (re) sources of conflict both within and between species. Let us think analogically: a biological system can be equated with information. This might be done by simply transcribing the genome of a particular creature to a set of symbols. Loss of both the symbols and the original creature is akin to information death. We see in the biological world a bewildering array of mechanisms for culling organisms. Interestingly there are many defined varieties of biological death (Schneider and Matakas, 1971):

- necrobiosis: individual cell death (but not necessarily of hosting multi-cellular organism)
- necrosis: death of a group of cells such as an organ
- apoptosis: programmed death of a cell within an organism
- brain death: total necrosis of the central nervous system
- extinction: death of a species

In the case of apoptosis: “most, if not all animal cells have the ability to self-destruct by activation of an intrinsic cell suicide program when they are no longer needed or have become seriously damaged” (Steller, 1995). Death is pervasive in biological systems both within organisms and between organisms.

One edge case is that of *Turritopsis nutricula*, a species of jellyfish that can reverse the life-cycle typical of jellyfish. If not killed by a predator, the jellyfish can repeat the process of returning to a polyp and again becoming sexually mature indefinitely. This leads to one of the few examples in the biological world of a degree of immortality (Piraino, 1996). However, the topic of this article is self-sacrifice so we will now turn our attention to biological self-sacrifice.

One variety of self-sacrifice is kin altruism in which an individual performs actions or undertakes a strategy to their own detriment but to the benefit of genetic relatives. Kin selection is exemplified in sterile ants, bees, wasps, and termites that spend their lives protecting and feeding related offspring (Griffin and West, 2002).

One spectacle of kin selection is the forming of rafts, plugs, ovens, walls, and bridges by ants and beetles. Such assemblies may have a variety of immediate uses for the reproducing members of the kin groups for instance: "mantle shields against rain," "thermoregulatory clusters," "swarms (against desiccation)," or "rafts (against flooding)." In some of these assemblies the organisms give up their lives to benefit the survival of the group (Anderson et al., 2002).

Within evolutionary biology Hamilton's rule is used to model cases in which kin altruism will cause genes to propagate in the genome of a species.

When J. B. S. Haldane remarked, "I will jump into the river to save two brothers or eight cousins," he anticipated what became later known as Hamilton's rule...This ingenious idea is that natural selection can favor cooperation if the donor and the recipient of an altruistic act are genetic relatives (Nowak, 2006).

Hamilton's rule is stated as follows:

$$r > \frac{c}{b}$$

Here r is the probability of relatedness, c is the cost (reproductively) of performing an altruistic action, and b is the benefit (reproductively) to the recipient of the action. If relatedness outweighs the ratio of cost to benefit then according to Hamilton's rule an action should be performed and potentially inherited by offspring. Would it be worthwhile if such a rule existed for information entities?

Another interesting biological example related to the intrinsic value of information is the debate surrounding the destruction of smallpox stocks (see both Jolick, et al., 1993 and Mahy et al., 1993). One group argues that the smallpox virus stocks should be saved for research purposes (or as information for future generations) while a second group argues that the potential for weaponization or accidents outweighs the benefit of preserving the biological specimens. We speak in the next section to provide a frame for arguments concerning destruction of information and its benefit.

Intrinsic value and societal value

Floridi argues for the moral worth of information objects (2002). We are taking the position that in some cases the value of destroying information outweighs its intrinsic moral value. Indeed, Floridi later hints at this possibility with the following quotation:

Nobody would ever argue that this is equivalent to saying that a spider's and a human life are equally worthy of respect. Culling, for example, is an ethical duty in environmental ethics (Floridi, 2008).

Extending from arguments in the section above it is easy to see the biological systems benefit from self-sacrifice. The soldier ant queen is able to continue the colony because she rides upon drowning workers during the flood.

Himma (2004) has previously critiqued "the moral value of things qua information objects." We will make a different claim that some types of information systems do not function without constant information destruction. A slight extension of this is the claim that it is worthwhile for information to destroy itself. Responding to the needs of a society of information entities or biological entities may in some cases enjoin a duty to destroy.

Imagine the following predicament: the grey goo meme. This is a piece of information that duplicates exponentially. An information system into which grey goo is input begins to broadcast the grey goo to other information systems. Such a piece of information would behave as a computer worm does.

One way in which packet switching networks, such as Internet Protocol networks limit their vulnerability to such rouge information is assigning a time to live for information encapsulated. Another concept packet filtering seeks to discard data packets according to certain criteria (Chang, 2002). Internet routers kill packets of stale or unwanted information with a prodigious efficiency. How might we square intrinsic moral value of information at the same time with networking protocol designer's propensity to define ways of destroying information?

A counterpoint to consider is that even malignant worms and viruses that exist in information systems are not entirely worthless. A script kiddy may achieve social esteem by successfully attacking and co-opting personal computers. Indeed, Robert Tappan Morris, the author of the synonymous worm has even been appointed a professorship perhaps in part for his knowledge of such distributed systems. The creators of anti-social exploits may be financially rewarded by groups seeking botnets for nefarious activities. Security researchers catalog and intentionally entrap such pieces of information to write software to protect and harden networks and computing systems.

Our argument may be reduced to the following statement: it is good that information is capable of being destroyed. Let us revisit cases in which information is duplicated without check: computer worms, runaway cell growth, and the grey goo doomsday scenario (Pheonix and Drexler, 2004). Some types of information entities produce negative outcomes for the (eco)systems in which they reside.

We are not arguing that all instances of such information should be destroyed. Obviously cancer researchers benefit from examining cancer cells and a policy of constantly destroying the cells wholesale is ludicrous. We are taking the more moderate view that it would be better if information entities in such systems could knowingly self-sacrifice themselves when the surrounding society, environment, or situation warrants this course of action. While wholesale destruction of all variants of an information entity is not advocated, selective self-destruction of information entities to maintain the function of a society is advocated.

The Signal to noise ratio of information entities

We have examined some more and less absurd situations in which non-destruction of information has negative consequences. Conversely, we've examined some cases (often drawn from biology) in which self-sacrifice has a positive effect on the biological community.

As themes, artificial death and artificial self-sacrifice raise many interesting philosophical and design questions. For instance: should robots destroy themselves in certain circumstances? Also: should the robot have control over the choice to destroy itself? As information entities, they may have intrinsic moral worth, but they may also have duties. Is the act of negating an information entities own existence among these duties?

REFERENCES

- Anderson C.**, et al. (2002). Self-assemblages in insect societies. *Insectes Sociaux* 49(2):99-110.
- Andel E. & Gütze C.** (2005). 5VOLT CORE - [Shockbot] "Corejulio". <http://5voltcore.com/typolight/typolight257/index.php?id=1&articles=1>
- Borges J. L.** (1999). *Collected Fictions*. Penguin.
- Chang R. K. C.** (2002). Defending against flooding-based distributed denial-of-service attacks: a tutorial. *Communications Magazine*, IEEE 40(10):42-51.
- Chow P.** (1986). MIPS-X instruction set and programmers manual. Technical Report, Stanford, CA, USA. <ftp://reports.stanford.edu/pub/cstr/reports/csl/tr/86/289/CSL-TR-86-289.pdf>

- Clarke A. C.** (1984). 2010: Odyssey Two. Del Rey, first edition.
- Floridi L.** (2002). On the intrinsic value of information objects and the infosphere. *Ethics and Information Technology* 4(4):287-304.
- Floridi L.** (2008). Information ethics: a reappraisal. *Ethics and Information Technology* 10(2&3):189-204.
- Griffin A. S. & West S. S.** (2002). Kin selection: fact and fiction. *Trends in Ecology & Evolution* 17(1):15-21.
- Himma K.** (2004). There's something about Mary: The moral value of things qua information objects. *Ethics and Information Technology* 6(3):145-159.
- Joklik W. K., et al.** (1993). Why the smallpox virus stocks should not be destroyed. *Science* (New York, N.Y.) 262(5137):1225-1226.
- Langton C. G.** (1986). Studying artificial life with cellular automata. *Phys. D* 2(1-3):120-149.
- MacCorkle R. A., et al.** (1998). Synthetic activation of caspases: artificial death switches. *Proceedings of the National Academy of Sciences of the United States of America* 95(7): 3655-3660.
- Mahy B. W., et al.** (1993). The remaining stocks of smallpox virus should be destroyed. *Science* (New York, N.Y.) 262(5137):1223-1224.
- Nowak M. A.** (2006). Five rules for the evolution of cooperation *Science* 314(5805):1560-1563.
- Phoenix C. & Drexler E.** (2004). Safe exponential manufacturing. *Nanotechnology* 15(8):869-872.
- Piraino S., et al.** (1996). Reversing the Life Cycle: Medusae Transforming into Polyps and Cell Transdifferentiation in *Turritopsis nutricula* (Cnidaria, Hydrozoa). *Biological Bulletin* 190(3):302-312.
- Powers R.** (2004). *Galatea 2.2: A Novel*. Picador.
- Ramplell C.** (2008). 'Xerox Creates Self-Erasing Paper. *The Chronicle for Higher Education Wired Campus Newsletter*. <http://chronicle.com/wiredcampus/article/2956/xerox-creates-self-erasing-paper>
- RAND Corporation (2002) *A Million Random Digits with 100,000 Normal Deviates*. American Book Publishers.
- Schneider H. & Matakas F.** (1971). Pathological changes of the spinal cord after brain death. *Acta Neuropathologica* 18(3):234-247.

Steller H. (1995). Mechanisms and genes of cellular suicide. *Science* 267-(5203):1445-1449.

Stephenson N. (1996) 'Mother Earth Mother Board. *Wired* 4(12). <http://www.wired.com/wired/archive/4.12/ffglass.html>

Sullins J. (2005). Ethics and Artificial life: From Modeling to Moral Agents. *Ethics and Information Technology* 7(3):139-148.

Mixed Feelings Computing

Nikola Šerbedžija*
Fraunhofer FIRST Berlin

Abstract

Introducing awareness and responsiveness to human senses into computing, a realm known to be rigid and explicit in its present form, significantly enriches technological systems and their communication capabilities. However, the processing based on our feelings, mental state or physical condition goes beyond the already weakened privacy boarder and raises ethical concerns. The paper discusses both technical feasibility and ethical impacts of the novel approach.

Keywords: Pervasive Adaptation, Smart Systems, Software Engineering, Ethical Issues, Privacy

What you feel is what you get

Remember the times when Orwell's "Big Brother is watching you" was a metaphor. Nowadays, the technological systems offering their services based on collection and intelligent analyses of human psycho-physiological data are often ahead of science fiction. This raises mixed feelings.

Mixed feelings (or reflective) computing is an interdisciplinary approach to support building smart systems that study users and their surroundings and react according to the situation and user's cognitive, emotional and physical state. The approach combines the cutting edge achievements in emotional/ physiological/ neuro- computing, software technology and man-machine interface, joining technical science and humanities in a common goal to develop genuine user-centric systems. A new generation of smart systems should understand user's needs, intentions and social situation and provide appropriate assistance in a discrete and personalized manner.

* Nikola Serbedzija is a senior scientist at Fraunhofer FIRST in Berlin and professor at University of Arts, Berlin. His major research areas are: Middleware Architectures, Internet Computing, Pervasive and Ubiquitous Systems, Embedded and Real-time Systems. He is involved in numerous multidisciplinary projects and initiatives concerning new generation smart control systems.

The biocybernetic loop (Pope, 1995; Serbedzija and Fairclough, 2009) is the core component of a reflective computing system. The loop is initiated by the collection of psychophysiological data from the user via ambulatory, remote or wireless sensors (Serbedzija and Fairclough, 2009). These data are quantified to diagnose relevant psychological constructs, e.g. frustration, user engagement, and alertness. The functional goal of the biocybernetic loop is to derive real-time adaptations that appear both timely and intuitive from the users' perspective. Implementing biocybernetic loop in a technical system brings a radical change into the man-machine interface. Explicit interaction as the usual way of controlling the technical system is substituted by implicit, seamless control. In traditional systems, user friendliness is often considered to equate with ease of use (e.g. for word processing – WYSIWYG – what you see is what you get). This new approach to user friendliness, however, presupposes an implicit uncovering of user needs via real-time interaction. In another words, the motto is “what you feel is what you get”.

Seamless, non-explicit interaction promises applications that can be used in various domains: from entertainment, education and ambient assistance to embedded real-time systems. The use of new technology should have a great impact on future products and redefine the way of living. It should transform a music player into a mood player, an on-line learning platform into an effort-sensitive personal teacher, a shopping mall into a shop-mate. Furthermore, traditionally senseless places can be enriched with personalized characteristics: e.g. a kindergarten plays a role of a baby sitter, an elderly house becomes a supportive life-assistant, a vehicle acts as a friendly co-driver.

However, diagnosing and/or recording of persons' emotional or cognitive characteristics combined with massively interconnected devices able to influence human behaviour opens yet another Pandora's box. What if a “mood player” reinforces negative emotions or a “shop mate” persuades shopper to spend more money, or any other “smart” assistant behaves not in users' but in someone else's interest.

The rest of the paper explores the dilemma: to what extent should we allow smart technology to interfere with everyday life, discussing both technical and ethical issues. It is organized in five sections. After the introduction, the second section deals with software technology that makes reflective computing possible, illustrating it with concrete technical description that relies on psychological research. The third section offers application scenarios that indicate how new technology can re-enforce comfort, safety and joy in everyday situation. The fourth section introduces another side of the coin focusing on possible misuses, discussing ethical concerns and the ways to preserve privacy in a world of global com-

puting. In conclusion, a multi-disciplinary approach is suggested that gathers scientist from different fields in efforts to extend each other's knowledge and jointly deal with ethical dilemmas that a brave new world faces.

Mixing the feelings

The basic computer principle has always been input-process-output. From the early days, computers have been endlessly cycling these three simple steps, faster and faster, becoming a non-avoidable factor of everyday life. Presently, however, we are experiencing an interesting shift: emerging of a new computer systems functioning in a “sense-analyse-react” manner. This innovation allows for development of genuinely user-centric systems. It equips the computers with means to determine psycho-physiological image of the users, as well as the environmental situation, and adapt the system functioning accordingly. These systems are often called pervasive adaptive systems (Costa, 2008)

Under the motto “the best assistant is the one you do not notice”, the reflective approach (Reflect, 2009) investigates ways of realizing pervasive-adaptive environments. A generic software framework is being developed with a set of practical tools for building context-aware, self-organized control systems featuring seamless collaboration with users and reflective computing. The driving force behind the sense-analyse-react paradigm is affective computing (Picard, 1997; Byrne, 1996), a rapidly evolving discipline that investigates how to capture and interpret affects (emotions with accompanying movements), postures (ways of person's holding/carrying body) and gestures (expressive bodily movements). It also relies on so called psycho-physiological variables (PPV) that can be measured with modern microelectronic sensors. By the analysis of the collected data it is possible to precisely determine user's emotional, mental or physical state, especially in situations of a predictable behaviour. Once the user state is evaluated, the system can be tuned to provide a supportive reaction.



Figure 1 Spinning the senses

Developing software to control pervasive and adaptive systems includes real-time sensor and actuator control, user and context-awareness, affective computing, self-organization and adaptation. To perform these tasks, modular middleware architecture is being designed that simplifies the construction of dynamic sense-analyse-react behaviour. Figure 1 illustrates the reflective middleware with three layers:

- Tangible layer: controls sensor and actuator devices. This layer makes its atomic services available to the rest of the system.
- Reflective layer: combines the atomic services of the lower layer with user profiles and scenario descriptions. This allows for components of far higher complexity, which can evaluate the user's emotional, cognitive and physical states as well as his environment and consequently propose a system (re-) action.
- Application layer: presents applications scenarios and combines low-level and high-level software components from other layers to run and control the whole system.

Such a system can adapt through iterative self-tuning, accounting for both the state of the user and the influence of the previous system response (closed-loop control).

Bio-cybernetic loop

The design of a physiologically based system relies upon the biocybernetic loop. The loop describes how psychophysiological data from the user is captured, analysed and converted to a computer control input in real-time. The function of the loop is to monitor psychophysiological changes in order to initiate an appropriate adaptive response. The biocybernetic loop is designed according to a specific rationale, which serves a number of meta-goals. For instance, the biocybernetic loop may be designed to:

- promote and sustain a state of positive engagement with the given task
- minimise any health or safety risks inherent to the human computer interface

The biocybernetic loop is equipped with an array of adaptive interventions to promote these meta-goals (Gilleade, 2005) e.g. to provide help, to give emotional support, to manipulate task difficulty. The implementation of these interventions is controlled by the loop in order to 'manage' the psychological state of the user. Correspondingly, the way in which person responds to each adaptation is how the user 'manages' the biocybernetic loop. This is the improvisatory crux that achieves human-computer collaboration by having person and machine respond dynamically and reactively to each other. It may be useful for the loop to monitor how users respond to each intervention in order to individualise and refine this dialogue. This generative and iterative model of HCI emphasises the importance of equipping system with an elaborate repertoire of adaptive responses that covers the full range of possible outcomes within the human-computer dialogue over a period of sustained use. The latter point is particularly important

for ‘future-proofing’ the physiological computing system as user and machine are locked into a co-evolutionary spiral of mutual adaptation.

The interaction between user and system via the biocybernetic loop may be differentiated in terms of timescales for system adaptation.

- awareness of user state (seconds/minutes/hours)
- adaptation to stable traits (hours/days/weeks)
- adaptation to trait changes (months/years)

Given that the meta-goals of the biocybernetic loop are to engage and protect the user, how should the loop response to cases when both goals are incompatible? For example, when the player of a computer game registers boredom because of an extended period of play? If the primary goal of the loop is to engage the player, the system may respond with a stimulating increase of task demand (Serbedzija and Fairclough, 2009). With the goal of protecting health in mind, the loop may suggest that the player takes a rest break. This scenario draws attention to the requirement for a primary directive or meta-goal for the loop. The designer must decide whether the biocybernetic loop emphasises engagement, health, or safety as the “bottom line.”

The biocybernetic loop may use two inherent dynamics: negative or positive feedback control. This is another important design option for Physiological Computing systems. Negative control loops create stability by reducing the discrepancy between the input signal (real-time psychophysiological measure of engagement) and a desired standard (the desired level of engagement). Negative feedback control is perfect if the system has been designed to keep the user within a ‘safe’ zone, such as avoiding extremes of fatigue or stress.

The biocybernetic loop requires a sensitive and reliable representation of the user and the user state in order to function. This representation may be multi-layered, representing state changes in seconds or minutes due to discrete events at the interface, as well as representing the personality or proficiency of the user on a longer time scale. It is important for the system to differentiate dynamic and sporadic changes in user state (awareness) as well as tracking slower changes against a background of stable user traits. This provision allows the biocybernetic loop to “sense-analyse-react” on several levels simultaneously in order to feed the co-evolutionary dynamic between user and system. As a result, the whole system exercises adaptation on several time scales, from short term (awareness) via medium term (adaptation) to a long term (evolution).

Sensuous applications

Reflective systems can be applied in a number of complex situations. The reflective approach does not redefine the existing systems; it rather provides another “reflective” dimension to it. Actually any system that uses computer to control the environment can be enriched with sensuous behavior: (1) reflective mobile phones that offer hints (whom to call, whose call to ignore, where to go, what to do) depending on the owner’s state and given situation; (2) reflective music player that selects the music according to users’ mood (Schroeder et al., (2008); (3) reflective advertising that adapts the content and the way of presentation according to a number, kind and interests of viewers; (4) reflective home for elderly people that recognizes needs and weaknesses of the inhabitants and assists and/or calls for assistance.

Vehicle as a Co-driver

A driving experience is an obvious example where reflective system can be used. Generally speaking it is always better to drive in company. A co-driver usually observes the driver carefully, keeps an eye open on viewpoints the driver cannot see, makes a vivid atmosphere on a longer trip and assists the driver significantly. However a group ride is not a frequent case. The reflective vehicle should overcome possible shortcomings of a solitary drive by overtaking the role of a friendly co-driver (Serbedzija et al. 2008). Its task is to observe the driver and taking into account driver’s emotional, cognitive and physical state as well as vehicular, driving and traffic conditions, optimize the vehicle configuration and actively participate in a complex driving process. Reflective vehicle concept aims at implementing adaptive control in vehicles, achieving more secure, more pleasant and more effective driving.

Mood Player

Home ambient is yet another environment which is more and more exposed to computer control. So called smart homes already offer intelligent kitchens that assist inhabitants in everyday house work, control air-condition, heating and lightening according to the outside temperature, time of the day and the tenant’s habits. The reflective home goes a step further trying to improve the ambient and comfort people according to their current emotional, mental and physical situation.

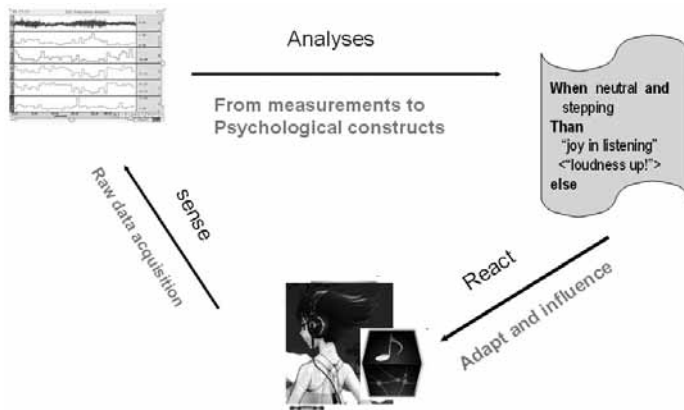


Figure 2: Reflective player

Figure 2 illustrates the functioning of a mood player, as a part of a more complex reflective home system. Its biocybernetic loop drives the music/video player at home. The reflective player observes diagnoses and responds to the affective, mental and bodily states of a person in a closed-loop fashion. The biocybernetic loop consists of:

- Sensing: a) The person's psycho-physiological characteristics (using facial expression, speech and movement detectors) and b) player settings and lighting information (music player and lighting system are used as both sensors and actuators)
- Analysing: Diagnosing the person's emotional and physical condition (positive emotion and tapping to the music are obvious signs that a listener likes the music)
- Reacting: Changing the player settings and ambient light to adapt to the person's condition and enjoyment of the music being played (if a person likes the music, lighting can be changed and the volume of the music may increase)

In a closed loop, the system re-examines the effects of its (re-)actions to further refine its behaviour. The ultimate goal is to ensure that listening to the music reinforces joy.

The application area of reflective technology is wide – it can be used in any situation where computer control is deployed, e.g.:

- Automotive: reflective vehicle, reflective road
- Communication: reflective phone, reflective browser

- Education: reflective learning
- Entertainment: reflective player, reflective computer game
- Urban infrastructure: reflective advertisement display, reflective guide
- Ambient Assisted Living: reflective home for elderly people, reflective kindergarten
- Robotics

The examples of sensuous applications illustrate how reflective systems can help people in every day situation. However, since bio-cybernetic loop controls the ambient according to the given “meta-goals”, it is not difficult to contemplate that the effective use of the whole system can be manipulated and is heavily dependent on the pre-scribed purpose of the system.

Mixed feelings

The ultimate goal of reflective approach is to make future control systems genuinely friendly and personalized to suit the needs of individual users. To achieve this goal it is not enough to create interfaces that suit an average users, but to make it really serves the personal needs, taking into account current necessities as well psycho-physiological state of the persons being involved in real-time. At one hand, emerging mixed feelings systems that are reactive to human senses can be interpreted as our technical “sixth sense”, at another we should be aware that our personal identity is being exposed to the interconnected world whose meta-goals may be different from our expectations.

This is exactly what raises mixed feelings!

Knowing how interconnected the digital world is, recording what we buy and eat, where we travel, sleep, rent-a car (e.g. credit card institution), what we read (e.g. Amazon), what we write or visit at the Web (our internet provider), how much fun we had (e.g. Flickr), how we socialise (e.g. Facebook), how our biometric signature look like (e.g. custom control), our medical record (health insurance smart cards) we can fully relax and let “Infosphere” (Floridi, 2007) assist in most of our every day activities. The digital divide is sharpening, “digitalize or perish” seems to be the motto – those who are left behind the digital curtain belong to third (digital) community. With the technology described here as well as with neuro-science that is making a huge progress in the domain of brain-computer interface we are about to submit the last fence of our privacy to the “digital consideration”. Namely our feelings, mental constitution, even our thoughts.

Certainly, the use of modern technology facilitates the life, provides efficiency, comfort and smooth communication in a way it could not be even perceived a

few decades ago. Especially the applications in medicine and ambient assisted living domain represent clear justification of the recent achievements. Nevertheless, a potential for misuses is wide and what is even more disturbing, neither possibilities nor risks are properly understood.

The problem of the impact of the new technology to our life and society as a whole is neither new nor unsolvable (Hettinger, 1989, Floridi, 2007). Due to rapid developments it is sharpening though, and needs to be addressed from different viewpoints involving wide and cross-disciplinary discussion. The roles are traditionally divided among:

- artists who picture the universe in a free and imaginative way (Orwel, 1949) with a free mind calling for re-thinking, re-consideration, re-involvement, esthetic, *l'art-pour-l'artism*, ...
- scientists and technology providers who are not only responsible for the development of novel ideas but also for hints on how to technically deal with possible ethical concerns (Rulon, 2009).
- philosophers and sociologists as leaders in considering the impacts of new technology to the society as a whole (Hettinger, 1989, Floridi, 2007).
- law makers and politicians to insure efficient legal background and deployment (EC, 2008)
- practitioners and industry to respect regulation and ethical norms while making commercial and other use of modern technology.

In one of his articles, Floridi (2007) introduces a neologism *Infosphere* as a collection of informational entities that inevitably constitutes the environment that supports our life. To further understand the impact of the *Infosphere* to us (as *Infogs* – informational organisms), the author introduces “re-ontologization”, referring to fundamental transformation of our environment as a consequence of digitalization and *Infosphere*. Such a complex approach allows for thorough reasoning about the meaning, impact and ethics of *Infosphere* as a non-avoidable part of our ecological system. In that context the observation given here can be described as “nano-Ethics” focusing only at a small segment of the problem – where technology may help. For example, making reflective systems closed by strict separation of psycho-physiological and administrative data, and making it technically impossible to exchange information with other systems – would be a step in right direction in protecting the privacy.

Conclusion

The paper presents a novel approach in building smart systems illustrating both its technical background and its applicability. Finally, the consequences of fast and uncontrolled deployment of smart technology are considered, warning that the “digital victory” may turn into a Pyrrhic one, as the use of massively interconnected digital devices may endanger our privacy exposing all aspects of our behaviour from the every day activities, work competences, habits, feelings, intentions up to our inner thinking

It seems that the one of the main controversies a modern society is confronted with is not the rapid technology development which sometimes goes even beyond science-fiction, but rather the slow pace at which humanities answer, or are allowed to answer to the new moral challenges. Their main task remains to be to warn about possible nightmare scenarios. Orwell’s greatness (Orwell, 1949) cannot be measured by the fact that he envisaged many abuses already present nowadays, but rather by his warnings which helped that his worst predictions haven’t (yet) became true. Therefore, an interdisciplinary approach is a *conditio sine qua non* in all future research programmes.

As technology and science advance, the spectrum of further challenges is wide. In a cross-disciplinary endeavor, psychologists are needed to put more light on affective analyses, visionary practitioners to create more application scenarios, philosophers and sociologists to consider what is right and what is wrong, and computer scientists to ensure that the “smart assistant” is genuinely personal, not a friend of the Big Brother.

Acknowledgement

Most of the work presented here has been done under the REFLECT project (project number FP7-215893) funded by the European Commission within the 7th Framework Programme, pervasive adaptation initiative.

REFERENCES

- Pope, A.T., Bogart, E.H, and D.S. Bartolome (1995). Biocybernetic system evaluates indices of operator engagement in automated task. *Biological Psychology*, Vol. 40, 187-195.
- Serbedzija, N. and Fairclough, S. (2009). Biocybernetic Loop: from Awareness to Evolution, Proc, IEEE Congress on Evolutionary Computation, IEEE CEC 2009, May, Trondheim, Norway

Costa de C.A. et al. (2008). Toward a General Software Infrastructure for Ubiquitous Computing, Pervasive Computing, January-March 64-74.

Picard, R. (1997). Affective Computing, MIT Press.

Byrne, E. and Parasuraman, R. (1996). Psychophysiology and adaptive automation. Biological Psychology, Vol.42, 249-268.

Reflect (2009). The REFLECT project - Responsive Flexible Collaborating Ambient, accessed 30 April 2009, <http://reflect.first.fraunhofer.de>.

Gilleade, K.M., Dix, A. and Allanson, J. (2005). Affective videogames and modes of affective gaming: assist me, challenge me, emote me. in Proceedings of DiGRA 2005.

Schroeder, A et al. (2008) Middleware Architecture for Human-Centred Pervasive Adaptive Applications, Proc. 1st PerAda Workshop at SASO 2008, 21. Oct. 2008, Venice, Italy.

Serbedzija, N.S., Calvosa, A.M. and Ragnoni, A. (2008). Vehicle as a Co-Driver, Proc. First Annual International Symposium on Vehicular Computing Systems – ISVCS, Dublin.

Hettinger, E. (1989). Justifying intellectual property. Philosophy and Public Affairs, Vol.18, 31-52.

Floridi, L (2007). A look into the future impact of ICT on our lives, Information Society 2007, Vol. 23/1, 59-64.

Orwell, G. (1949). Nineteen Eighty-Four. A novel. London: Secker & Warburg.

Rulon, M. (2009). New Technology Raises Privacy Concerns, USA Today, accessed 30 April 2009, http://www.usatoday.com/tech/news/surveillance/2006-08-31-rfid-privacy_x.htm

EC (2008). The European Group on Ethics in Science and New Technologies to the European Commission, Ethically speaking, No. 9, http://ec.europa.eu/european_group_ethics/publications/docs/ethically_speaking_9.pdf

Ethical Issues of Pervasive Computing in the Insurance Industry

Oliver Siemoneit*

Institute of Philosophy,
Chair of the Philosophy of Science and Technology,
Universität Stuttgart, Stuttgart, Germany

Abstract

The recent progress of nanotechnology, microelectronics, and communication technology has given rise to a vast field of new application. Every day objects, clothes, cars, buildings, or even the human body can be invisibly equipped with tiny sensors as well as with information and communication technology systems which collect, receive, process, save and post data. As these technologies have become cheaper and more mature in the course of the past few years, quite a number of insurance companies have started to introduce them for individual risk inference and individual premium calculation. It is the aim of this paper to analyse the impact of Pervasive Computing technologies on insurance from an ethical point of view. The threat to privacy as well as the misuse and rededication of data are especially focused on here. Finally, some technical solutions are to be outlined which present, from an ethical point of view, good and acceptable solutions, i.e. they avoid most of the drawbacks of current technical solutions but nevertheless do not cut off all their advantages.

Keywords: Pervasive Computing, Ubiquitous Computing, Insurance, Risk Calculation, Actuarial Fairness, Ethics, Privacy, Surveillance

Introduction

It was Mark Weiser, who in 1991 first introduced the idea of so-called Ubiquitous Computing in his paper "The Computer of the 21st Century" [Weiser (1991)]. There, Weiser propagated a profound integration of information and communication technology (ICT) into everyday life. Everyday objects, clothes, cars, buildings, or even the human body should be invisibly equipped with tiny

* Oliver Siemoneit, born 1975, studied business management and engineering and is now working since 2003 at the Institute of Philosophy, Chair for the Philosophy of Science and Technology, Stuttgart, Germany. Fields of expertise: Philosophy of Science and Technology, Information Ethics, Business Ethics.

sensors as well as with ICT systems which collect, receive, process, save and post data, thus easing our life and opening up new possibilities. Setting the research agenda by this, in the aftermath a lot of different terms have been coined for the same idea: Pervasive Computing, Ambient Intelligence, Sentient Computing, Calm Computing, just to mention a few. Although these different terms may suggest different things, they all have something in common. They name a certain kind of visionary ICT usage which could be characterized as follows:

- **Miniaturisation:** ICT components and assemblies become much smaller and thus more portable.
- **Embedding:** Miniaturisation and falling prices allow for the embedding and/or massive proliferation of ICT systems in everyday objects.
- **Interconnectedness:** Embedded ICT systems are normally interconnected, communicate wirelessly with each other, locally through ad-hoc networks or globally through infrastructure-based networks and the internet.
- **Context-Awareness:** Embedded ICT systems collect or gain information about their environment and/or the user and adapt their behaviour accordingly (thereby becoming “smart”).

All in all, the vast progress of nanotechnology, microelectronics, and communication technology in the course of the past few years has permitted turning the mere vision of Pervasive Computing into reality. Especially during the past few years, insurance companies have started to introduce Pervasive Computing technologies so as to offer new innovative products for their customers. It is the aim of this paper to give a in-depth discussion on this, to point out the pros and cons and to suggest technical solutions which, from an ethical point of view, may also be considered good, fair and acceptable solutions.

Pervasive computing in the insurance industry

The extensive proliferation of wireless sensor technology as well as embedded ICT systems in everyday objects has two main effects for insurance companies: a) in the field of risk monitoring and b) in the field of risk control and prevention. Concerning a) risk monitoring, Pervasive Computing technologies for the first time allow for the detailed, real-time collecting of risk-relevant data of the insured object, which could serve as the basis for a more accurate insurance premium calculation based on the actual size of risk and the actual probability of damage. Concerning b) risk control, ubiquitously available, intelligent, smart ICT systems may help prevent cases of damage or reduce the amount of loss by early warnings and detections or by proactive action. It is often said that this applies both to “non-life” insurance business (e.g. cars, buildings) as well as to “life” in-

insurance business (insurance of individuals like health insurances, disability insurances, life insurances in the narrow sense of the word) [Filipova (2007), p. 4; Müller and Zimmermann (2003), p. 5].

Regarding “non-life” insurance business, starting in 2002 quite a number of insurance companies around the world have introduced more or less sophisticated Pervasive Computing technologies in motor insurance. A “black box” in the insured vehicle records risk-relevant data such as

- driver and duration of a trip,
- time of a trip (day? night? rush-hours?) and driving route (main road or country road),
- duration of trip and possible breaks,
- changes of direction, braking and acceleration characteristics, speed, and other telemetric data (e.g. indicating skidding),
- weather and road conditions,
- distance to the car in front aso.

Furthermore, matching the current position of the vehicle with maps should allow for checking whether speed limits have been violated or whether road signs have been ignored. All these data are collected, recorded thoroughly and finally sent to the insurance company which uses them to estimate the individual risk and accordingly raises or lowers insurance premiums.

From an economic point of view, permanent and detailed risk monitoring is said to be a good and fair concept, since it a) discharges high premium from responsible drivers and b) also reduces so called “asymmetric information” in contractual relationships and thus prevents “adverse selection” and market failure [Filipova and Welzel (2005), Filipova (2006), Filipova (2007)]. In case of risk control, smart intelligent vehicles should help preventing damage by measuring the driver’s alcohol level or tiredness, reporting low air-pressure of wheels and indicating the need for technical service. All in all, Pervasive Computing technologies in vehicles are often made palatable to the customer by bundling black boxes in vehicles with other services, thus offering added value through increased convenience and safety. E.g. in case of an accident or breakdown, automatic emergency calls at different levels may be sent out informing the rescue services about location and amount of damage in detail. Also gathered information of different cars with black boxes may be merged together to create a digital, real-time representation/model of road condition and traffic [as it is the case with the Falcon project, see CERT (2009)], helping forecast and prevent traffic jams and accidents by fore-

warning other road users or by using new ways of advanced automotive vehicle navigation.

However, the idea of using Pervasive Computing technology is said to be not only restricted to “non-life” insurance business. It is also supposed to be applicable to “life” insurance business. There, as the vision says, smart artefacts could also monitor the “insured object” (i.e. the insured person), thus allowing for detailed risk inference and premium calculation. At the same time, this opens up the market for new service providers, as companies do not only sell products which serve as the basis for individual risk inference but also come forth with new services [Müller and Zimmermann (2003), p. 7]. “Adidas could offer, manage and control individualized training programs based on the data acquired from communicating sensors in Adidas sports shoes. This information could be used by the health insurance company, and prices would be adapted according to the training efforts of an insurance holder.” [Müller and Zimmermann (2003), p. 7]. Furthermore, “telemetric data” (as in automotive insurance) like pulse, blood parameters could be sent directly to the family doctor, allowing for an all-embracing health care / health examination and quick help in cases of emergency.

Actuarial fairness in life and non-life insurance business

At a first glance, there seems to be no problem with using an actuarial conception of fairness in motor insurance. Differentiation of premiums for individuals according to their risk level rewards responsible driving, sanctions risky driving and prevents catastrophic driving by completely denying insurance. This seems to be ethically justifiable, since usually you have a choice to drive in a different manner. However, the question becomes already more complicated if you consider mobility a kind of basic right (which should be affordable to everyone in a free society), and if you enter into the difficult discussion on the extent to which young drivers and old drivers “have a choice” to change their risk level. They often have not. And the same is mostly true for basic health issues, since people do not have complete control over the development of their individual healthiness. Using only private insurance concepts in this realm means to fully place the costs of misfortune on the unfortunate. And that is why here, from an ethical point of view, a concept of actuarial fairness could be considered to be quite unfair (and thus ethically hardly justifiable).

Privacy, “costs of privacy” and risk perception

The main problem of Pervasive Computing technology in insurance is the loss of privacy, though. Not only does the detailed, real-time recording of risk-relevant data allow for an exact calculation of individual risk level, but it also allows for a

profound profiling of a person's activities, preferences, habits. This is quite problematic and ethically challenging in two ways. First, these days people do not seem to care anymore about privacy at all. Privacy is often considered an anachronistic concept, a concept of unsavoury people who have to conceal something. It is the duty of ethics to emphasize the value of privacy for the constitution of a self-determined person in a free society [Rössler (2005); Heesen and Siemoneit (2007)]. Second, often it seems that the consequences when agreeing on the revelation of personal data are underestimated, thus raising the question whether the informed consent is really informed. Risk perception and "cost estimation" is also the blind spot of the above mentioned economic modelling: There, it is often stated that the "costs of privacy" change "in some way", however this question is never asked: How could a loss of privacy be monetarily quantized at all? How could the "costs of privacy" be measured intersubjectively? After all: How rational is the subjective measurement of costs, since it is well known from risk research that habitual risks (e.g. traffic participation), creeping risks (such as smoking), risks easy to observe or those being actively taken are often underestimated [Renn (2008), p. 93 et seq.]. Badly observable risks (electromagnetic pollution), risks that have not been actively taken but one is exposed to (asbestos) or risks which result in heavy casualties (airplane crash) are often overestimated [Renn (2008), p. 93 et seq.]. In the context of Pervasive Computing this could mean: Becoming a victim of surveillance activities of the economy or of the government seems to be heavily underestimated, just as the consequences resulting from the revelation of personal data in the long run. (We know this discussion already from the internet.) Creeping risks, such as slowly developing into a surveillance society, or actively taken risks which are based on an "informed consent" such as using pay-as-you-drive-solutions (which from the ethical point of view are still based on an uninformed consent) seem to be underestimated as well. It is thus the duty of ethics to carry out more awareness training about the risks of the information society.

Justified and justifiable technical solutions for pervasive computing in the insurance industry

Hence, to prevent the misuse of collected data by the economy as well as to keep the state from an extensive, unnecessary rededication of personal data for criminal prevention, in the following different technical options should be outlined which would meet these requirements. Thereby the focus is on automotive insurance only, since this is the realm with the most mature concepts being ready for practical deployment in everyday life.

Basically, so as to assure privacy and prevent extensive surveillance, there are two operating levers: First the reduction of data quality/granularity and second the

change of data update rates (whereas both “levers” can be combined in different ways). One feasible technical solution for automotive insurance would thus be to replace the permanent tracking of cars and transmission of data by sending out only aggregated data to the insurance company at bigger time intervals. This aggregated data would also allow for a detailed risk inference. However the data would be of a quality which, beyond risk calculation, is of no worth either for the insurance company, governmental authorities or others. Privacy would be preserved, surveillance complicated or impeded, nevertheless not cutting off all the advantages of pay-as-you-drive for both, the insurer and the insured. Another more radical technical solution would be to restrict the revelation of (aggregated) data to cases of loss only, i.e. conceptualizing pay-as-you-drive as complete “off-line-solutions”. Only in cases of accident the insurance company is allowed to have access to the black box in situ, i.e. directly at the car by connecting a cable to it and thus receiving data about the driving manner. This technical solution would not produce either a “glassy” driver but just an obedient one, for whom the permanent monitoring of his driving manner has become a common, everyday circumstance. Both depicted solutions are appropriate and sufficient to reach the basic aim to get something to know about the risk level of an insured object, but they do not come along with an inadequate, unnecessary massive collection of different types of high quality data which could be used and/or misused for other purposes.

Conclusions

It was the aim of this paper to give an overview of the evolving topic of Pervasive Computing in the insurance industry, pointing out the pros and cons from an ethical point of view. On the one hand, it has been made clear that there are no vital concerns about actuarial fairness if the insured person has a choice to alter his/her risk behaviour. Concepts of actuarial fairness, however, yield at the same time to a new form of paternalism, mass-education and uniformisation of people to certain standards, which are considered “normal”, “desirable”, “healthy”, “socially acceptable” also by certain social groups [sic!]. On the other hand it has been stressed that the possibility of having a choice is often not given in the realm of basic health issues. There, the availability of private health insurances only based on the concept of actuarial fairness would mean placing the complete costs of misfortune on the unfortunate, which is hardly justifiable from an ethical point of view. Hence, universal public health insurances are indispensable. However, conceptions of actuarial fairness based on Pervasive Computing technology gain more and more explosiveness in the light of the question to which extent public insurances are (still) affordable. It is the duty of the societal discourse to decide to which extent an “insured claim” should be covered by private or public

insurance concepts and to which extend an insurance system should be economically profitable and promising. Both public and private insurances are “risk communities” (based on some kind of “solidarity”), whereby the former is often based on massive cross-subsidisation from different financial sources, while the latter is economically feasible on its own and even promises earnings. But if concepts of actuarial fairness based on Pervasive Computing technology are employed, the acquired data should be sufficient and adequate to reach the purpose of risk inference only. Unnecessary data collection should be avoided. Collected data should be kept locally in an aggregated manner and not in an all-comprising, central spatial model (as it is the case with the Falcon Project of the United Arab Emirates). Only these measures could really assure privacy and seem to be compatible with the core values of Western liberal and democratic societies, thus keeping second and third parties from misusing the collected data for other (“own”) purposes.

REFERENCES

CERT (2009): Falcon: A Versatile Telematics Solution. Available at: http://www.certinfotrack.com/includes/Falcon_brochure.pdf. Last access: 19.3.2009.

Filipova, L. (2007). Monitoring and Privacy in Automobile Insurance Markets with Moral Hazard. Volkswirtschaftliche Diskussionsreihe, Institut für Volkswirtschaftslehre der Universität Augsburg, Paper No. 293. Available at: <http://www.wiwi.uni-augsburg.de/vwl/institut/paper/293.htm>. Last access: 19.3.2009.

Filipova, L. (2006). Endogenous Information and Privacy in Automobile Insurance Markets. Volkswirtschaftliche Diskussionsreihe, Institut für Volkswirtschaftslehre der Universität Augsburg, Paper No. 284. Available at: <http://www.wiwi.uni-augsburg.de/vwl/institut/paper/284.htm>. Last access: 19.3.2009.

Filipova, L. and Welzel, P. (2005). Reducing Asymmetric Information in Insurance Markets: Cars with Black Boxes. Volkswirtschaftliche Diskussionsreihe, Institut für Volkswirtschaftslehre der Universität Augsburg, Paper No. 270. Available at: <http://www.wiwi.uni-augsburg.de/vwl/institut/paper/270.htm>. Last access: 19.3.2009.

Heesen, J. and Siemoneit, O. (2007). Opportunities for Privacy and Trust in the Development of Ubiquitous Computing. International Review of Information Ethics, Vol. 8, No. 8. Available at: http://www.i-r-i-e.net/inhalt/008/008_8.pdf. Last access: 19.3.2009.

Müller, Ch. and Zimmermann, H.D. (2003). Beyond Mobile: Research Topics for upcoming Technologies in the Insurance Industry. Proceedings of the 36th Hawaii International Conference on System Sciences.

Renn, O. (2008). Risk Governance: Coping with Uncertainty in a Complex World. Earthscan, London.

Rössler, B (2005): The Value of Privacy. Polity Press, Cambridge.

Weiser, M. (1991). The Computer for the 21st Century, Scientific American, Vol. 265, No. 3, 94 – 104.

MyChoice & Traffic Lights of Trustworthiness: Where Epistemology Meets Ethics in Developing Tools for Empowerment and Reflexivity

Judith Simon*
Institut Jean Nicod

Abstract

One major trend in software development has been labelled social software. A key feature of it is that social networks or trust relationships between users of a system are used for the selection and evaluation of the quality of information provided on the web. Based on such observations, I will examine the relationship between knowledge and trust in the web from an epistemological point of view, focusing on recommender systems to elucidate my claims. I will argue that as soon as knowledge is regarded to be the result of socio-epistemic practices, as is the case on the web, epistemology has to meet ethics and politics in analyzing and amending these practices. In the second part of this paper, I will introduce MyChoice and traffic lights of trustworthiness as widgets to be included into social software applications. Their goal is to raise epistemological as well as ethical and political awareness among its users about the impact of - possibly implicit or minor - programming decisions on the information they obtain and on epistemic justice. I will conclude by showing how such widgets can enhance critical awareness and reflection among users while empowering them to make informed, context-dependent epistemic choices.

Keywords: MyChoice, ethics, epistemology, trust, recommender systems, epistemic injustice, feminist epistemology, trust propagation, transparency, reflexivity

* *Judith Simon* is a fellow at the Institut Jean Nicod in Paris since April 2009. Previously, she has worked at the Department of Philosophy, University of Vienna for four years and spent two years in technology assessment of biomedicine in Berlin. Her current areas of interest are concepts of knowledge, trust and sociality in epistemology and social software. She holds a Master in Psychology from the Free University Berlin and has been a visiting scholar at Stanford University's Forum on Contemporary Europe for half a year in 2008.

Introduction: Epistemic Trust - Trusting to Know

Trust in Science

Trust has been a topic in ethics (McLeod 2006, Baier 1992) for quite some time before receiving attention in epistemology, i.e. the philosophical discipline concerned with the process of knowing and criteria for knowledge. In his seminal paper "The role of trust in knowledge" (1991), John Hardwig exposes the function of trust for knowledge creation in science. He states that in classical epistemology, knowledge and trust are conceptualized antithetically: we trust when we do not know; when we know, we do not have to trust. However, "[m]odern knowers", he argues "[...] cannot be independent and self-reliant, not even in their own fields of specialization" (Hardwig 1991, p.693). His analysis departs from the observation that the majority of research is nowadays conducted in teams and he presents two examples of major scientific achievements in physics and mathematics as case studies in support of his claims. Co-operation in science is supposedly needed to overcome time pressure and to handle rising specialization. As a consequence, in scientific co-operations scientists have to trust the competency and the honesty of their colleagues, because they do not only lack the time to perform every subtask of their research on their own, but mostly they also lack the necessary expertise in the respective area of research. Thus, in order to successfully operate in science, scientists need to assess their colleagues not only epistemically but also morally.

Unfortunately, such assessment is not immune against injustice and the attribution of trustworthiness is frequently influenced by social categories, such as gender, class or race as feminist epistemologists have shown (e.g. Scheman 2001, Alcoff 2001).

Accordingly, the task for a sound epistemological analysis of the attribution of trustworthiness should be twofold. First, reasonable analyses of epistemic trust have to address the empirical question of whether trustworthiness actually is attributed adequately and fairly. However, the assessment of adequacy and fairness depend upon a prior discussion of different possible ideals of rational assessment and attribution of trustworthiness. Secondly, based on these analyses, epistemologists have to develop normative standards of how epistemologically just attributions of trustworthiness can be achieved. Since even an empirically informed epistemology proper can only show differences in how epistemic trustworthiness is applied, ethical considerations have to be taken into account to decide upon what epistemic justice is supposed to be. Thus, the development of normative standards and procedures to enforce them depends on a) the epistemological, ethical and political discussion about different possible forms of epistemic justice and b) empirical data about forms of epistemic injustice rooted in current

epistemic practices. Elizabeth Fricker has offered quite recently an interesting analysis of the ethical dimension of epistemic practices. In her book «Epistemic Injustice» she differentiates two forms of epistemic injustice, testimonial and hermeneutical injustice, and develops normative standards to ensure that these epistemic injustices can be countered (Fricker 2007).

Trust in Everyday Life and the World Wide Web

If we compare those insights about scientific reasoning to reasoning in everyday life, we soon realize that similar processes of knowing, trusting and attributing trustworthiness take place there as well. Taking into account how many decisions you have to make in your daily life, where you lack the necessary expertise, it should become quite obvious that trusting others for epistemic purposes, trusting others to know is an extremely prevalent phenomenon pervading all possible aspects of our lives: we ask our doctors about possible illnesses, mechanics about our cars, our insurance brokers about necessary insurances - and we will probably cross-check with other offers, different agents and maybe some independent agencies. The effort that we put into cross-checking will depend on what is at stakes: potential loss of money, potential danger to life or health, e.tc.

Similar processes occur when you search for information on the World Wide Web. You might have a default to trust resources on the web, but depending on the topic, the stakes, the spare time you have at the moment, you will spend more or less effort on finding supporting or contradiction information elsewhere. For instance, if you are interested in the weather conditions in Vienna at the moment, you might use a search engine. Maybe you will simply trust the first information you get shown on the top of your search results. If you are a bit more suspicious, you might check, whether the information was provided by some agency trustworthy with respect to weather forecast^{1,2}.

Trust and Recommender Systems

Let's turn to a specific aspect of trust and social software. There are two very distinct approaches towards concetualizing the relationship between trust and software in web science. The field that I am not going to deal with is trust in software, which focuses especially on security and privacy issues. In the following I will focus exclusively on trust in other people via software and use trust-aware recommender systems (RS) to exemplify my claims. RSs in general are systems

1. An aspect I want to stress here is that trust concerning competency is of course context-specific. You might trust your mechanic on his opinion about the brakes of your car, but probably not about the best treatment for your liver disease.

2. Please confer: <http://www1.epinions.com/about> [last access: 16.4.2009].

that suggest new items to users, which he or she might like (e.g. books, music, etc). Classical RS techniques have several shortcomings, including the so-called cold start problem, i.e. the difficulty to generate recommendations for new users.

When a new user enters a system, the system does not “know” anything about this new user and this ignorance makes it difficult to generate appropriate recommendations for her. To counteract this problem, traditionally, new users have been asked to rate a few items so that the system can “learn” something about the user in order to provide personalized information on interesting items for her. However, especially in large databases necessary correlations are scarce and thus, this procedure often turns out to be quite ineffective. In consequence, Massa & Bhattachasjee (2004) have developed an algorithm for “Trust-aware Recommender Systems”, arguing that this problem can be solved by implementing a notion of trust between users into the system (Massa & Bhattachasjee 2004). The difference between traditional RSs and trust-aware RSs is quite simple: “[w]hile traditional RSs exploit only ratings provided by users about items, Trust-aware Recommender Systems let the user express also trust statements, i.e. their subjective opinions about the usefulness of other users” (Massa & Avesani 2006). This seemingly minor change proves to be highly effective to remedy the cold start problem because “it is able to exploit trust propagation over the trust network by means of a trust metric” (Massa & Avesani 2006).

It becomes especially obvious that the work of Massa and his colleagues is also interesting from an ethical point of view in the case of “controversial users”. It is here that the underlying values and possible consequences of different trust metrics become visible. In an empirical study, Massa & Avesani (2007) analyzed data from Epinions.com, a web site, where people can publish reviews about a variety of products and rate reviews of others. The goal of Epinions.com according to its self-description is to help “[...] people make informed buying decisions. It is a [...] reliable source for valuable consumer insight, unbiased advice, in-depth product evaluations and personalized recommendations”³. Users of Epinions.com can assign binary trust statement to other users, indicating whether they in principle trust or distrust their reviews. This process leads to webs of trust. Controversial users are users that receive diverging trust statements from other users of the community, i.e. many users trust them while many others express their distrust in them. Trust metrics are techniques for answering questions such as “Should I trust this person?” in virtual communities and in this inquiry they tackle the philosophical question of whether trust is warranted in a certain situation.

3. Please confer <http://wikidashboard.parc.com/> [last access: 16.4.2009].

In their analyses, Massa and Avesani distinguish between global and local trust metrics and define them as follows: “[while] [g]lobal trust metrics assign to a given user a unique trust score, the same independently of the user that is evaluating the other user’s trustworthiness [...], a local trust metric provides a personalized trust score that depends on the point of view of the evaluating user.” (Massa & Avesani 2007, p. 40).

It is of epistemological, ethical and political interest that controversial users are valued very differently in these two different metrics. Local trust metrics explicitly stress and appreciate the individuality and situatedness of every trust statement and state that controversial users by definition do not have a global trust value for the whole community. By contrast, global trust metrics suggest a fictitious consensus between users by calculating an averaged trust value for each user. Through this process, the controversial user is rendered “unreliable” and gets statistically eliminated.

These different types of trust metrics do not only have different underlying assumptions about the value of those users and about deviation from the mean - or norm - more generally. They also have an impact on which information you receive and whose opinions are included. And they might even retroact on cultural and societal values on how to deal with minority views. Averaging out controversial users by means of statistics has a similar effect as other mechanisms of sorting out (Bowker & Star 1999) and silencing: they exclude those from participation that deviate too much from the norms or do not fit in ready-made categories.

Trust and Wikipedia

The second major source of inspiration for this paper and the development of my own widget in the next section have been the works of Ed Chi and his colleagues of the Socially Augmented Cognition Group at the Palo Alto Research Center (PARC). In the following, I will give a brief description of one of their tools and the related experiments they have conducted concerning the relationship between social transparency, accountability and trustworthiness on the web.

Ed H. Chi, Aniket Kittur, Brian A. Pendleton and Bongwon Suh from PARC have developed a tool called WikiDashboard⁴ «[...] that visualizes the social dynamics and editing patterns of every article and editor of Wikipedia» (Chi, Suh & Kittur 2008). By unveiling the evolvement of articles and the role and amount of conflict, this tool is aimed at raising the social transparency and accountability and

4. Deciding on a localized indicator or an almost universal one are two options you can choose from. Even though I prefer the universal traffic light in this example, I do not want to suggest that this is always the better option. The best alternative might depend on the context and has to be decided case by case.

by this the trustworthiness of Wikipedia (Suh, Chi, Kittur & Pendleton 2008). The basic idea behind the development of this tool was that the fact that anyone can edit any Wikipedia article does not necessarily have to be regarded as a threat to reliability, but also as a source for it. Suh et al. (2008) argue that it is precisely the possibility to put ideas into discussion, to examine and challenge each others' claims, that is crucial for knowledge generation in science and that similar processes also occur on the web. Reliability and growth of knowledge might thus be advanced by discussions and mutual criticism in combination with practices increasing social transparency, such as attribution, indication of past performance and provision of sources. How relevant the revealing of sources is for an epistemically valid attribution of trust, the assessment of the quality and the identification of the potential bias of information was shown by the WikiScanner (wikiscanner.virgil.gr). By tracking the IP addresses of anonymous editors, this tool unveiled that numerous organizations were editing a diversity of Wikipedia articles anonymously in a way that served their particular interests.

WikiDashboards exist for users and for articles. Thus the editing activity of a specific user or a specific article is visualized and can be used as a cue for assessing the trustworthiness of an article at a given moment or as a proxy for the trustworthiness of a user. The WikiDashboard embedded within each article of Wikipedia is intended to make to user aware of interesting editing patterns, he might otherwise not notice. Examples would be sudden bursts of edits due to recent events in the case of articles. WikiDashboards on user sites might indicate the user's specific editing habits as well as the range and variety of topics she has contributed to.

Referring to theories of social translucence (Erickson, et al. 2002, quoted from Suh et al. 2008), they consider three things essential for effective communication and collaboration: «[...] making socially significant information visible and salient; supporting awareness of the rules and constraints governing the system; and supporting accountability for actions.» (Suh et al. 2008, p. 1039). They conclude that the WikiDashboard might be a useful tool for supporting social translucence and that it might not only benefit the readers in trying to assess the trustworthiness of a Wikipedia article or user, but that it might also have an impact on the behavior of editors and authors.

What I am particularly interested in is the role, visualization can play for rational - and just - attribution of trustworthiness and its benefit for informed decisions about the value of information on the web. It is not some fancy visualizations that I have in mind, but rather the basic process of rendering things visible. And this is often a matter of degree and location: what is how visible for whom?

For instance, the discussion and history pages of Wikipedia, which serve as input data for Chi et al.'s (2008) tool are in principle easily accessible to all users of

Wikipedia. However, many people do not look at these pages. Maybe because it is too much effort; maybe because they do not want to be bothered; maybe because they do not understand the interface; maybe because they are overwhelmed by the sheer number of revisions or discussion entries; and maybe also because they do not understand how this information might be useful to them. It is especially for this very last fraction of users that the WikiDashboard as well as the application that I have conceptualized might be useful (cf. section 2). However, even some of those from the 'I don't care'-category might be turned into 'Well, now-I'm-interested', if the indication of trustworthiness was a salient feature of the website and if it had proven useful to them in their first few trials.

So the tool developed at PARC should be regarded not simply as just another web2.0 widget. Rather it should be seen as a tool for empowerment, a tool that raises awareness about the functioning - and possible malfunctioning - of a system that is widely used for information gathering. The skeptic, and especially a Wikipedia skeptic, might still ask why such a tool should be useful. Please allow me to take you on a brief excursus.

Excursus: A personal note on using Wikipedia and self-observation as tools for reasoning

It is quite often very useful and informative to take a look at one's very own methods and practices of information searching, knowledge acquisition and trust attribution in order to get a clearer view on epistemological problems. When reading epistemological papers, articles on knowledge and trust on the web as well as apocalyptic prognoses about the future of human knowledge, I frequently get the impression that authors assume that they are among the only ones who reason - at least approximately - accurately. I think there is no reason to be that pessimistic; the epistemic situation of humankind is in all probability not quite as disastrous as critics and worriers often want to make us believe.

Wikipedia for that matter has many critics, among them philosophical, pedagogical, economic and political thinkers and controverters (e.g., Keen 2008, Sanger 2009, Waters 2007). Every once in a while, there are even political cries to ban Wikipedia and often it seems that some critics tend to throw the baby out with the bathwater. Despite the fact that I do not think efforts, such as banning Wikipedia from academia are enforceable, I do also think that these reactions are neither reasonable nor are they based on adequate empirical knowledge about the actual usages of Wikipedia. I use Wikipedia frequently myself and I think it is a very valuable tool. So, the issue should not be whether to use it, but rather how to use it.

According to some of its critics, Wikipedia is predominantly used with the default of blind trust. I do not agree with this speculation unless someone has provided convincing evidence that indeed the majority of users trusts Wikipedia more blindly than other, similar resources. Rather, I consider it to be paternalistic and arrogant to assume that all other users of Wikipedia are less skilled than I am myself in assessing the quality of information. I am convinced that most users assess - however tentatively and provisionally - the reliability and trustworthiness of the information they receive on a certain topic, at a certain time from Wikipedia (or any other source if you wish). We all have to rely on proxies and indicators of reliability (such as overlap with aspects of the topic we already knew before, e.tc.).

However, depending on what is at stake, I will raise my epistemic standards of accepting information (cf. also Origgi 2008). If I have a lot to lose, I will probably look for supporting or contradicting information elsewhere and if I am still not convinced I will reject the information or - if possible - suspend judgment. For instance, if I want to have a first idea about what might be causes of a minor skin irritation, I might search Wikipedia or just type the symptoms into Google. It might actually really be the case - as an example of 'epistemic luck' (cf. Pritchard 2005) - that I find the correct information and maybe even a solution to relieve me from the itching with this random search. However, except from rare occasions, (e.g. I am alone in the middle of nowhere or the last human being on earth, who has only access to the internet, but not to other people or experts) will I rely solely on the output of Google. And I will even less so be satisfied with this information, when I experience more serious symptoms. I have oversimplified things in this example. I might for instance trust information on an illness that is provided by a medical department of a university more than some random website without any institutional 'voucher'. When it comes to trusting information about effects and especially side effects of medication, however, I might actually end up putting more trust into the newsgroup of a patients' association than on the medical department's website or even the supposed expertise of a pharmaceutical company selling this medication.

Since I assume that I am not a rare example of a rationality and that other people deliberate what to believe as well, there seems to be hope and no need to become desperate about the future of human knowledge and knowing. However, there is of course, room for improvement of epistemic practices as well as of software. There is a lot of work to do to raise transparency and accountability, to support education, reflection and empowerment. And hope the widget depicted in the next section will make a contribution to this effort.

Developing Tools for Empowerment and Reflexivity

The goal of this second part will be to conceptualize tools that accommodate for all the insights gained from the Part 1: the relationship between knowledge and trust in science, everyday-life and the web in particular; the danger of unjust attribution of trustworthiness; the epistemological duty to empower people to make informed decisions about which information they want to trust how much in which context.

The fact that one of the definite books on usability design is entitled «Don't make me think» (Krug 2005) is certainly not very encouraging for an epistemologist trying to improve the epistemological awareness among users by implementing a new widget. However, as noted before, I prefer to stay optimistic about the willingness of people to think. If they can see a clear advantage of a new feature that is easy and intuitive to use and does not cause them much extra hassle, most people I know appreciate epistemic support. Thus, instead of being pessimistic, the lesson learned from the success of this book should rather be that when developing a new tool, simplicity and ease of use as well as perceived usefulness for the user have to be taken into account.

Should I Stay or Should I Go? Traffic Lights of Trustworthiness

So what are the conclusions that can be drawn from my previous analyses for the development of a tool to improve the information seeking behaviour on the web both epistemologically and ethically? I think that the WikiDashboard is a good example of a tool that empowers its users and enables them to make more informed decisions about the information they are being provided (Suh et al. 2008). However, I also think that the interface of the WikiDashboard is quite complex and that a majority of users might not use it, because they might feel still overwhelmed by the huge amount of data provided even in this aggregated format. So what I am interested in is whether there might be even simpler tools that still raise the epistemological and ethical reflexivity of many differently skilled and interested users.

Let's for instance assume that you see a traffic light either in red, yellow or green on the top of each Wikipedia article. Wouldn't this make you stop and wonder for a second what this signal is trying to tell you? Whether you should treat the information of the article differently depending on whether the lights are on red or green? The reason why I chose the traffic light is that it is an almost universally recognized warning system. Given how widely used Wikipedia is around the globe, the only alternative would be to localize the symbol you intend to use as a cue for trustworthiness depending on the language of Wikipedia. However, if you only take into account in how many countries English is spoken and in how

many more the English Wikipedia site is used, e.g. because the version in their own language is much smaller, the potential problems around localization become clearer⁵.

The traffic lights are simply a much more condensed and intuitive format than the rather complex interface of the WikiDashboard. The traffic light signal would be a dynamic, automatically generated indicator of controversy which preceded the temporal fixation of the article of Wikipedia at the moment you read it. This indicator can then be used as a proxy for the trustworthiness of the current state of the article⁶. That this indicator is dynamic and automatically generated is crucial for any website as dynamic as Wikipedia. If the traffic lights cannot potentially change with each revision of the article, they would soon be rendered unreliable and thus useless as indicators of trustworthiness of a potentially constantly changing article. But this is exactly what the user is interested in: a quick assessment of the quality of the article they see at the moment they see it⁷.

MyChoice: Empowering users, raising reflexivity of users and developers

The second tool I have conceptualized in more detail is as simple as the traffic light example, but it is meant to amend searches for recommendations. I have labelled my widget MyChoice for two reasons. First it is supposed to indicate that epistemologically and ethically relevant choices are constantly being made in the course of developing software. Thus, by the label MyChoice, users should be made aware that these decisions are built into software and have certain effects. Secondly, MyChoice is a tool that empowers users to make informed choices of their own where it is normally the programmer who has decided for them.

MyChoice comprises of a dual search-function and a visualization option and can be implemented in principle into all websites using recommender systems. In the following, I will explain the features of MyChoice and show how they correspond to the previous epistemological and ethical considerations of this paper. I have also sketched a user interface to illustrate this tool. For reasons of exemplification, I have used the epinions.com-website as a background into which I

5. I am aware that the level of controversy is only one possible cue for assessing trustworthiness. Thus it would be possible and plausible to aggregate different algorithms and merge them into the ternary symbol of the traffic light.

6. Others who are more interested and would check the history and discussion pages anyway are not the target group of such an application and also not those who do not even care about the traffic lights, because they either blindly trust or distrust or randomly trust the information provided on Wikipedia. But I would assume that remaining group of users is quite large, so that the development of such tools has a good chance of getting used.

7. This is true at least for those RSs that allow for different trust metrics.

have inserted the new features of MyChoice. In figure 1 you can see the normal starting page of Epinions.com.



Figure 1: Screenshot taken from the website Epinions.com [<http://www.epinions.com/>;
Date of access: 16.04.2009]

MyChoice is depicted in figure 2. For the purpose of illustration, I have superimposed the features of MyChoice on the background of the Epinions.com starting page.

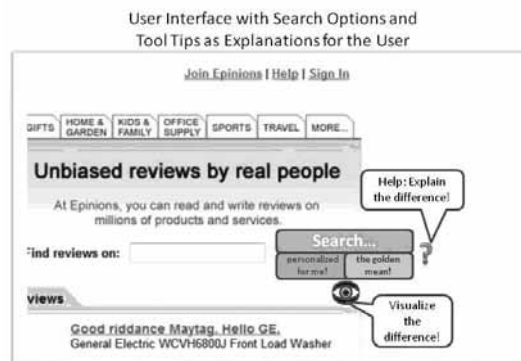


Figure 2: Sketch of the user interface of MyChoice with search options
and tool tips as explanations for the users]

Basically, MyChoice has two distinct features. One is a dual search button, by which you can choose between two different trust metrics to generate recommendations for your search query. The labels that I have chosen are "Search... personalized for me" versus "Search... the golden mean". They correspond to the local and global trust metrics as described in Massa & Avesani (2007), but other metrics are also possible. You can decide which trust metric you want to set as your

default, however, by clicking on one of the two buttons, you can change this for each new query.

The second feature is an “open-eye”-button. If you click on this button, the differences between the two trust metrics are visualized. There are various options on how to visualize this difference. The simplest version would be to just show the different search results next to each user (similar to the function of comparing versions in Wikipedia). However, I would opt for a graphic display that is more visually stimulating and intuitive. It should be possible to display – at least in parts - the different trust propagation patterns. The resulting displays should resemble social networks graphics. Since those graphics are by now embedded into many popular websites, such as Facebook.com, e.t.c., I would assume that many users will be familiar with this type of graphic display. The crucial point will be to show the differences that result from using the local as opposed to the global trust metric. The simplest option might be to show both graphics next to each other. However, it might be more instructive to superimpose the two graphics. Or it would be possible to switch slowly between the two metrics, so that differences become dynamically visible and the users can follow the changes.

In the end, there are different visualization options and it would be necessary to develop different prototypes and test them for their respective usability. It might be even possible to let the user decide upon her preferred mode of visualization. However, I think one should avoid providing too many choices, because this might eventually alienate some users, especially the less experienced ones.

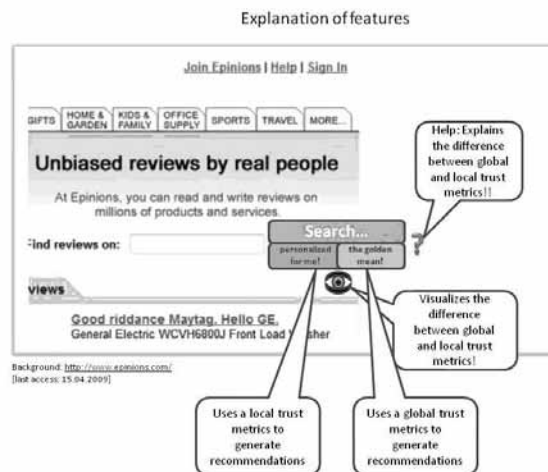


Figure 3: MyChoice: explanations of its features

Another feature that should be included into the graphic displays is the use different colour markers for different types and groups of users. It might be informative to see how different user types change in their positions and their impact depending on the metrics used. The fate of the controversial user in global trust metrics has been depicted before, but there are other possible examples. What happens to the user who always rates each product she evaluates with the highest possible value? What about her picky counterpart? How are these biases accounted for? What happens to the user that very rarely writes reviews, but those he writes are highly valued but many users? What about his prolific neighbour, who writes five reviews a day, one as useless as the other? Highlighting those groups of users who are most affected by changes in trust metrics will render the discriminative consequences of different metrics visible. Moreover, this information about different users and user types might not only be illuminative for the users, but also for developers trying to improve their metrics and algorithms.

For the moment, I can only hint at different visualization options, which have to be developed and tested in the future. The guiding principle should be that the resulting visualizations should be as instructive and informative as possible while remaining usable and intuitive. Thus, over-complex as well as under-complex graphic displays should be avoided.

So what would be the utility of such a tool? I assume that once people start using this tool and become aware of the differences in results that are caused by a simple click on a differently coloured search button, they will start to think about this. They will realize that different metrics and algorithms can be used and are used in different applications. And that these different metrics have an impact on which information they receive, whose voice is heard and who is automatically silenced or sorted out. The presence of the two search buttons and the visualization of differences will raise awareness about the huge consequences of seemingly minor programming decisions on information retrieval and epistemic justice.

MyChoice is also a tool for empowerment, enabling the user to make more informed decisions about the information she wants to receive. For instance, it would be possible to switch between the global and the local metric depending on the context. Users might opt for the more situated, local option when looking for movie recommendations, but for the rather universalist, global one when they want to learn something basic about computing, statistics or gardening. In fields where users are novices or for certain reasons more interested in mainstream recommendations, they might press the “Golden Mean”-button. And for other questions, they might prefer the “Personalized For Me”-search. In the end this decision is up to the user. But I think that such a simple tool as a button to decide which trust metric to use and have the differences displayed would be ex-

tremely valuable. By using MyChoice, users will learn about the functioning and the consequences of different metrics and algorithms. And this will have positive epistemological and ethical consequences.

Let me summarize the possible benefits of MyChoice. Besides enabling the user to decide on a case-to-case basis which metric he prefers depending on the context, MyChoice also has several pedagogic functions. People using this system would on the one hand be empowered to decide upon which metric they want to use, i.e. they have a choice of which information they want to be displayed. Moreover, people using this system would become much more aware of how implicit assumptions and values of the programmers are inscribed into technology, i.e. they would realize the impact of certain programming decisions of the retrieval of information. This effect would already be triggered by the dual search option, but it would be amended by the visualization of the different search mechanism and differences of retrieved information. Knowing what happens to controversial users once you start averaging people out that deviate too far from the norm, might also raise critical awareness of processes of silencing minorities more generally. Just remember that depending on the context, controversial users might just as well be labelled «pointed users».

MyChoice renders visible the possibly far-reaching consequences of seemingly minor programming decisions implemented into systems. Further, if the use of tools like MyChoice becomes more widespread; if people use it, because they like this option and see its benefit for their usage, this would possibly lead to a re-thinking and more awareness among programmers and software developers as well. It should be an epistemological goal to empower people to make rational decisions about how much trust to put on information they receive in general and on the web in particular. And it is an epistemological and ethical duty to raise awareness about epistemic justice and to provide tools that support it. I think MyChoice is a starting point for this endeavour and I hope it will ignite more and possibly different reflection about decisions made in software development and their epistemological, ethical and political consequences.

Conclusions

I hope to have shown that the notion of epistemic trust is a topic where epistemology has to meet ethics. Assessing the quality of information, deciding whom to trust and whom to distrust is not limited to information obtained on the WWW. However, it becomes all the more obvious in an environment in which information can be exchanged with high speed over long distances, enormously increasing the amount of interactions with people we do not know personally but whom we have to trust – or decide to distrust – nonetheless. Taking these

developments into account, a thorough ethical, epistemological and political analysis of the relationship between trust and knowledge should be indispensable – for philosophy just as much as for web science and software development. Simple widgets, small changes in programming can have huge epistemological, ethical and political consequences. To give inspiration and concrete ideas for the development of normative standards as well as tools that improve epistemological as well as social and political justice on the web and in society more generally should be a central task for contemporary epistemologists. I hope this paper convinced you both of the necessity and feasibility of such an endeavour and that it serves itself as a first small contribution to it.

Acknowledgements

I would like to thank the three anonymous reviewers as well as Noah Holtwische for their very encouraging and constructive feedback on earlier drafts of this paper. The research for this paper was enabled by two grants: the ANR2008 grant (Agence Nationale de la Recherche, France) CSOSG- CAHORS for a Project on «Information Evaluation, Analysis, Organization and Ontologies for Intelligence and Security» and a research scholarship from the University of Vienna, Austria for a project on notions of knowledge, sociality and trust in social epistemology and social software (Project number: F-405).

REFERENCES

- Alcoff, L. M.** (2001). On Judging Epistemic Credibility: Is Social Identity Relevant? Engendering Rationalities. N. Tuana and S. Morgen. Albany, SUNY Press: 53-80.
- Baier, A. C.** (1992). Trust. The Tanner Lectures on Human Values, Princeton University 6-8.3.1991. available at: <http://www.tannerlectures.utah.edu/lectures/documents/baier92.pdf>. Last access: 15.04.2009
- Bowker, G. C. & Star S. L.** (1999). Sorting Things Out: Classification and Its Consequences. Cambridge, MIT Press.
- Chi, E. H., B. Suh,** et al. (2008). Providing social transparency through visualizations in Wikipedia. Social Data Analysis Workshop at CHI 2008, Florence, Italy.
- Fricker, M.** (2007). Epistemic Injustice. Power and the Ethics of Knowing. Oxford, Oxford University Press.
- Hardwig, J.** (1991). The role of trust in knowledge. The Journal of Philosophy 88(12): 693-708.

- Keen, A.** (2008). *The Cult of the Amateur*. New York, Doubleday.
- Krug, S.** (2005). *Don't Make Me Think! A Common Sense Approach to Web Usability*, New Riders.
- Massa, P. & Bhattacharjee, B.** (2004). Using Trust in Recommender Systems: an Experimental Analysis. Proceedings of iTrust2004, International Conference.
- Massa, P. and P. Avesani** (2006). Trust-aware Bootstrapping of Recommender Systems. Proceedings of ECAI, Riva del Garda, Italy.
- Massa, P. and P. Avesani** (2007). «Trust metrics on controversial users: balancing between tyranny of the majority and echo chambers.» *International Journal on Semantic Web and Information Systems* 3(1): 39-64.
- McLeod, C.** (2006). Trust. *Stanford Encyclopedia of Philosophy*. Edward N. Zalta (ed.), URL = < <http://plato.stanford.edu/entries/trust/> > [last access: 15.4.2009]
- Origgi, G.** (2008). Trust, authority and epistemic responsibility. *Theoria* 61: 35-44.
- Pritchard, D.** (2005). *Epistemic Luck*. Oxford, Clarendon Press.
- Sanger, L. M.** (2009). The Fate of Expertise after Wikipedia. *Episteme* 6(1): 52-73.
- Scheman, N.** (2001). Epistemology Resuscitated: Objectivity as Trustworthiness. *Engendering Rationalities*. N. Tuana and S. Morgen. Albany, SUNY Press: 23-52.
- Suh, B., E. H. Chi, et al.** (2008). Lifting the Veil: Improving Accountability and Social Transparency in Wikipedia with WikiDashboard. 26th Annual ACM Conference on Human Factors in Computing Systems, Florence, Italy, NY: ACM.
- Waters, N. L.** (2007). «Why You Can't Cite Wikipedia in My Class.» *Communications of the ACM* 50(9): 15-17.

The Rights and Duties of Lawful User in EU Copyright Law

Tatiana-Eleni Sinodinou*
Dr in Law, Attorney at law

Abstract

The emergence of the concepts of lawful user and of lawful use in European copyright law could be regarded as an additional element to the process of gradual deconstruction of the author - centered model of copyright protection. It is the first time in copyright history that the end-user of a protected work of mind is identified as a distinct entity that can claim the application of the exceptions to copyright. This phenomenon should not be overestimated as the introduction of these concepts have been established until now only to specific copyright regimes that have been introduced by the European Directives for computer programs and databases and, thus, they have not become part of the horizontal harmonization of copyright law. However, the discrete arrival of the lawful user in the field of copyright law has not only terminological or purely theoretical significance. It is accompanied by a new more dynamic dimension in the way copyright exceptions could be interpreted and applied. The recognition of an individualised person that could claim the application of copyright exceptions and the reinforcement of certain copyright exceptions in the sense that they have been shielded against the risk of their contractual exclusion marks a mutation to their legal nature. Instead of being qualified as simple spaces of freedom of use with a priory defending and passive character, certain copyright exceptions are enhanced in order to approach the legal nature of rights. The aim of this paper is to highlight and analyze this evolution and to contribute to the definition of the concept of lawful user. This analysis opts for a flexible approach of the concept of lawful user that could be constructed not only on the basis of specific legal provisions and contractual

* *Attorney-at-law, Bar Office of Thessaloniki, Greece. Bachelor (U.B) of Law, Law Faculty of Aristotle University (Thessaloniki, Greece). LL.M in Mass Media Law. (D.E.A. Droit des médias), Law Faculty Aix-Marseille III, (Aix-en-Provence, France). PhD in Copyright Law, Law Faculty of Aristotle University (Thessaloniki, Greece). Postdoctoral studies in Copyright Law, Aristotle University (Thessaloniki, Greece). Visiting Lecturer to University of Macedonia (Thessaloniki, Greece). Author of three books and many essays in the fields of copyright law, information law and commercial law.*

terms, but also on the basis of the interpretation of the contract under the spectre of good faith and fair practice.

Keywords: Lawful user, lawful use, lawful acquisition, computer program, database, exceptions, limitations, three-step test, good faith, fair practice, information ethics, copyright harmonization, “rights” of the lawful user.

Introduction

The research of the position of lawful user in copyright law could be considered as heretic in EU copyright law. Copyright doctrine is characterized by the absence of the user (Cohen, 347). This controversy stems mainly from the dominant author – centered approach of European continental copyright law (the so called “author’s right” approach). According to this approach, the natural person of the author of the intellectual creation is the cornerstone of the awarded protection. Public interest is satisfied by the instauration of strictly defined exceptions or limitations to copyright (Lucas/Lucas, p. 33, n°30, Lepage, p. 5, Strowel, p. 20-21). These exceptions or limitations are not granted in favor of a legally recognized individual entity, but in a general and abstract way in favor of the public. In other words, the end- user of the works of intellect is not recognized as an individual entity that can claim the application of exceptions or limitations to copyright. Moreover, exceptions or limitations are not traditionally considered as rights of the end-users (Lucas/Lucas, p. 260, n°22).

This paper analyzes the introduction of the concept of “lawful user” in European Intellectual property law and the recognition of a new category of rights, the so-called “rights of the lawful user”.

The paper is organized as follows. In the first part, the analysis will focus on the emergence of the concept of lawful user in copyright law and the recognition of the rights of the lawful user due to the harmonization of EU copyright law. Furthermore, the specific rights of the lawful user will be examined in order to find their legal nature and their real extent.

The second part deals with the recognition of some limits to the exercise of the exceptions to copyright and to “rights of the lawful user” introduced by EU Directives. Two major points will be used for the delimitation of exceptions to copyright and of the rights of the lawful user. The first one is the definition of the concepts of “lawful user” and “lawful use” of the works of mind and the second one is the application of the three step- test.

The emergence of the concept of lawful user in the eu copyright law

The introduction of the concept of lawful user and the establishment of the rights of the lawful user constitutes a legal evolution that signifies a profound mutation of the European continental tradition of copyright law.

The technological evolution has been one of the factors that accelerated this mutation. The basic characteristic of this change has been the gradual transformation of intellectual property law from a legal mechanism of protection of the author to a legal instrument of protection of the investments made by the cultural industries. In order to better understand the significance of this mutation it is crucial to make a reference to its legal and social background.

The “depersonalization” of continental copyright law in the new economy

According to the traditional approach of the continental “droit d’auteur” system of protection (the “author’s right” approach), the end-user of the work of mind is not recognized as an individual entity and does not possess any individual right to the use of the work (Lucas, 260). The end-user can use the work only in the context of strictly defined limits that have been established by the author or the producer (Guibault, 17).

The legal situation of the user could be resumed as following: in order to use the work the user has to obtain the express authorization of the author or to claim the application of one of the exceptions to copyright which are established by national copyright legislations. If her behavior is not covered by the author’s authorization or by a precise legal exception or limitation to copyright she violates copyright law.

New technologies have brought significant changes to copyright law. In the digital information society, works of mind are emancipated from their physical support and can be easily and rapidly distributed worldwide via Internet. However, the emancipation of the works from the natural person of the author is even more important. The organic bond between the author and the work is often devaluated. The works of mind tend to be considered less as intellectual creations that derive from the unique personality of their author and more as simple consumer products. The transaction of intellectual goods gains more importance than their cultural value. This dominant tendency marks a highly progressive “depersonalization” of intellectual property law.

The alienation of the creation from its source takes also other forms. The author gradually loses the control over the exploitation of the intellectual creation. The exploitation of the works of mind is often exercised by the producers or by pow-

erful and impersonal collective societies that impose their own rules not only to the users but also to the author (Boyhta, 384, Gaudrat, 85). In that context, the author is obliged to stay to the margin and cannot follow or influence the life and the evolution of the intellectual creation.

At the same time, it is undeniable that technology offers many new possibilities for the infringement of copyright law. The use of peer-to-peer technology and social networks for the infringement of copyright law are two significant examples of the wide use of technology against the interests of authors and copyright holders. Free unauthorized use of intellectual creations has become a trend for Internet users who seek to justify illegal file-sharing with arguments based on their right to privacy and on freedom of expression (Biegala, p.74).

The legal response to the technological threat was the reinforcement of copyright in international and EU level. The prerogatives of the author were broadened or re-defined in order to cover new ways of exploitation of the works of mind. In some special cases, such as in the database sector, specific regimes made their appearance. Secondly, a new layer of protection was added to the legal monopoly of the author and to the pure technological protection: the legal protection of technological measures of protection. In that context, the circumvention of technological measures of protection and the import, manufacture, sale and distribution of devices and the provision of services which are primarily used for the circumvention of technological measures of protection were recognized as illegal acts. This layer of protection in combination with an extremely broad definition of the scope of the right of reproduction reveal a new reality for copyright users that is characterized by the emergence of a new right to control the access to copyrighted works of mind (Dusollier, *L'utilisation légitime de l'œuvre: un nouveau sésame pour le bénéfice des exceptions en droit d'auteur ?*, p. 18, Heide, p. 469, Koelman, p. 616)

Nowadays the reinforcement of the protection of copyright constitutes a source of permanent dispute between copyright holders and users. The latter constantly demand more exceptions and a broader freedom of use. It is in that framework that the concepts of the "lawful user" and of "lawful use" and the concept of "rights of the lawful user" made their appearance in copyright law.

The advent of the concept of "lawful user" in the "first generation" of EU copyright law Directives

We could distinguish two different steps in the process of harmonization of EU national copyright legislations by the EU institutions. The first one was the harmonization of specific issues of copyright protection. The so-called "first generation" Directives address sectorial issues of copyright harmonization as they apply

only to certain categories of intellectual creations which are strongly connected to the new technologies' sector or certain categories of rights or focus on a particular situation or address a particular feature of copyright protection (Becthold, p. 343). This was the case of the Software Directive and the Database Directive whose evident purpose, along with market integration, was that of fostering the industrial production of these valuable information goods and afford them an equal protection throughout the Community in order to avoid conditions of distorted competition (Mazzioti, p. 37). The aim served by these two "first generation" Directives was a vertical harmonization of copyright protection that considered only specific subject matter, such as computer programs and databases (Benabou, p. 237).

The "second generation" EU Directives have a larger effect, as they provide horizontal harmonization, which concerns all kinds of intellectual creations. One of the most significant pieces of the "second generation" Directives is the Directive 2001/29/CE, the so-called "Infosoc" Directive.

The concept of "lawful user" made its first appearance in the Software Directive. It reappeared five years later in the Database Directive. The affirmation of the concept in Directive 96/9/CE marks the determination of EU institutions to establish the notion of lawful user as a EU copyright norm.

The introduction of the concept of lawful user in these two Directives constitutes the expression of a new perception of the delimitation of copyright monopoly. This perception is consistent with the contractual reality of the transaction of information goods. Indeed, the use of the concept of lawful user in both Directives is strongly linked to the nature and the commercial features of the subject matter that is protected by these two "first generation" Directives.

Computer programs and electronic databases are mostly distributed through preformatted, standard and non-individualized licenses that prescribe in a very precise and restrictive manner the terms of use and provide very broad and detailed obligations for the licensee. Adhesion contracts that have been prepared by the copyright holder and agreed to by the end-user without negotiation portray a contractual framework where the end-user cannot really benefit of the contractual freedom of negotiation but he has to act in a take-it- or -leave -it basis (Guibault, p. 120).

In this context, the rights of the copyright holder tend to be considered as contractual duties of the user, while copyright exceptions are mostly regarded as contractual rights of the specific person who was licensed the right to use the computer program or the electronic database and less as a legal possibility to act that is generally awarded indifferently to all members of the public. The legal

terminology is strongly influenced by the contractual scheme that presupposes at least two distinct parties, the creditor and the debtor. In that context, the abstract notion of the public is replaced by the more specific concept of a particular entity that has the right to use the computer program or the database: the “lawful user”.

In parallel, the almost total preeminence of the contract as the legal instrument of regulation of terms of use in information goods brought to the surface the necessity to legally guarantee a minimum space of free use inside the contractual perimeter of restricted uses. This was achieved by the legal recognition of “the rights of the lawful user”.

Comparison between the terminology of the Software Directive and the Database Directive

The Directive about the legal protection of computer programs does not establish a clear terminology about the determination of the person who could claim the application of the exceptions. The term “lawful acquirer of the program” or descriptive definitions such as the “person having a right to use the computer program” or the “person having a right to use a copy of a computer program” are used indifferently in order to determinate the person who can lawfully invoke the application of copyright exceptions. On the other hand, in the Directive about the legal protection of databases the person who can claim the application of the exceptions is defined constantly as the “lawful user of a database”.

The lack of uniformity in the deployment of terms in the two Directives should not mislead to legal differentiation (See contra Bently, p. 324). In our opinion, all the terms used in the Software Directive and the term “lawful user” which is used in the Database Directive should be regarded as identical from a legal point of view. They all correspond to a sole entity that can be globally defined as “lawful user. This approach is also dictated by practical reasons. It is the only one that is consistent with the technological reality of co-existence of electronic databases and computer programs in the same medium. A differentiation in the way lawful use is evaluated in cases of joint use of computer programs and databases due to their functional interaction could probably lead to legal uncertainty.

The “rights of the lawful user” of a computer program

The Directive 91/250/CE on the legal protection of computer programs introduces for the first time the concept of the “lawful user” of a computer program. The lawful acquirer of a computer program is the only person who can take benefit from the exceptions to copyright. He can use the program in accordance to its intended purpose, he has the right to create a back-up copy of the program, to proceed to the observation, study and testing of the program and to decompile the program in order to achieve the interoperability of an independently cre-

ated program with other programs¹. In other words, the exceptions to software copyright protection cannot be invoked by every user, but only by the users who have the legal status of the “lawful user”. It is the first time that the exceptions to copyright are reserved only in favor of a restricted category of members of the public. Certain of these exceptions are considered as reinforced, in the sense that they cannot be overridden by opposite contractual agreement (See article 9 of the Software Directive).

The rights and obligations of the lawful user of a database

The concept of the lawful user reappears in the Directive 96/9/CE for the legal protection of databases. A lawful user of the database can claim the exceptions to copyright law and to *sui generis* database right that is granted to the maker of a database. Firstly, the lawful user may perform restricted acts that are necessary to access the database or for the normal use of the database. Moreover, member states have the option for providing for limitations to copyright protection in the case of reproduction for private purposes of a non-electronic database, in the case of use for the sole purpose of illustration for teaching or scientific research, in case of use for the purposes of public security or for the purposes of an administrative or judicial procedure.

The recognition of the place of lawful user is even more eminent in the framework of *sui generis* database protection. First of all, the exceptions to database *sui generis* (extraction for private purposes of the contents of a non-electronic database, extraction for the purposes of illustration for teaching or scientific research, extraction and/or re-utilization of the contents for the purposes of public security or for the purposes of an administrative or judicial procedure) right can be exercised only by lawful users. Secondly, article 8 of the Directive establishes some rules of good behavior of the lawful user or towards the lawful user that are named in the title of the article as “rights and obligations of the lawful user”.

However, not all the legal possibilities awarded to lawful users have the same legal force. As it will be indicated in the following paragraph, some of these spaces of unauthorized act as regarded less as exceptions to copyright or to *sui generis* right and more as “rights” of the lawful user. In fact, by establishing the binding nature of certain provisions of the Database Directive, article 15 of the Directive guarantees certain minimum user rights to lawful users of a database (Hugenholtz, p. 340).

1. For the analysis of these provisions see: Vivant M., *Lamy droit de l'informatique et des réseaux*, 2005, n°199-202).

The general problematic of the determination of the legal nature of copyright exceptions

Another originality of both Directives is the determination of some exceptions as obligatory. Article 9 of the Directive 91/250/CE states that any contractual provisions that limit or abrogate the possibility to create a back-up copy of a computer program, to proceed to the observation, study and testing of the program and to decompile the program in order to achieve interoperability shall be considered as null and void. Article 15 of the Directive 96/9/CE also declares the binding nature of some exceptions. Any contractual provision contrary to articles 6 par. 1 and 8 of the Directive shall be null and void.

The proclamation of these exceptions to copyright as *ius cogens* has a specific significance in the determination of their legal nature. The attribution of a mandatory character to exceptions or limitations to copyright injects a new perspective into copyright exceptions. It is a point of approach of copyright exceptions with the legal nature of “rights”². This evolution could be considered as a sign of an indirect recognition of the category of users’ rights as the essential counterbalance to copyright protection. It could also be seen as a part of a more subversive approach that considers copyright monopoly as an exception to the general principle of free use of works of mind (Benabou, p. 258).

The determination of a distinct legal subject, the person of the “lawful user”, who can claim the application of copyright exceptions and the recognition of the exceptions as “legal prerogatives” of the lawful user that cannot be overridden by the contractual will, marks the advent of a new more active approach of copyright exceptions in EU copyright law. Under the light of this evolution, apart from the concept of “lawful use” a new category of “legal prerogatives” also emerges: the “rights of the lawful user”. These “rights” could be considered as legal hybrids between exceptions and rights. Even if they can be judicially enforced, these “rights” could not be equated to the rights that are granted to authors or other copyright holders because of the dominant philosophy of the continental European copyright law.

Indeed, one of the most controversial issues of modern copyright law is the determination of the legal nature of the exceptions to copyright law. Do they grant to copyright users a simple legal possibility to act that is founded on permissive

2. Article 6 par. 4 of the Directive 2001/29/EC also opts for the partial guarantee of the exercise of some exceptions as it offers to national lawmakers the possibility to take appropriate measures to ensure that right holders make available to the beneficiary of some exceptions the means for benefiting from these exceptions to the extent necessary to benefit from them and where the beneficiary has legal access to the protected work or subject-matter concerned.

legal rules or do they grant real enforceable rights? Apart from the Belgian copyright legislation where all copyright exceptions are promoted to *ius cogens* (Dusollier, 503, Guibault, 219), no other national legislation in Europe declares all copyright exceptions as mandatory.

The way the exceptions are formulated in law could be used as an indicant about their legal nature (Caron, 273). However, it should not be overestimated. It must be combined with other elements of higher importance, such as the dominant philosophy about the place of the author and the justification of copyright law in each specific national legal tradition. Moreover, particular significance should be given to the degree of recognition in every legal tradition of the principles of freedom of contracts and of the autonomy of the parties and to the philosophical and legal justification of each separate exception (Geiger, *Droit d'auteur et droit du public à l'information, Approche de droit comparé*, p. 121, Dusollier/Pouillet/Buydens, 25).

The concept of “lawful use” in the “Infosoc” Directive

The Directive 2001/29/EC does not define as only beneficiary of the copyright exceptions the lawful user. However, the appearance of a neighboring concept, the concept of lawful use of a work of mind takes place in article 5 of the Directive 2001/29/EC. According to the provision of paragraph one of article 5 of the Directive, temporary acts of reproduction, which are transient or incidental [and] an integral and essential part of a technological process and whose sole purpose is to enable a transmission in a network between third parties by an intermediary, or a lawful use of a work or other subject matter to be made, and which have no independent economic significance, shall be exempted from the reproduction right. Recital 33 of the Directive defines the term “lawful use”. A use should be considered lawful where it is authorised by the right holder or not restricted by law.

By extending the cases of lawful use to every use that is not prohibited by law, the Directive has opted for a broad interpretation of the notion of lawful use. For example, temporary and transient or incidental copies which do not have any independent economic significance shall be covered by the exemption of paragraph one of article 5 if they are necessary in order to benefit from a copyright exception or limitation (Bechtold, 372) provided that the application of the limitation fulfils the criteria of the three-step-test. The joint delineation of the concepts of “lawful use” and “lawful user” will be conducted in the next part of this paper.

Moreover, the idea of the identification of the end-user as a distinct autonomous member of the public is also underlying in the conceptual framework of the making available right that is established in article 3 par. 1 of the Directive 2001/29/

EC. In the emerging interactive reality, the recognition of the active role of an identified end-user who defines with his choices the circumstances of individual communication of works of mind to his own space introduces an innovating vision in the classic scheme of exploitation of works of mind through mass communication conduits.

Limitations to the “rights of the lawful user”

The “rights of the lawful user” of a computer program or a database are not boundless. Inner limits deriving from the definition of the concepts of “lawful user” and of “lawful use” (3.1) and the application of the three step-test or analogous mechanisms of control (3.2) of their exercise aim to assure their proper interpretation and exercise in each specific situation.

Inner limits deriving from the definition of the concepts of “lawful user” and “lawful use”

Potential interpretations

The concept of the “lawful user” appears as a new riddle in EU copyright law. It is necessary to highlight the importance of a clear definition of the term of “lawful user”. Provided that neither of the two EU Directives set a definition of the term, the delimitation of the concept is left to the interpretative skills of legal doctrine until an unambiguous definition is established by case law. In that framework, it is obvious that in order to avoid interpretative disparities the interpretation of the term by the European Court of Justice could be more than valuable.

The concept of lawful user should be differentiated by the general concept of the end-user of the work or by the abstract concept of the public. In other words, they are not all the users “lawful users, neither the attribute of lawful user is recognized in general to every member of the public as the final unidentified recipient of protected works of mind. Consequently, essential elements for the definition of the concept of lawful user is how it is possible to become lawful user (positive definition) as well as in which cases the end user is not a lawful user (negative definition) (Vanovermeire, 65).

The concept of “lawful use” should also be distinguished from the concept of “normal use” of a database (Directive 96/9/EC, article 6 par. 1) or from the concept of “utilisation of a computer program according to its intended purpose” (Directive 91/250/EC, art. 5 par. 1). As we will demonstrate in the following paragraphs, in both cases the “lawfulness” or “the normalcy” of the use could possibly be appreciated on the basis of similar abstract conceptual criteria such as good faith, but “lawful use” and “normal use” refer to clearly defined consecutive phases of use. The evaluation of the existence of lawful use precedes the evalua-

tion of the existence of normal use. Lawful use constitutes an essential condition for the existence of “normal use” that will also eventually be determined by various factors, such as objective good faith or transaction customs.

In regard to the delineation of the concept of lawful user, two focal interpretations have been supported. According to a first doctrinal thesis, as lawful user should be defined any person that has acquired the right to use a protected work of mind on the basis of a contract, such as a license, or the person that uses the work on other legal grounds, such as on the basis of the legal exceptions to copyright or thanks to the principle of exhaustion of the right of distribution (Koumantos, 101, Dusollier, *L'utilisation légitime de l'œuvre: un nouveau sésame pour le bénéfice des exceptions en droit d'auteur*, 19, De Saint Affrique, 27).

In this framework, lawful use exists each time where the right of use is acquired through a contract regardless of the type of the contract and each time the use is not prohibited by law. This interpretation covers mainly every case of contractual lawful acquirement of a computer program or a database including acquirements through third parties and not especially directly from the right holder (Derclaye, 123), but it is in some extent broader as it does not presupposes obligatorily the existence of a contact of lawful acquirement since it also accepts the justification of lawful use on other legal grounds.

The first theory is consistent with the definition of the concept of “lawful use” in the recital 33 of the Infosoc Directive that considers as lawful use any use authorized by the right holder or not restricted by law. It is a broad interpretation that could cover a lot of different situations. In case the lawfulness of the use is justified by the lack of restriction provided by law, the comprehensiveness of the first interpretation knows some gradations. It can be narrower if the lawfulness of the use is founded on the lack of restrictions provided exclusively by copyright law, but broader if the lack of restrictions that are provided generally by law is taken into consideration. On the basis of this theory, lawful use could also exist in case of acquisition of the work of mind or of a copy of it through inheritance (Dusollier, *L'utilisation légitime de l'œuvre: un nouveau sésame pour le bénéfice des exceptions en droit d'auteur*, 18)..

This thesis presents significant advantages. For example, by founding the existence of lawful use not only on contracts but also on other legal grounds it provides a guarantee for the freedom of expression and the free trading of information and ideas. Indeed, the obligatory condition of existence of contractual authorisation for the ascertainment of the existence of lawful use could place at serious risk the fundamental principle of article 10 of the European Convention of Human Rights (Vanovermeire, 67). Nevertheless, this interpretation presents certain contextual weaknesses and it could lead to conceptual impasse, since in

case where lawful use is grounded on copyright exceptions the exception is recognized firstly as a condition of the existence of lawful use, but also simultaneously as the content of rights that are granted to lawful users exclusively thanks to the existence of lawful use. In other words, the use is lawful because there is a copyright exception, but the exception can profit only the lawful user (Koumantos, 125-126, Derclaye, 121). Nevertheless, this confusing situation does not appear in case of temporary copies that are necessary in order to benefit from a copyright exception (article 5 par. 1 of Directive 2001/29/EC) outside of the field of protection of computer programs and databases, because the prerequisite to be a lawful user in order to invoke the exceptions has only been provided in the Software and the Database Directive.

Accordingly, however, with a second interpretation, the existence of lawful use presupposes obligatorily a special agreement of concession of use which fixes at greater length the terms of use, while a use that is founded exclusively on the activation of exceptions or on other legal provisions cannot be a "lawful" one (Gaster, 38-39). The possibility of acquisition of the quality of lawful user only in case of contractual concession of the right to use the work of mind is conform with a logic of the absolute control of the use of copyrighted works and of copyright exceptions by the private will. This logic is consistent with the changes taking place in the ways of distribution of intangible goods in the interactive environment of the economy of access. However, it is foreigner to the philosophy of continental intellectual property law system, which excludes the regulation of copyright limitations and consequently of the scope of the copyright monopoly by the contract (Lucas, in: Lucas A./Devèze J./Frayssinet J., n°549).

The necessity for a comprehensive and flexible definition of the concept of lawful use

As it has been shown, these interpretations present at the same time advantages and weaknesses. In our opinion, it should be preferable to reject the most restrictive second interpretation that accepts the existence of lawful use only in case of license and to set a comprehensive and flexible definition that is based on the first theory. This definition should mainly cover uses based on every kind of contract regardless if they are written or not (license, sale or resale, donation, public lending, rental, etc) or uses by third parties justified by the existence of a contract, for example, use of the lawfully acquired employer's software by the employee inside the frame of his duties stemming from the labor contract. Uses that are justified by the existence or the aim of the contract could be those that are not prescribed explicitly by concrete contractual terms, but can be deduced by the interpretation of the contract in accordance with the implied will of the parties and in accordance with the principle of good faith.

Moreover, it should cover use of computer programs and databases justified. Jointly by contract and by other legal grounds outside the perimeter of copyright law, such as inheritance law, insolvency law or attachment. It is obvious that in most causes the conclusion of contract is necessary. However, in rare cases lawful use could be justified by other legal grounds, such as in case of lawful acquisition of the physical carrier of the work by the person who found it (see for example, article 1088 of the Greek Civil Code). On the contrary, it is obvious that every person who uses the program or the database without having lawful access, such as the hacker, the person who downloads these protected works without authorization by Internet, every person who steals a work incorporated in a tangible medium or acquires a stolen copy of it or even a person who has unauthorized access to the program in the office of a colleague shall in no circumstances regarded as a lawful user (Strowel/Derclaye, p. 226, Derclaye, p. 125).

Despite the legal dogma of the clear distinction of the intangible asset of “work of mind” from its physical carrier, the lawful acquisition or the lawful possession of a physical copy of a work could constitute a basis of lawful use. One of the most representative situations is the case of inheritance. Indeed, unless it is expressly stipulated that the decease of the licensee terminates the contract, the inheritor could normally be substituted to the user rights of the inherited according to general civil law provisions and to inheritance law. Nonetheless, lawful use by the inheritor of the licensee could be eventually be justified in some exceptional cases even if the contract between the inherited and the right holder has expressly excluded the substitution of the heir to the user rights of the inherited or if it has been stipulated that the decease of the licensee terminates the contract. This could happen, for example, in case of inheritance of a physical copy in which standardized software is integrated. In such contracts, the identity of the licensee is not a substantial element of the contract. The lawful acquisition of the copy of the software through sale grants to the acquirer of the copy the right to use the software. Under this specific contractual framework, inheritance rules could justify not only the transfer of the real property over the physical carrier of the software but also its use by the inheritor, provided that the use does not prerequisite more permanent reproductions of the program. Thus, in case of more inheritors who enjoy joint-ownership over the physical carrier of the computer program, it shall not be possible to duplicate the program in order to confer to each of the inheritors the possibility to use the program in her own personal computer. On the opposite, in case of use of an electronic database, the express exclusion of the substitution of the inheritor to the contractual rights and obligations of the inherited shall prevail and the use of the database by the inheritor shall be considered as not lawful.

Insolvency law could eventually provide a legal basis for the preservation of lawful use. For example, in a contract of license of an electronic database, in case the licensor goes bankrupt specific insolvency law provisions could provide proper solution for the continuation of the effects of the license and thus maintain the possibility to proceed to lawful use. In that case, the legal foundation of the lawful use is the contract of license and not directly insolvency law. However, by maintaining the force of the contract insolvency law succeeds to guarantee the continuation of lawful use. This is the case of the Greek Code of Insolvency Law that maintains the validity of the permanent contracts unless otherwise provided by contract or by law (Kotsiris, p. 330)³ and the case of United States Bankruptcy Code. According to Section 365(n) of this Code, in case of bankruptcy of the licensor, if the licensor rejects an executory license, the licensee has the option of treating the license as terminated by virtue of the rejection and asserting damages for breach, or retaining its rights under the license for its duration and any applicable extensions.

In the field of justification of lawful use exclusively on the basis of a contract, it is of high importance to safeguard flexibility in the way contractual terms shall be interpreted. Despite the eventual but inevitable risk of legal uncertainty, in cases of lack of precise or clear contractual provisions that set the conditions and terms of lawful use every attempt to evaluate the existence of lawful use should take into the consideration fundamental principles of private law, such as the interpretation and performance of contracts in accordance with the principle of good faith and fair practice. Interpretation of the lawfulness of use compliant with good faith and fair practice implies an element of calibration of the interests of parties, -right holder and user- that is consistent with the interpretation of copyright protection under the light of fundamental constitutional values, such as the principle of proportionality. This approach contradicts the classic denial of an eventual function of the legal mechanism of balancing of interests in continental copyright law (Against this approach see: Lepage, p. 1-2, Geiger, *L'avenir des exceptions au droit d'auteur, Observations en vue d'une nécessaire adaptation et harmonisation du système*, 2154). However, it is in line with the emergence of the concept of lawful user in copyright law and with the growing need to reconsider the current mechanisms of copyright law protection.

The standard of good faith that could be taken into consideration shall primarily be objective good faith which could be defined as an objectively accepted standard of reasonableness, while a subjective criterion that evaluates good faith on the personal inner belief of the user about his honesty and his "right to use the

3. Article 31 par. 3 of the Greek Code of Insolvency Law.

work” should be avoided⁴. Under that specter, lawful use could eventually be asserted in cases of private use, regardless if positive law establishes the private copy exception. Let’s take the example of an electronic database which is commercialized either on line through licenses or through the distribution of hard copies. In both cases the legal ground of lawful use is a contract. The Database Directive excludes the possibility of private copy exception in both cases (See article of Directive 96/9/EC). As every possible use of the database necessarily presupposes a series of provisory reproductions, the restrictive recognition of the quality of lawful user only to the licensee or the purchaser of a copy of the database would result to the denial of existence of lawful use in cases of uses of the database in the private sphere of the licensee. The wife or the children of the licensee could not be qualified as lawful users even if they use the database through the licensee’s or the purchaser’s personal computer and thus without permanent reproduction of the database. However, private use is normally outside of the scope of copyright monopoly. The inner limitation of the scope of copyright protection by the corrective application of the principle of good faith in the interpretation of the contract of license could possibly justify the existence of lawful use even in this case, despite the lack of other specific legal grounds. On the contrary, the subjective belief of a user who has acquired a copy of a computer program by a file-sharing network without the right holder’s authorization that her conduct is lawful because this is a common habitude in the circle of Internet users shall in no circumstances be deemed as good faith that could serve as a legal ground for lawful use.

This diverse version of the first theory seems to be consistent with the findings of an unsuccessful attempt of definition of the concept of lawful user that appeared in the Report of the EU Commission on the implementation and effects of Directive 91/250/EEC on the legal protection of computer programs. According to the Report, lawful user is not only a purchaser, a licensee or a renter but also a person authorized to use the program on behalf of one of the above (Report from the Commission to the Council, the European Parliament and the Economic and Social Committee on the implementation and effects of Directive 91/250/EEC on the legal protection of computer programs, par. 5).

4. As regards generally the role of good faith in copyright law, the existence of subjective good faith due to erroneous appreciation of the facts or of the law can be taken into consideration during the evaluation of the act of copyright infringement in order to dismiss the charges (Lucas/Lucas, n° 958-959), while in some jurisdictions, such as in the USA, the good faith of the infringer is expressly established by specific provisions as a “fair use defense” (see Section 504(c)(2) of chapter five of US Copyright Law).

From a doctrinal point of view, the possibility of ascertainment of lawful use on the basis of interpretation of the contract in accordance with the principle of good faith is in terminological conformity with the concept of lawful user itself. Indeed, EU authorities chose the term “lawful” user instead of the term “legal” user. The term “lawful” should be considered as referring to a situation that is not strictly defined absolutely by specific legal or contractual provisions. It refers to what is provided by law or not restricted by law or by contract. Therefore, the drafting of the conceptual borders of “lawfulness” should not fall into strict categorization, but they should be evaluated in an open-ended and flexible manner. The concept of “lawfulness” should be interpreted in a way that also takes into consideration elements of natural justice. The latter is represented in private law mainly through abstract legal principles and concepts, such as good faith (according to Ghestin, good faith is a means to achieve an ideal of justice in contracts: Ghestin J., *L’utile et le juste dans les contrats*, Dalloz 1982, chr., p. 1), while in the field of constitutional law it is represented through fundamental constitutional values, such as the principle of proportionality.

Moreover, the evaluation of the concept of lawful user on the basis of these criteria attributes to the concept of lawful user a dynamic character. In fact, defining the lawful user only on the basis of specific legal grounds (legal provisions, specific contractual terms,) is accompanied by the risk of condemning the concept to remain static. This could possibly lead to consider erroneously the obtaining for a first time legal access to a copyrighted work as the unique criterion of the evaluation of the existence of lawful use.

On the contrary, taking into consideration of these principles could be used not only to broaden the concept and avoid unjust effects, but could also function in the opposite direction as an inner limit of the lawful use itself. Indeed, good faith and fair practice could be used as criteria in order to judge if the lawful use is really lawful or if it still remains lawful. In this framework, the bad faith of a priory lawful user should be considered as a situation of “abuse of rights” and lead to the loss of the quality of lawful user. It would, consequently, result to the loss of the possibility to invoke the rights of the lawful user under certain specific circumstances. We are in presence of an example of a so-called “limitative function” of good faith⁵. Let’s take the example of a lawful acquirer,- such as a purchaser-, of a copy of a software who stores a back up copy of the software on an insecure server where everyone can have free access and thus intentionally or by negligence offers to other users of the server the possibility to reproduce the program. This user violates the principle of good faith and abuses the right to make

5. For a doctrinal distinction of the functions of good faith, see: Masse, Rapport général, in *Travaux de l’Association Henri Capitant, La bonne foi*, Paris, Litec, 1994, p. 224-227.

a back up copy of to program. Consequently, even if the lawful acquisition of the copy of the computer program has made him a lawful user, he could not be still considered as a lawful user under these specific circumstances.

The potential of application of the principle of good faith in order to effectuate a constant control of the lawfulness of the use results to indirectly insert to the evaluation of the exercise of the users' "rights" a tool of balancing of interests. It introduces an element of flexibility without deviating from the principles and mechanisms of continental private law. Good faith implies taking into consideration the legitimate interests of the other party and showing honesty in the exercise of the contractual duties and rights in a way that combines to contain communitarian values and to guarantee party autonomy (Storme, p. 3 - 4).

Except for good faith, fair practice could also be taken into account in order to affirm the "lawfulness" of the use in each specific case. Recital 22 of the Software Directive that refers to the right of decompilation is really revelatory on this point. According to this provision, "it has therefore to be considered that in these limited circumstances only, performance of the acts of reproduction and translation by or of behalf of a person having a right to use a copy of the program is legitimate and compatible with fair practice and must therefore be deemed not to require the authorization of the rightholder". Recital 22 considers that only if the specific conditions of article 6 of the Software Directive are met the performance of acts of decompilation of the program by lawful users is compatible with fair practice. Even though the provision does not aim to define the lawfulness of the use, it reveals the potential role of "fair practice" in the evaluation of the lawfulness of the use⁶. Indeed, fair practice is considered as a standard of conduct that has to be guaranteed in order to assure that the use is still lawful. Fair practice could complement the interpretation of contractual rights and duties of the user in accordance with good faith by reinforcing the objective dimension of the evaluation of existence of good faith.

6. The exigency of compatibility with fair practice is not an unknown element in copyright law. See for example article 3 of the Belgian Law on Copyright and Neighboring Rights: "the assignee shall be required to exploit the work in accordance with the fair practice of the profession". See also article 19 of Greek Copyright law 2121/1993 that provides as following: "Quotation of short extracts of a lawfully published work by an author for the purpose of providing support for a case advanced by the person making the quotation or a critique of the position of the author shall be permissible without the consent of the author and without payment, provided that the quotation is compatible with fair practice").

The role of information ethics

Besides, it seems that information ethics could also be taken into account as in most cases appear to be the deontological representation and specialization of abstract legal concepts such as good faith and fair practice. Indeed, if an extra-contractual application of the principle of good faith is rejected (for the possibility to apply the principle of good faith in the phase of the formation of the contract on the basis of an existence of an implied moral contract between two persons who enter in contact, see: Darmaisin S., *le contrat moral*, LGDJ 2000, p. 232), information ethics could serve as a safe indicator of existence of lawful use in case there is no pre- contractual or contractual relationship between the right holder and the user. Even if information ethics do not have a legal bounding effect, they can serve as helpful guidelines in order to evaluate the “lawfulness” or the “normalcy” of the use of a protected work of mind. For example, in case of an electronic database (e.g. a Web site which meets the criteria of the definition of database) which is made freely accessible by the right holder through Internet without technological protection or other restrictions, according to Internet customs every user should be considered as a lawful user even if there is no explicit consent of the right holder or any contractual relationship between the right holder and the user and no other specific legal grounds which serve as a basis for the use. In that case, the lawfulness of the use can be founded on the recognition of an implicit consent of the right holder (Derclaye, p. 125) that can be deducted or confirmed by information ethics.

Limitations to the rights of lawful user imposed by the application of the three-step test

Despite the conceptual evolution towards the legal nature of rights instead of exceptions, the legal powers granted to lawful users both in the Software and the Database Directive principally maintain a defensive function to block the legal claims of the right holders which is comparable to the function of exceptions or limitations to copyright protection. In that context, it is not surprising that the application of these rights is controlled under the three-step-test, which is a legal instrument of control of copyright exceptions. Moreover, the application of the three-step-test in the case of the “rights of lawful user” is consistent with the assessment of the three-step-test as a legal mechanism of control of abuse of rights (Sinodinou, p. 74).

Indeed, the most important key for the definition of the limits of exceptions to copyright is the three-step test. The three-step test has become an international and EU community standard for the evaluation of the exceptions to copyright law. According to the test (art. 5 (5) of Directive 2001/29/EC), the exceptions shall only be applied in certain special cases which do not conflict with a normal

exploitation of the work or other subject matter and do not unreasonably prejudice the legitimate interests of the right holder.

The possibility of control of the exercise of the “rights of lawful user” under the conditions set by the three-step test is clearly affirmed in both the Software and the Database Directive. Express references mainly to the second and the third criteria of the test appear in certain provisions of these Directives, however, in a fragmented way and without a clear systematization. Article 6 par. 3 of Directive 91/250 expressly states that the right of decompilation shall not be interpreted and applied in a manner which unreasonably prejudices the right holder’s legitimate interests or conflicts with a normal exploitation of the computer program. Directive 96/9 on the legal protection of databases sets in diverse provisions the criteria of the three-step test as guidelines for the evaluation of the application of the rights of lawful user. Article 6 par. 3 requires that any right of the lawful user established by article 6 has to comply with the second and third criteria of the test, while at the same time makes an express reference to the matrix of the three-step test in international copyright law, article 9 par. 2 of the Berne Convention. Furthermore, article 8 par 2 and 3 limit the rights of a lawful user of a database which is made available to the public in the same way as the three-step test limits the scope of exceptions to copyright (Hugenholtz, p. 332). According to article 8 par. 2, a lawful user may not perform acts which conflict with normal exploitation of the database or unreasonably prejudice the legitimate interests of the database maker. Article 8 par. 3 disposes that a lawful user may not perform acts which unreasonably prejudice the holder of a copyright or related right in respect of the works or subject matter contained in the database. In the Infosoc Directive, the possibility to control the mandatory temporary copying exception of article 5 par. 1 that is justified by the lawful use of a work or other subject-matter under the three-step test is expressly affirmed by article 5 par. 5 of the Directive.

The application of the three-step test in order to define the limits of exercise of copyright exceptions is strongly connected to the philosophy of balance of interests (Ficsor, p. 145, also Ficsor, p. 6, Senftleben, p. 35, Geiger, The role of the three-step-test in the adaptation of copyright law to the information society, p. 16, 18). The second and especially the third step of the test set up a mechanism of balancing the legitimate interests of copyright holders and the legitimate interests of the public, of copyright users.

It is crucial to understand, that both copyright protection and the exercise of copyright exceptions or of the “rights of lawful user” -where they are recognized as such- should be limited. The abstract conceptual criteria of “normalcy” of the exploitation of works of mind, of “unreasonable” prejudice and of the “legitimate” interests of the right holders have to be meticulously concretized in order

to evaluate the impact of the application of any specific exception in a particular context. This evaluation must take into consideration not only the justification of each particular exception or “right” but also the ultimate justification and the objectives of copyright protection in order to define the obligations of users. The research of the objectives of copyright protection should be considered as inherent in the evaluation of the “normalcy” of the exploitation and the “legitimacy” of the interests of the right holders. The appreciation of the destination of each exception or each right of the lawful user under the light of the destination of copyright law should be taken into account not only through the phase of the three-step, the examination of the prejudice of the exception or the right for the interests of the right holders, but already from the second step of the test, the consideration of the conflict with the normal exploitation of the work (Sinodinou, 71).

Under that specter, the exercise of the exceptions to copyright that are granted only to lawful users and of the “rights of the lawful users” has to respect the criteria of the three-step test. The lack of compatibility of the exercise of the exception or the “right” with the criteria of the test, render the priory lawful use unlawful. The importance of this compatibility is even higher if we consider that due to the implementation of the Infosoc Directive in national copyright legislations, the compliance of the exercise of the exception or the user’s “right” with the test can be evaluated directly by courts⁷. “The principles of good faith and of fair practice and the application of the three-step-test set on an effective mechanism of casting control of the lawfulness of acts of use under the light of the fundamental principle of proportionality which implies the balancing of the interests of the copyright holder and the user”.

Conclusion

The concepts of lawful user and of lawful use have emerged as new copyright norms in pieces of EU legislation which regulate the application of copyright law in the sector of new technologies. It is undeniable that the definition of these concepts is a very significant and delicate task. Indeed, the breadth of the spaces of freedom that are granted to end-users depends on this definition. However, the phenomenon of the appearance of the concept of lawful user in EU intellectual property law should not be overestimated. The concepts of “lawful user” and

7. For the perspectives and the risks of direct application of the three-step test by courts see: Gautier P.Y., *L'élargissement des exceptions aux droits exclusifs, contrebalancés par le « test des trois étapes »*, Communication-Commerce Électronique, novembre 2006, p.11. See also: Hugenholtz B., *The Implementation of Directive 2001/29/EC in the Netherlands*, RIDA, n°206, octobre 2005, p. 127. According to this author, the three-step test should only be applied by the European Court of Justice.

of “lawful use” and the instauration of the “rights of the “lawful user” still remain marginal in copyright law. They are used to determine reinforced copyright exceptions (in the sense that they are “ius cogens”) only in specific cases and they have not become fundamental concepts of copyright law. On the other hand, the emergence of these concepts in copyright law should not be neglected. The advent of lawful use in copyright law could be regarded as a precursor of a new perspective of the place of the user in copyright law that could eventually lead to a gradual transformation of all copyright exceptions to user’s rights.

REFERENCES

Bechtold, Directive 2001/29/EC, in Dreier/Hugenholtz, European Concise Copyright Law, Kluwer Law International (2006).

Benabou V.L. (1997), Droits d’auteur, droits voisins et droit communautaire, Bruylant Bruxelles.

Biegall S. (2001), Beyond Our Control? Confronting the Limits of Our Legal System in the Age of Cyberspace, MIT Press, London.

Boyhta G., Whose right is copyright?, GRUR Int. 1983, p. 384.

Cohen J., The place of the user in copyright law, 74 Fordham L. Rev. (2005), p. 347-74.

Derclaye E. (2008), The legal protection of databases, A comparative analysis, Edward Elgar Publishing.

De Saint Affrique D., Le droit sui generis sur les bases de données, Communication-Commerce Électronique, février 2004, 27.

Dusollier S., L’utilisation légitime de l’œuvre: un nouveau sésame pour le bénéfice des exceptions en droit d’auteur?, Communication-Commerce Electronique, novembre 2005, 17.

Dusollier S., Exceptions and Technological Measures in the European Copyright Directive of 2001-An Empty Promise, IIC 1/2003, 70.

Dusollier S., Technology as an Imperative for Regulating Copyright: From the Public Exploitation to the Private Use of the Work, EIPR. 2005, 201.

Dusollier S./Pouillet Y./Buydens M., Copyright and access to information in the digital environment, A study prepared for the Third UNESCO Congress on Ethical, Legal and Societal Challenges of Cyberspace, Infoethics 2000, Paris, 17 juillet 2000.

Ficsor M., EBU Copyright Symposium, Quo vadis, Copyright, Barcelona, March 31, 2006, How did we arrive here? The evolution of copyright legislation (The end of ~ ?), www.ebu.ch/

Ficsor M., How much of what? The three-step test and its application in two recent WTO dispute settlement cases, RIDA 2002, n°192, p. 111.

Gaster J., La protection juridique des bases de données à la lumière de la discussion concernant le droit d'auteur et les droits voisins dans la société de l'information, in: Libertés, droits et réseaux dans la société de l'information, Bruylant, Bruxelles, 1996, 38-39.

Geiger Chr. (2004), Droit d'auteur et droit du public à l'information, Approche de droit comparé, Litec, Paris.

Geiger Chr., The role of the three-step-test in the adaptation of copyright law to the information society, e-Copyright Bulletin, January-March 2007, <http://portal.unesco.org>.

Geiger, Chr., L'avenir des exceptions au droit d'auteur, Observations en vue d'une nécessaire adaptation et harmonisation du système, J.C.P., G, n° 47, étude, 23 novembre 2005, 2154.

Gaudrat Ph., Les démêlés intemporels d'un couple à succès: le créateur et l'investisseur, RIDA, octobre 2001, n°190, 71.

Ghestin J., L'utile et le juste dans les contrats, Dalloz 1982, chr., p. 1.

Guibault L. (2002), Copyright Limitations and Contracts, An Analysis of the Contractual Overridability of Limitations to Copyright, Kluwer Law International.

Heide Th., Copyright in the E.U. and United States: What "Access Right"?, EIPR 10/2001, p. 469.

Koelman K., Copyright Law and Economics in the EU Copyright Directive: Is the Droit d'auteur Passé ?, IIC 6/2004, p. 616.

Kotsiris L. (2008), Insolvency Law, Sakkoulas publications (in Greek).

Koumantos G., Les bases de données dans la directive communautaire, RIDA, janvier 1997. n°171, 101.

Lepage A., Vue générale sur les exceptions et les limitations au droit d'auteur dans l'environnement numérique, eBulletin du droit d'auteur, janvier-mars 2003.

Lucas /Devèze J./Frayssinet J. (2001), Droit de l'informatique et de l'Internet, PUF.

Lucas A./Lucas H.J. (2006), *Traité de la propriété littéraire et artistique*, 3e édition ; Lexis Nexis, Litec.

Masse (1994), *Rapport général*, in *Travaux de l'Association Henri Capitant*, La bonne foi, Paris, Litec.

Senftleben M. (2004), *Copyright, Limitations and the Three-step test*, An analysis of the Three-step test in International and EC Copyright Law, Kluwer Law International.

Sinodinou T., *Voyages des sources du test des trois étapes aux sources du droit d'auteur*, *Revue Lamy Droit de l'immatériel*, août/septembre 2007, p. 67.

Spinello R. & Bottis M., *A Defence of Intellectual Property Rights*, 2009, Edward Elgar Publishing.

Storme M., *Good faith and the contents of contracts in European private law*, *E.C.J. L. vol. 7.1*, march 2003, <http://www.ejcl.org>.

Strowel A. (1993), *Droit d'auteur et copyright, Divergences et convergences*, *Etude de droit comparé*, Bruylant, Bruxelles, L.G.D.J. Paris.

Strowel A./Derclaye E., *Droit d'auteur et numérique : logiciels, bases de données, multimédia*, *Droit belge, européen et comparé*, Bruylant, Bruxelles 2001.

Vanovermeire V., *The Concept of the Lawful User in the Database Directive*, *IIC 1/2000*. 65.

Report from the Commission to the Council, the European Parliament and the Economic and Social Committee on the implementation and effects of Directive 91/250/EEC on the legal protection of computer programs, COM/2000/0199 final.

Information, Knowledge and Wisdom: Groundwork for the Evaluation of Digital Information and Its Relation to the Good Life

Edward H. Spence*
Department of Philosophy, University
of Twente, Netherlands

Where is the Life we have lost in living?
Where is the wisdom we have lost in knowledge?
T. S. Eliot's Choruses from *The Rock* (1934)

Where is the knowledge we have lost in information?

Abstract

The paper provides a general meta-philosophical groundwork for the theoretical and applied normative evaluation of digital information in relation to the good life. The overall general aim of the paper is to address the question how computer ethics can be expanded to more centrally include the issue of the quality of life or the good life, for individuals and for society. In offering this groundwork the paper, due to constraints of space and time, will not provide a detailed examination and evaluation of specific normative issues, which arise in the production, dissemination and use of digital information. It will, however, provide a methodological approach of how different types of some major practical manifestations of digital information (henceforth information) can be evaluated using the meta-theoretical framework proposed in this paper.

The paper comprises three inter-related parts. Part (1) provides a summary of an argument I presented at CEPE 2007 (San Diego) whose primary aim is to demonstrate a meta-philosophical model, the Dual Obligation Information Theory (DOIT) to be used in the analysis and evaluation of digital information in terms of a cluster of normative categories. Those categories are the epistemological, the

* *Edward Spence*, BA (Hons.1st), PhD, is a senior lecturer in moral philosophy at Charles Sturt University, Australia. He is a Senior Research Fellow at the Centre for Applied Philosophy and Public Ethics (CAPPE), Australia and the 3TU. Ethics of Technology Centre, Netherlands. He is author of: *Advertising Ethics* 2005, and *Corruption and Anti-Corruption: A Philosophical Approach*, 2005, published by Prentice Hall, USA; and author of *Ethics Within Reason*, 2006, Lexington Books, USA

ethical, the axiological and the eudemonic. The overall objective of Part (2) is to examine and evaluate the theoretical and practical relationship between information, knowledge and wisdom. Crucially, this is a new and the innovative focal direction that this paper will take, using the meta-philosophical analysis provided by DOIT. Wisdom understood as a form of meta-information or meta-knowledge provides a direct conceptual and practical link between the concepts of information and the good life and more generally a direct link between computer ethics and the good life. As such, the concept of wisdom, as a type of meta-knowledge, allows for a direct evaluation of the axiological and eudemonic aspects of information in relation to a conception of a good life. Following on from parts (1) and (2), part (3) will provide a brief theoretical outline by way of offering a methodological approach of how different types of some major practical manifestations of digital information (henceforth information) can be normatively evaluated in relation to a conception of a good life through the application of the concept of wisdom as developed in section (2).

A Meta-Philosophical Evaluation of Information: A Universal Model for Evaluating the Normative Quality of Digital Information

The objective of this part of the paper is to describe and demonstrate a meta-theoretical framework for the normative and practical evaluation of digital information.

The meta-theoretical framework (DOIT) consists of two main inter-related models that together demonstrate the universal normative character of information and its global normative applicability: (A) The Inherently Normative Account of Information model (INAI) and (B) the Model of the Unity of the Right, the Good and the Good Life (MURG). Together these two models are designed to demonstrate and explain the dual-normative structure of information by disclosing the underlying epistemological, ethical, axiological and eudemonic commitments to which it gives rise and by virtue of which all informational agents are universally bound. INAI does so in terms of disclosing the epistemological and ethical principles and values inherent in information as a process of communication (Spence 2009 in press; and 2007a); and MURG does so in terms of universal rights (freedom and wellbeing) (Spence 2006; Gewirth 1978; 1996; and Beyleveld 1991) to which all informational agents are entitled by virtue of the inherent normative structure of action generally and information action specifically. In addition, on the basis of universal rights MURG demonstrates that these rights, in turn, give rise to prudential commitments to virtues, values and moral sentiments (the good) and to happiness, self-fulfillment (Gewirth 1998) or the preferred term used in this approach, eudemonia as the primary conditions for a good life (Spence

ce, 2006, chapters, 5 and 10; Spence, 2007a and 2007b). In sum, together INAI and MURG are designed to be applied to objectively and universally evaluate the normative quality of digital information in terms of its ethical (universal rights), epistemological (knowledge and truth) axiological (values) and eudemonic (virtues, self-fulfillment and happiness) aspects.

Evaluating information using the proposed meta-theoretical framework provided by DOIT as instantiated jointly by INAI and MURG is essential for addressing all the key normative features of information (epistemic, ethical, axiological, and eudemonic) as they relate to and impact on all aspects of the lives of individuals and those of societies generally. The evaluation of information in terms of only one or two of those features will leave out something important that is of interest and concern to its disseminators.

The Inherent Normative Structure of Information and Knowledge

Before embarking on the exploration of the conceptual and practical connection between information and knowledge on the one hand and wisdom on the other in Part (2), I will first provide an argument for the normative structure of information with regard to its epistemological, ethical, axiological and eudemonic dimensions. In describing the Dual Obligation Information Theory (DOIT) used for the evaluation of information that comprises the INAI and MURG meta-theoretical models as outlined above, the paper will employ an epistemological account of semantic information based on a minimal nuclear definition of information (Dretske 1999, 45). Following Luciano Floridi it will define information as “well formed meaningful data that is truthful” and following Dretske it will define information as “an objective commodity capable of yielding knowledge” and knowledge, in turn, will be defined as “information caused belief”.

According to Dretske’s nuclear definition of information,

A state of affairs contains information about X to just that extent to which a suitably placed observer could learn something about X by consulting it. This, I suggest, is the very same sense in which we speak of books, newspapers, and authorities as containing, or having, information about a particular topic, and I shall refer to it as the nuclear sense of the term “information”... Information is what is capable of yielding knowledge, and since knowledge requires truth, information requires it also (1999, 45).

According to Dretske’s notion of knowledge,

“K knows that s is F=K’s belief that s is F is caused (or causally sustained) by the information that S is F...s [is to be understood to be something] K perceives, something at an informational source about which K receives information. If K has a belief about this object, the belief that it is F, then this

belief qualifies as knowledge if and only if that belief is caused (or causally sustained) by the information that it is F (1999, 86)....the knowledge that s is F requires (because it is required as a cause of belief) the information that s is F (1999, 105).

What is necessary for both information and knowledge, therefore, is truth. For information without truth is not strictly speaking information but either misinformation (the unintentional dissemination of well-formed and meaningful false data) or disinformation (the intentional dissemination of false "information").

Using the minimal account of information described above, we can now develop an inherent normative account of information (INAI), which demonstrates and describes the generic epistemological and ethical commitments that necessarily arise in the dissemination of semantic information. A central claim of the paper is that all informational processes comprising the dissemination of information, specifically as a process of communication, commit all rational agents to both epistemological and ethical conduct; specifically, insofar as information of necessity has to be true, it commits all agents to epistemological values such as accuracy, truth, reliability, verifiability, objectivity and justification, among others, and corresponding ethical values and virtues such as sincerity, honesty, truthfulness, trustworthiness, justice or fairness.

Briefly, the argument is as follows: Insofar as information is a type of knowledge (it must be capable of yielding knowledge, one must be able to learn from it) it must comply with the epistemological conditions of knowledge, specifically, that of truth. And insofar as the dissemination of information is based on the justified and rightful expectation among its disseminators and especially its users that such information should meet the minimal condition of truth, then the disseminators of information are committed to certain widely recognized and accepted epistemological criteria. Those epistemic criteria will in the main comprise objectivity as well as the independence, reliability, accuracy and trustworthiness of the sources that generate the information. The epistemology of information in turn commits its disseminators to certain ethical principles and values, such as honesty, sincerity, truthfulness, trustworthiness and reliability (also epistemological values), and fairness, including justice, which requires the equal distribution of the informational goods to all citizens. Thus in terms of its dissemination, information has an intrinsic normative structure that commits everyone involved in its creation, production, search, communication, consumption and multiple other uses to epistemological and ethical norms and these norms being intrinsic to the normative structure of information with regard to all its disseminating modes are rationally unavoidable and thus not merely optional.

However, the following objection could be raised: not all forms of information might be fully accounted and explained by the epistemological account of information proposed. Information as a form of personal expression – expressive information, for example, personal information one might post on the Internet through Blogs or create for one's avatar in a Virtual World, or post on YouTube, Face Book or My Space, might be such types of e-information – “I am John and my avatar in Second Life is a female called Jane”. But again, this statement could be true or false (for example, as a matter of fact, “I don't have an avatar” or “my avatar's name is not Jane but Petunia”) and will therefore have to meet the minimal normative conditions of the nuclear account of information (both epistemological and ethical conditions).

Information and Universal Rights

The goal of the following argument is to show that apart from committing its disseminators to unavoidable epistemological and ethical standards by virtue of its own inherent normative structure, information commits its disseminators to respect for peoples' rights. That is, information, must not be disseminated in ways that violate peoples' fundamental rights to freedom and wellbeing (generic rights), individually or collectively, or undermine their capacity for self-fulfilment (Negative Rights). In addition, information must as far as possible be disseminated in ways that secure and promote peoples' generic rights and capacity for self-fulfilment (Positive Rights) when those rights cannot be secured or promoted by the individuals themselves and can be so secured and promoted at no comparable cost to its disseminators. But from where does this authority come and what are the fundamental rights to which I refer? Alan Gewirth's Principle of Generic Consistency (PGC) offers a description and prescription for both the rational authority (based primarily on instrumental and deductive rationality) and the content of the fundamental rights (freedom=FR and wellbeing=WB) that persons have necessarily and only by virtue (sufficient reason) of being purposive agents.

Due to constraints of space, I will not attempt to provide a justification for Alan Gewirth's argument for the Principle of Generic Consistency (PGC) on which his derivation of rights is based, as this is well beyond the scope and limits of this paper. I will, however, offer a brief summary of the rationale of the argument for the PGC by way of a schematic outline of the three major steps of that argument.

The Rights of Agents: The Rationale for Alan Gewirth's Argument for the Principle of Generic Consistency

Gewirth's main thesis is that every rational agent, in virtue of engaging in action, is logically committed to accept a supreme moral principle, the Principle of Generic Consistency. The basis of his thesis is found in his doctrine that action has

a normative structure, and because of this structure every rational agent, just in virtue of being an agent, is committed to certain necessary prudential and moral constraints.

Gewirth undertakes to prove his claim that every agent, qua agent, is committed to certain prudential and moral constraints in virtue of the normative structure of action in three main stages. First, he undertakes to show that by virtue of engaging in voluntary and purposive action, every agent makes certain implicitly evaluative judgments about the goodness of his purposes, and hence about the necessary goodness of his freedom and wellbeing, which are the necessary conditions for the fulfillment of his purposes. Secondly, he undertakes to show that by virtue of the necessary goodness which an agent attaches to his freedom and wellbeing, the agent implicitly claims that he has rights to these. At this stage of the argument, these rights being merely self-regarding are only prudential rights.

Thirdly, Gewirth undertakes to show that every agent must claim these rights in virtue of the sufficient reason that he is a prospective purposive agent (PPA) who has purposes he wants to fulfill. Furthermore, every agent must accept that, since he has rights to his freedom and wellbeing for the sufficient reason that he is a PPA, he is logically committed, on pain of self-contradiction, to also accept the rational generalization that all PPAs have rights to freedom and wellbeing. At this third stage of the argument these rights being not only self-regarding but also other-regarding, are now moral rights. The conclusion of Gewirth's argument for the PGC is in fact a generalized statement for the PGC, namely, that all PPAs have universal rights to their freedom and wellbeing.

Applying the PGC to information, we can now make the further argument that information generally and digital information specifically, must not be disseminated in ways that violate informational agents' rights to F and WB, individually or collectively, (Negative Rights). Moreover, information must as far as possible be disseminated in ways that secure and promote the informational agents' rights to F and WB (Positive Rights). Conceived as the Fourth Estate, this places a significant and important responsibility on the disseminators of information and in particular the media, especially journalists, both offline and online.

Information generally can be epistemologically and ethically evaluated internally by reference to its inherent normative structure. That structure commits its disseminators, to ethical and epistemological norms. This is especially true of professional communicators (Journalists and PR Consultants, for example, on-line and off-line).

Insofar as the ethical values to which the inherent normative structure of information gives rise require that the informational agents' rights to F and WB should

be respected, secured and promoted, those values are also mandated by the PGC and thus information can also be externally evaluated by reference to the PGC. Expressive Information can also be evaluated either internally or externally or both, in this way. For example, identity theft on the Internet is morally wrong both because it is untruthful (internal evaluation) and because it can cause harm (external evaluation).

Information and a Good Life

Although the PGC is a meta-ethical model of morality that describes what is right, there is still a matter of describing the other two components of this model, namely, the good and the good life.

Let us first begin with the good (specifically, I will only be concerned here with the notion of a good person): insofar as the PGC requires all agents, including informational agents, to act ethically or at least acknowledge that they ought to act ethically in respecting the rights to freedom and wellbeing of other agents including their own, and insofar as virtues of character such as the cardinal virtues of justice, courage, moderation and prudence, as well as the Humean moral sentiments, such as sympathy (positive) and remorse (negative), can be conceived as enabling dispositions that allow agents generally and informational agents specifically to act ethically in compliance with the PGC, then the inculcation of those virtues and cultivation of those sentiments are also rationally required, at least prudentially (Spence 2006).

Secondly, a good life is one that is at least minimally capable of enabling a person to attain self-fulfilment or eudemonia. For insofar as self-fulfilment, happiness or eudemonia is the ultimate object in life as Aristotle claimed, it is difficult to conceive a life that was not at least capable of leading to the attainment of self-fulfilment, as good – what would it be good for if it were incapable of realizing one's ultimate objective in life? A good life in turn is capable of attaining self-fulfillment or eudemonia if it at least accords with the minimal requirements of morality in accordance with the PGC. Those requirements can more successfully be complied with through the inculcation of the virtues and the moral sentiments, in accordance with an indirect application of the PGC. That is, a good life capable of resulting or at least contributing to self-fulfillment is more likely to be realizable if one is good – that is, if one has a good character comprising both the moral virtues and moral sentiments. And by being good one is also more likely to comply with the requirements of the PGC by respecting the rights of other people including his/her own.

The following is a summarised outline of my construction of an argument for self-fulfilment, based on Gewirth's theory of self-fulfilment (Gewirth, 1998):

- Gewirth distinguishes between self-fulfillment as aspiration – fulfillment and self-fulfillment as capacity – fulfillment. The former is the satisfaction of one's deepest desires, the latter, the process and goal of making the best of oneself (Gewirth 1998).
- According to Gewirth, "Moralitygives counsel and precepts for the self's having a good life through personal development of one's capacities whereby one makes the best of oneself" (1998: 107).
- A good life is one that is at least minimally capable of enabling a person to attain self-fulfillment (both aspiration-fulfillment and especially, capacity-fulfillment) and eudemonia.
- A good life in turn is capable of attaining self-fulfillment or eudemonia if it at least accords with the minimal requirements of morality in accordance with the PGC.
- Those requirements can more successfully be motivationally complied with through the inculcation of the virtues and the moral sentiments, in accordance with an indirect application of the PGC.

In conclusion of this section, the application of the Unified Model of Rights, the Good and Good Life (MURG) to information provides a convergence of Rights (Freedom and Wellbeing) plus the Good, specifically the Good Person (Virtues of character and moral sentiments) plus a Good Life (Self-fulfilment, or Happiness, which collectively I refer to as Eudemonia). Thus, the notion of a good life used in the paper is a eudemonic conception of a good life. A life, that is, that is at least capable of leading to the attainment of eudemonia. Accordingly, the application of MURG to information requires that information must not be disseminated in ways that violate persons' rights to F and WB, individually or collectively, or undermine their capacity for Self-fulfilment or Eudemonia (Negative Rights). Information must as far as possible be disseminated to secure and promote persons' rights to F and WB and their capacity for Self-fulfilment or Eudemonia (Positive Rights).

Information, Knowledge and Wisdom

Introduction

This is the focal and main part of the paper whose aim is the examination and evaluation of the relationship between the concepts of information (well formed meaningful data that is true or truthful – Floridi 2005), knowledge (roughly, information that we believe to be true and have good reasons based on justified and demonstrable evidence or testimony to believe it to be true) and wisdom

(a type of meta-knowledge that is used in the application of information and knowledge to make right judgments in reaching appropriate decisions that are of value and good for us personally (prudentially and eudemonically good) and that are of value and good for others (ethically good). The overall objective of this line of enquiry is to determine to what degree, if any, information contributes to a good life.

In this part of the paper, MURG will be applied to the examination for determining more closely the relationship between information and the good life via the concept of wisdom: if wisdom is a primary and essential condition for an individual in (a) determining what a good life is or ought to be (meta-knowledge- that and meta-knowledge- why) and (b) a primary and essential condition in providing us with guidance and direction, as individuals and societies generally, of how to live such good lives and (c) moreover, practically enabling us to live such good lives for the attainment of eudemonia (meta-knowledge-how), to what extent and in what ways then, if any, does information contribute to wisdom and by extension to the good life?

An initial hypothesis of this paper is that one way of evaluating the value of information (its axiological goodness) is by determining the degree to which it contributes or is capable of contributing to the attainment of a good life epistemologically (its capacity to yield knowledge) prudentially (its capacity to contribute to one's ability in making sound judgments concerning particular practical matters – I shall refer to this type of prudence as *phronesis*, a type of Aristotelian practical wisdom) ethically (its ability to contribute to the moral good of others both negatively by causing no unjustified harm to others, and positively by causing positive good for others) and eudemonically (its capacity to contribute to both the conception and the attainment of a good life). Having demonstrated in (Part I) the inherent conceptual connection between information and the epistemological and ethical commitments to which it gives rise as a product and process of communication and by virtue of which it universally bounds all informational agents, the primary aim of this focal part of the paper is to articulate and hopefully demonstrate the essential conceptual connection between information and a good life through wisdom acting as the necessary lynchpin between the two.

Analyzing information through the application of the concept of meta-knowledge (knowledge-that, knowledge-how and knowledge-why) of what is good or evil for us and others - how it contributes or is capable of contributing to a good life for us and others for the attainment of eudemonia - is what the paper will initially postulate as Wisdom. To avoid confusion, I will use the term wisdom to define a general and holistic notion of wisdom, a notion similar to the Greek notion of *Sophia*. The concept of wisdom understood this way, can then be used to evalu-

ate the qualitative goodness of information: its axiological value for both human beings and other species. For wisdom, at least in principle, should at the very least be capable of enabling us to live and act wisely for the attainment of eudemonia not only with regard to ourselves and other human-beings but also with regard to all informational entities within the infosphere, roughly understood in this paper in Floridian terms .

In sum, insofar as a good life should at least be capable of leading to self-fulfilment or eudemonia (otherwise what is it good for?), especially self-fulfilment as capacity-fulfilment (making the best of oneself as a human-being – Gewirth 1998) then wisdom (understood as a type of meta-knowledge, the acquisition of which enables one to create, communicate and use information so as to render oneself and others, whenever possible, capable of achieving self-fulfilment and eudemonia) is a necessary condition for a good life.

An important qualification to the claim made in this paper that wisdom is a necessary condition for a good life is that such a life is conceived eudemonically. For the notion of wisdom developed in this paper and applied in evaluating the axiological goodness of information is itself a eudemonic conception of wisdom. However, such a eudemonic notion of wisdom is not unlike our commonsense and pre-theoretical understanding of wisdom, namely, an overarching reflective capacity the possession of which allows one to lead a good life and moreover enables one to guide others in leading fulfilling and good lives. This eudemonic notion of wisdom is akin to the notions of wisdom defended by philosophers such as Plato, Aristotle, the Epicureans and the Stoics, and later Kant and Gewirth who although postulated and defended somewhat different notions of the good life, can nevertheless collectively be thought of as offering eudemonic accounts of the good life. It can be said that a common denominator for all eudemonic accounts of wisdom is their subordination of the concepts of pleasure and desire to that of virtue. The essential link between pleasure and desire on the one hand and virtue on the other might be weaker in the case of the Epicureans and stronger in the case of the Stoics but whatever the strength of that relationship, the link between pleasure, desire and virtue is an essential characteristic of a eudemonic conception of a good life and also a eudemonic conception of wisdom.

Importantly, the relationship between wisdom and a good life proposed in this paper under a eudemonic conception of a good life is reflexive. For wisdom guides one to the choice of a eudemonic conception of a good life and the pursuit of such a life, and a eudemonic conception of a good life, in turn, guides and motivates one to the acquisition of wisdom as an enabling disposition, in the form of an overarching reflective virtue, which is necessary for the attainment of a eudemonic life. This should not surprise us. For although wisdom acts initially instrumentally, as a necessary enabling virtuous disposition for the attainment of

a eudemonic life, once attained a eudemonic life becomes inseparable from the state of wisdom that enabled its attainment. This reflexivity between eudemonia and wisdom allows us then to say that a wise person is generally a eudemonic person and a eudemonic person is generally a wise person. However, I don't wish to exclude the logical possibility that one could be wise but unhappy although pragmatically, given our common understanding of wisdom, that would be an odd thing to say and in practice I think, an unusual occurrence.

What is Wisdom?

Having examined in some detail what information and knowledge are and what the relationship is that exists between them (by way of an examination of the essential property that characterises both, namely, the property of truth) it is now time to turn our attention to the notion of wisdom so as to explore further the conceptual relationship that holds between wisdom and information.

According to Nicholas Maxwell,

"The central task of inquiry is to devote reason to the enhancement of wisdom – wisdom being understood here as the desire, the active endeavour, and the capacity to discover and achieve what is desirable and of value in life, both for oneself and for others. Wisdom includes knowledge and understanding but goes beyond them in also including: the desire and active striving for what is of value, the ability to see what is of value, actually and potentially, in the circumstances of life, the ability to experience value, the capacity to help solve those problems of living that arise in connection with attempts to realize what is of value, the capacity to use and develop knowledge, technology and understanding as needed for the realization of value. Wisdom, like knowledge, can be conceived of, not only in personal terms, but also in institutional or social terms. We can thus interpret the philosophy of wisdom as asserting: the basic task of rational inquiry is to help us develop wiser ways of living, wiser institutions, customs and social relations, a wiser world" (2007, 79).

What is of interest to in Maxwell's quoted passage for our present purposes is the relationship he draws between the concepts of reason, knowledge, understanding, and the desire, capacity, and active endeavour for the achievement (or as in my case attainment) of what is of value in life, for oneself and others. With the exception of understanding, for which I will have more to say in what follows, the other concepts to which Maxwell draws attention seems to anticipate and reflect both explicitly and implicitly, the concepts included in my own normative analysis of information and knowledge in terms of their epistemological, ethical, axiological and eudemonic dimensions, on the basis of the meta-theoretical

framework of DOIT that comprises the two meta-theoretical evaluative models of INAI and MURG.

Knowledge (for) Wisdom and Knowledge (as) Wisdom

An important distinction when enquiring into the conceptual connection between knowledge (K) and wisdom (W) is the distinction between knowledge for wisdom and knowledge as wisdom. Although related the two are quite different and their difference highlights an important and crucial distinction between (K) and (W). No doubt some general knowledge about the world acquired on the basis of reliable and veridical information that causes it and sustains it (Dretske 1999) is necessary for wisdom. For with the exception of Socratic ignorance, roughly understood here as having knowledge of one's ignorance (being aware of one's ignorance and humbly acknowledging one's lack of knowledge), ignorance of one's ignorance is not conducive to wisdom. For unlike Socratic ignorance that prompts and motivates one to acquire the knowledge of which one is ignorant (knowledge understood here as some minimal general knowledge about some basic aspects of the world, e.g., history, geography, science, mathematics, literature, art, e.tc..) they who are ignorant of their ignorance and falsely claim to know that for which they lack knowledge, are not in a position to be motivated to acquire the knowledge they lack; and moreover, the knowledge which is at least in a minimal and general sense partly necessary for the acquisition of wisdom and by extension, the attainment of a good life and eudemonia.

Thus at a minimum, and bracketing the possibility that "holy fools" though totally ignorant of facts about the world are nevertheless in some sense "wise", some minimal and general information and knowledge about the world is instrumentally necessary for the acquisition of wisdom. At least at a minimum, an attitude of Socratic ignorance might be necessary for the acquisition of wisdom. For the Socratic elenchus can be applied as a method for acquiring the knowledge one lacks, through first recognizing and acknowledging one's ignorance, and then being motivated to gradually acquire that knowledge of which one is ignorant, through critical enquiry and further investigation. According to John Kekes, "the elenchus enables its practitioners to progress from a special kind of ignorance---foolishness---to a special kind of knowledge---moral wisdom" (1995, 39).

By contrast, unrecognized and unacknowledged ignorance of our ignorance, individually and collectively, is no bliss for it precludes us for ever developing the capacity for the acquisition of wisdom. We can therefore say that the acquisition of such general minimal knowledge about the world or an attitude of Socratic ignorance when we lack it is instrumental to the acquisition of wisdom because it provides at least part of the necessary means, that is, the capacity for the acquisition of wisdom. Moreover, the acquisition of such minimal and general

knowledge of the world or in its absence, an adoption of an attitude of Socratic ignorance, is prudential to the acquisition of wisdom. Insofar as we consider the acquisition of wisdom desirable, valuable and essential for the attainment of a good life, we should (normatively) inculcate in ourselves the virtue of learning: the desire and active pursuit of the acquisition of at least a minimal and general knowledge about the world. Hence, some minimal and general knowledge about the world is necessary for wisdom.

What about knowledge as wisdom? What kind of knowledge could that be? At first approximation we could say along with Socrates, that wisdom is knowledge of good and evil. That wisdom is a type of moral knowledge. It is this type of knowledge that prompted Socrates in the *Apology* to exhort his fellow-Athenians to seek “wisdom, truth, and the best state of the soul” as a way of living a good life (Benson, 2000, 23). According to Benson, for Socrates ‘it is not merely he who lives knowledgeable [see my knowledge for wisdom above] who is happy [eudemon] but who lives “knowledgeably with respect to some particular thing” [Charmides, 173e6-10]. Benson goes on to say that for Socrates the particular thing that one needs to have knowledge of in order to be happy or eudemon is “knowledge of the good and the bad” (Benson 2000, 153). In the discussion that follows, I shall, following Benson, refer to that type of knowledge as moral knowledge. Interestingly, John Kekes (1995) refers to the acquisition of wisdom that follows from such moral knowledge, as moral wisdom.

Moral Wisdom

In the opening sentence of his book *Moral Wisdom and Good Lives* (1995, ix) Kekes tells us that “moral wisdom is a virtue – the virtue of reflection”. A more detailed characterisation of moral wisdom by Kekes, is that

“Moral wisdom is the capacity [a psychological capacity] to judge rightly what should be done in particular situations to make life better...Because this human psychological capacity, once developed, is likely to be lasting and important, it can be identified as a character trait....We can say, therefore, that people have moral wisdom if they regularly and predictably act wisely in the appropriate situations and if so acting is an enduring pattern in their lives... Whether an action is morally wise depends also on what the agents bring to the judgements they make, such as their particular conception of what would make life better. An action being morally wise depends therefore not just on the nature of the action and the situation, but also on the agent, and this invalidates generalisations of moral wisdom which ignore the character and beliefs of the agents.” (1995, 5-7).

According to Kekes, moral wisdom is a second order virtue whose primary concern, “[Is] the development of our character [emphasis added] in a desirable direction by strengthening or weakening some of our dispositions. First-order virtues guide our actions in view of what we think of a good life; second-order virtues guide our actions with a view of developing the kind of character that reflects a reasonable conception of a good life(1995, 9).

For Kekes moral wisdom is reflexive.

“The subject who has the virtue is the same as the object toward whom the appropriate actions are directed” (1995, 9). He goes on to say that being a second-order and a reflexive virtue moral wisdom “involves the direction of first-order virtues” (1995, 10).

Kekes goes on to explain that

“The primary motivation behind the second-order and reflexive activities prompted by moral wisdom is the desire to make our lives better [emphasis added]. We can do so, however, only if we have some conception of what a good life would be. The aim that governs our exercise of moral wisdom must therefore be the realization of our conception of a good life. The reason for directing our first-order virtues is to improve our lives by transforming our characters so as to improve the chances of achieving what we regard as a good life. This valued conception of a good life, however, may not be reasonable. Since moral wisdom is concerned not merely with means but also with ends [emphasis added], it has among its tasks the critical scrutiny of the conception of a good life which motivates us to act according to moral wisdom” (1995, 10-11).

Importantly for Kekes,

“Moral wisdom aims further than the transformation of our character; it aims as well at the development of a conception of a life which not only seems to us to be good but which is good [emphasis added]. And the character that we desire to have is not just an instrumental good whose possession would be conducive to the good achievement of this genuinely good life; it is also an intrinsic good, because having the sort of character that our conception of a genuinely good life requires is an essential part of the conception. For a conception of a good life guides us to live in a certain way; living involves acting; and we normally act in characteristic ways, that is, according to our character” (1995, 11).

Kekes concludes that,

Moral wisdom is a human psychological capacity to judge soundly what we should do in matters seriously affecting the goodness of our life. The judgement is made in the light of our conception of a good life, but it concerns the

evaluation of both the actions that exemplify the conception and the conception itself. Moral wisdom is thus sound judgement involving the application of knowledge of good and evil to the evaluation of both the means to and the ends constitutive of good lives" (1995, 14).

Wisdom, Judgment and Control

Other important components of moral wisdom are judgement and control. According to Kekes,

"The eudemonistic conception of a good life is not to be understood as the endorsement of a particular form of life. It is rather a regulative ideal that specifies some general conditions to which all good lives must conform" [emphasis added] (1995, 24).

This claim by Kekes is insightful and very much in keeping with the eudemonic conception of a good life proposed and argued for on the basis of MURG above. For MURG is intended only as a meta-theoretical regulative ideal that specifies some general conditions to which all good lives must conform regardless of the particular contexts and contingencies of those lives. Central to those general conditions to which all good lives are bound are the universal rights to freedom and wellbeing to which all agents are entitled; as well, the virtues of character and associated moral sentiments and values that are prudentially desirable and required as enabling general motivational dispositions for the attainment of a good life in line with the normative directions of MURG.

Kekes correctly claims that according to a eudemonistic conception of a good life, "Primary values [values that concern uniform and universal human goods and needs] may be thought of as establishing the moral limits and secondary values [values that vary across individuals in accordance with differences in cultural traditions, conceptions of a good life, and individual contingencies and circumstances] as establishing the moral possibilities that define good lives (1995, 25)....the former define a grid [emphasis added] within which human beings must endeavour to make a good life for ourselves, while the latter provide the ways in which individuals fill in the grid" (1995, 23)

According to Kekes,

"The role of moral wisdom is to acquire general knowledge of good lives and to go beyond it by bringing the general to bear on our particular character and circumstances. The component of moral wisdom that makes this possible is judgement.... The need for judgement, then, would indicate a kind of ignorance; judgement would be an expedient needed to take up the slack created by insufficient knowledge....Judgement is a process, therefore, by which a decision is reached about what to do or not to do, given the good the agent wants

to achieve and the agent's concrete situation" (1995, 25-26).... The way to moral wisdom is to endeavour to fall less short [of good judgement] by making our judgements better, and that involves increasing our control. The process requires enlarging the area of our lives that we can order so as to conform to our conception of a good life. It is a process of shaping ourselves to become the person our conception requires [emphasis added] (1995, 74).

I emphasised the phrase in Kekes' passage above in order to draw attention to its close similarity to Alan Gewirth's notion of capacity-fulfilment---becoming the best human-being possible---discussed earlier.

Kekes concludes,

"That growth in moral wisdom depends on increasing control. The reason for this is that moral wisdom consists in living in accordance with our conception of a good life, and that, in turn, consists in using our knowledge, evaluation, and judgement to transform the complex moral situations we encounter into simple ones. The capacities we have, the situations we face, and the judgements we make are, however, subject to the influences of permanent adversities. The appropriate exercise of our capacities, facing the situation in the right way, and making reasonable judgements requires coping with contingency, conflict, and evil, and that is possible, if at all, only by increasing our control" (1995, 94).

Wisdom and Self-knowledge

According to Kekes,

"The object of self-knowledge is the knower's character. As a first approximation, it may be said that character is composed of enduring patterns of motivation and action (1995, 115)...self-knowledge....is a mode of reflection, involving judgement, whose aim is to make our character less fortuitous and more deliberate. This is the same process as that of increasing control, for what moves our character in the desired direction is that we control it to approximate more closely than before our conception of a good life. This desirable transformation proceeds through the evaluation of our desires, capacities, opportunities, values, and actions with a view of forming out of them such enduring patterns as we regard conducive to a good life (1995, 127-28)...The process of acquiring self-knowledge is thus a mode of reflection directed toward the transformation of our character" [emphasis added] (1995, 136).....This transformation involves increasing our control over what desires, capacities, opportunities, values, and actions form the enduring patterns of our character. It is this way that judgement and control are connected. Control is increased by strengthening the motivational force of our conception of a good life and by weakening the internal obstacles in our character to living according to that conception. Self-knowledge

is essential to this process because it is through it that we learn what the internal obstacles are and what we can do to cope with them...that account provides our moral identity and motivates us to transform our character so as to make it less fortuitous and more deliberate. What directs this transformation is our conception of a good life" [emphasis added] (1995, 159).

I emphasised the last sentence in Kekes' passage above in order to re-direct attention the important reflexivity between moral wisdom and a conception of a good life to which I first drew attention earlier on. That is, a conception of a good life directs us to acquire wisdom (or moral wisdom in Kekes' case) so as to enable us in attaining it, and wisdom, in turn, informs the choice of our conception of a good life. This realisation is important, because it reveals a close fit between the notion of wisdom and the eudemonic conception of a good life. That is to say, if one is morally wise one would chose a eudemonic conception of a good life as that which has the closest affinity with wisdom; at least the notion of wisdom advocated in this paper. In turn, a eudemonic conception of a good life would direct one to acquire wisdom so as to attain and sustain such a conception of a good life. This might at first seem circular but it is not so. For it is the desirability for attaining a eudemonic life that motivates the seeker of such a life to acquire the necessary prudential and practical means for achieving that goal and the best candidate for providing those necessary means, is wisdom. Thus if one desires and values a eudemonic life, a life capable of leading to eudemonia, one would also desire and value wisdom as the necessary means for attaining it; and if one is wise they would desire and value a eudemonic life.

Wisdom and Understanding

According to Jonathan Kvanvig,

"Understanding requires the grasping of explanatory and coherence-making relationships in a large and comprehensive body of information. One can know many unrelated pieces of information, but understanding is achieved only when informational items are pieced together by the subject in question... the object of understanding is an 'informational chunk' rather than a number of single propositions...what is distinctive about understanding has to do with the way in which an individual combines pieces of information into a unified body...whereas knowledge can be piecemeal, understanding requires more completeness...such organisation is pragmatically useful because it allows us to reason from one bit of information to other related information that is useful as a basis for action, where unorganised thinking provides no such basis for inference...(Kvanvig 2003, 192, 197, 202-203).

Citing Sosa, Kvanvig relates the notion of understanding to a type of reflective knowledge, "Sosa describes reflective knowledge in terms of manifesting the understanding of a fact's place in a wider whole" (2003, 206). Comparing under-

standing to knowledge he asks if “understanding [is] thereby more valuable than knowledge” and concludes that although he has “not argued that this is so, there is a good case to be made for it” (2003, 206).

If understanding is a feature of wisdom, and I will argue that it is, and specifically as a body of information that concerns how one ought to live a good life for the attainment of eudemonia, then there is a case to be made that understanding is more valuable than the mere piecemeal accumulation of disparate and unrelated facts on many different topics. Where the latter might not always be necessary for wisdom (consider someone who is very good at trivia competitions but quite bad at knowing how to live a good life) the former is.

Interestingly, Hugh Benson interprets Socratic knowledge as a type of understanding. According to Benson,

“Socratic knowledge (wisdom or expertise) is a strong and complete grasp of distinct F-nesses...This grasp of the respective F-ness produces correct judgements involving F-ness that yield true cognitive states consistent with the knower’s other cognitive states involving F-ness as well as the ability to answer the Socratic ‘What is F-ness?’ question in a way consistent with those other cognitive states. Moreover, being the result of such a grasp..is both a necessary and sufficient condition for the resulting cognitive states to be knowledge states. The correct judgements are knowledge-yielding judgements if and only if they are produced by such a grasp. Such a grasp is of course likely to be difficult to obtain, but obtaining it would seem to be worth the effort. Having obtained such a grasp of F-ness, one’s judgements concerning F-ness would be completely reliable. One would be, in a sense, certain of one’s judgements concerning F-ness....While the Socratic model of knowledge I have uncovered may be implausible as an account of the contemporary model of knowledge understood as something like some kind of justified true belief, it may not be so implausible as an account of the contemporary (or otherwise) model of understanding [emphasis added] (2000, 211).

What emerges from the notion of understanding outlined above is its conceptual affinity with the notion of self-knowledge discussed earlier. For as with understanding, which requires a consistency in one’s body of epistemic beliefs about the world, self-knowledge also requires a consistency in one’s epistemic beliefs and cognitive states about oneself; for as Kekes correctly observes “the object of self-knowledge is the knower’s character” (Kekes 1995, 115). For it is through self-knowledge that we can hope to discover, by critical self-reflection, inconsistencies between our beliefs, desires and other cognitive states concerning ourselves, the world and others. As such, self-knowledge can also be thought of as a type of understanding – self-understanding. The Socratic adage “know thyself”

inscribed on the Apollo's temple at Delphi can therefore be thought of as an exhortation to self-understanding, an essential pre-requisite to the acquisition of wisdom.

The Essential Characterising Features of Wisdom

Following-on from the above discussion concerning the nature of wisdom, we can say in summary that wisdom comprises at least the following essential characteristics:

1. It is a capacity to discover and achieve what is desirable and of value in life; a capacity to use and develop knowledge and understanding needed for the realization of value. Thus wisdom includes knowledge and understanding (reference to Maxwell, 2007). It is a psychological capacity to judge rightly what should be done to make life better. This psychological capacity can be identified as a character trait (reference to Kekes, 1995).
2. It is the virtue of reflection. It is a second-order virtue whose primary aim is the development of our character. Whereas first-order virtues guide our actions in view of our conception of a good life, second-order virtues such as wisdom, guide our actions with a view of developing the kind of character that reflects our conception of a good life (reference to Kekes, 1995).
3. It is a second-order reflexive virtue. The subject who has the virtue is the same as the object toward whom the appropriate actions are directed. The primary motivation behind the second-order and reflexive activities directed by wisdom is the desire to make our lives better (reference to Kekes, 1995). Notice the similarity between Kekes and Maxwell in identifying the motivation for wisdom as the desire for making our lives better (see (1)). This establishes the essential conceptual and motivational connection between wisdom and the desire for a good life. As I mentioned also before, notice how within the conception of a eudemonic model of a good life the desire for a good life is motivated and guided by virtue: the overarching second-order virtue of wisdom.
4. It is concerned not merely with means but also with ends (reference to Kekes, 1995). This characterisation of wisdom is closely aligned and in keeping with Gewirth' Principle of Generic Consistency (PGC) as applied within the Model for the Unity of the Right and the Good (MURG) that judges the goodness of actions both with regard to their means as well as to their ends.
5. It aims at the transformation of our character. The character we desire to have in relation to our conception of a good life. It is an intrinsic good and

- not merely an instrumental good for the means of attaining our conception of a good life, because the character we desire to have is an essential part of that conception (reference to Kekes, 1995).
6. It is sound judgement involving the application of knowledge of good and evil to the evaluation of both means and ends constitutive of good lives (reference to Kekes, 1995). Judgement is a process by which a decision is reached about what to do or not to do, given the good the agent wants to achieve and the agent's concrete situation (reference to Kekes, 1995).
 7. It involves a process of increasing our control by enlarging the area of our lives that we can order so as to conform to our conception of a good life. It is a process that requires our becoming the person our conception requires (reference to Kekes, 1995). Notice how this process is in keeping with the notion of wisdom as a second-order reflective value whose primary aim is the development of character (see (2)).
 8. It is a type of self-knowledge whose object is the knower's character. It is a mode of reflection involving judgement whose aim is to make our character less fortuitous and more deliberate. The process of acquiring self-knowledge is thus directed at the transformation of our character. In this way, judgement and control are conceptually connected. Self-knowledge is essential to this process because it is through it that we learn what the internal obstacles are and what we can do to cope with them or overcome them. In this way we acquire our moral identity that motivates us to the transformation of our character in relation to our conception of a good life (reference to Kekes, 1995).
 9. It is a type of understanding that requires an individual to combine pieces of information into a unified body of knowledge. Unlike the accumulation of piecemeal bits of information and knowledge, understanding requires the grasping of explanatory and coherence-making relationships in a large and comprehensive body of information (reference to Kvanvig, 2003). As I mentioned earlier, understanding bears a close conceptual affinity with self-knowledge so that we can think of self-knowledge as a form of self-understanding. Ultimately, wisdom can be thought in a sense, as a type of self-understanding in relation to a conception of a good life that motivates and enables both its pursuit and its attainment.
 10. Finally, wisdom itself unlike information and knowledge is not transferable. Being necessarily embedded in one's character wisdom can only be acquired. Of course we can learn about wisdom and how to go about acquiring it, but its acquisition must be achieved by us individually. Thus

although knowledge about wisdom is transferable, wisdom itself is not. To use a stoic term, the acquisition of wisdom is a *kathorthoma* or achievement not unlike, for example, training for and successfully running and completing a marathon. You can learn of course what you must do in training for a marathon but ultimately it is only you who can run the marathon, no one else can do it for you. Similarly, the acquisition of wisdom is of necessity a uniquely individual achievement.

The Application of Wisdom for the Normative Evaluation of Digital Information in Relation to the Concept of a Good Life

Insofar as the ultimate purpose of a good life is the attainment of self-fulfillment leading to eudemonia or happiness (see part 1) then wisdom, which both informs the conception of a good life and directs its active pursuit for the attainment of eudemonia, is an essential condition for both the conception and the attainment of a good life. As the essential condition for both the conception and guided active pursuit and successful achievement of the good life, wisdom, as a type of meta-information that consists of at least the ten essential characteristics discussed above, is thus established as the essential conceptual connection between information and the good life. This, in turn, allow us to determine some of the generic implications of information for the conception of a good life, in particular, a eudemonic conception of a good life. Recall, however, that as I mentioned earlier and following Kekes, this eudemonic conception of a good life is only intended as a “regulative ideal that specifies some general conditions to which all good lives must conform” (Kekes 1995, 24) As such, the eudemonic account of a good life argued for in this paper is broadly speaking pluralistic as it is in principle compatible with other different conceptions of a good life that meet the same necessary general conditions to which any notion of a good life must conform.

Having discussed generically the epistemological and ethical implications of information as a process of communication for individuals and society generally in part (1), I will in this part of the paper outline only a few of the main axiological and eudemonic implications that the production, dissemination and communication of digital information might have for the good lives of individuals and society generally, through the application of the notion of wisdom as expounded in section (2).

More Information but not More Wisdom

Wisdom as a type of second-order knowledge involves reflection, sound judgment, and understanding in the use and development of information in the form of first-order knowledge (both knowledge-that and knowledge-how) to be ap-

plied with the aim of guiding and directing our thoughts and actions in the discovery and achievement of what is of value for us in life, in relation to our overall conception of a good life.

New Media, including the internet and computers generally have made it possible for us to access and use inordinate large quantities of information. However, what seems to follow from our characterization of wisdom above is that the uncritical access and use of so much information without the appropriate reflection, judgment and understanding, might not be conducive to wisdom and consequently might not be conducive to a good life or a better life. Thus, the uncritical accumulation and use of more information is not necessarily conducive to more wisdom and hence not more conducive to a good life. On the contrary, sometimes we might be better off with less information rather than more, especially if the former is directed by wisdom (less information) and the latter is not (more information).

Even in the case of critically accessing and using contextualized information, such information might also not be conducive to wisdom and to a good life, if that information is merely used instrumentally without a clear understanding of the ends which that information is intended to achieve or a clear understanding of the value of those ends. Recall that wisdom not only directs the means but also the ends of our actions. Thus the accessing and use of a lot of contextualized information on our I-Phone, for example, without a clear understanding of the value of the ends for which that information is to be used, might at best be neutral with regard to the goodness of our lives and worse detrimental to the goodness of our lives if it utilizes too much of our cognitive and social resources for the acquisition of information that ultimately is of little or no axiological or eudemonic value for us.

Face Book and You Tube: Only Fools Rush In

The uncritical and sometimes thoughtless dissemination of a lot of trivial and personal information of oneself and others on Social Network Sites (SNS) such as Face Book and You Tube, might be not only not conducive to wisdom and a good life overall, but detrimental to them if they encourage unreflective, foolish and reckless behavior with no apparent compensating axiological or eudemonic value for oneself and one's "friends". Given the pitfalls of creating and accessing unreflectively information about oneself and others, without any compensating realization of value in relation to one's conception of a good life, a wise person would thus be best served in exercising caution in using SNS. Such overall reflective caution is also in keeping with another of wisdom's characteristics, namely, control; that is to say, the ability to narrow down the areas in one's life over which one has little or no control so as to enlarge the areas of our lives that we can order

in conformity with our conception of a good life (Kekes 1995). This is a process of self-knowledge, a process of reflection involving judgments “whose aim is to make our character less fortuitous and more deliberate” (Kekes 1995, 127-28). When we allow ourselves to unreflectively disseminate information of ourselves and others on the internet we make ourselves more “fortuitous” and less “deliberate” and in so doing become hostages to fortune since we no longer have any control over that information, which now for ever floats beyond our control in cyberspace. Surely such unreflective conduct is not only not wise but foolish and ultimately self-defeating if, upon careful reflection, it undermines our considered conception of a good life.

Conclusion

The above examples are merely used illustratively and as proof of concept to demonstrate how the notion of wisdom as the essential link between information and a good life can be used methodologically to evaluate both axiologically and eudemonically the access, use and dissemination of any type of digital information on the internet and on computers generally. Some digital informational evaluations such as, for example, the use of Skype to communicate regularly with close friends and family will be positive given that relationships with friends and family are conducive to a good life and as such, their cultivation and maintenance is generally, wise-worthy. Other digital informational evaluations, such as the example from Face Book above, might prove negative because not conducive to a good life and therefore not wise-worthy.

Although descriptively it can be said that we are now living in the age of information, prescriptively we should, if we are wise, aim at promoting the Age of Wisdom, both for our own sake and that of others, and especially for the sake of future generations who might mistake mere information for wisdom.

REFERENCES

- Benson, H. H.** (2000). *Socratic Wisdom: The Model of Knowledge in Plato's Early Dialogues*. Oxford: Oxford University Press.
- Beyleveld, D.** (1991). *The Dialectical Necessity of Morality: An Analysis and Defense of Alan Gewirth's Argument to the Principle of Generic Consistency*. Chicago: University of Chicago Press.
- Brey, P.** (2000), *Disclosive Computer Ethics: The Exposure and Evaluation of Embedded Normativity in Computer Technology*, *Computers and Society*, 30(4), 10-16.

Bynum, T. (2008) Toward a Metaphysical Foundation for Information Ethics American Philosophy Association (APA) Newsletter on Philosophy and Computers, Fall 2008, Volume 08, Number 1, 12-16.

Dretske, F. (1999) Knowledge and the Flow of Information. Stanford, CSLI Publications.

Dreyfus, H. L. (2001) On the Internet. Routledge.

Eco, U. (1989) The Open Work, translated by Anna Cancogni. Cambridge, MA., Harvard University Press.

Finnis, J. (1983) Fundamentals of Ethics. Clarendon Press Oxford.

_____ (1980) Natural Law and Natural Rights. Clarendon Press Oxford.

Floridi, L. (2005), Is Semantic Information Meaningful Data? Philosophy and Phenomenological Research, Vol. LXX, No. 2.

_____ (2002), What is the Philosophy of Information? Metaphilosophy, Vol.33, pp. 123-145.

_____ (1999), Information Ethics: On the Theoretical Foundations of Computer Ethics. Ethics and Information Technology, 1(1), 37-56.

Floridi L. and. Sanders J. W. (2002), Mapping the Foundationalist Debate in Computer Ethics. Ethics and Information Technology 4(1), 1-9.

Gorniak-Kocikowska, K. (1996), The Computer Revolution And the Problem of Global Ethics, Science and Engineering Ethics.

Gewirth, A. (1996) The Community of Rights. Chicago, University of Chicago Press.

_____ (1998a) Self-fulfillment, NJ: Princeton University Press.

_____ (1978) Reason and Morality, Chicago, University of Chicago Press.

Kekes, J. (1995) Moral Wisdom and Good Lives, Ithaca, Cornell University Press.

Kvanvig, J.L. (2003) The Value of Knowledge and the Pursuit of Understanding. Cambridge University Press.

Maxwell, N. (2007): From Knowledge to Wisdom (Second Edition), London: Pentire Press.

Spence, E. (in press, 2009) The Epistemology and Ethics of Internet Information, in a Springer book of selected papers from the International Association of Information Systems (Italian Chapter) itAIS Conference, Paris, 13-14 December, 2008.

_____ (in press 2008) Media Corruption. *International Journal of Applied Philosophy*.

_____ (2008a) Understanding Luciano Floridi's Metaphysical Theory of Information Ethics: A Critical Appraisal and an Alternative Neo-Gewirthian Information Ethics. Invited refereed paper for the American Philosophy Association (APA) Newsletter on Philosophy and Computers, Fall 2008, Volume 08, Number 1.

_____ (2007a) What's Right and Good about Internet Information? A Universal Model for Evaluating the Cultural Quality of Digital Information. In Larry Hinman, Philip Brey, Luciano Floridi, Frances Grodzinsky and Lucas Introna (eds.), *Proceedings of CEPE 2007, The 7th International Conference of Computer Ethics: Philosophical Enquiry*, University of San Diego, USA, July 12-14 2007, ISSN 0929-0672.

_____ (2007b) Positive Rights and the Cosmopolitan Community: A Right-Centred Foundations for Global Ethics, *Journal of Global Ethics*, Volume 3 Issue 02, July/August 2007.

_____ (2006) *Ethics Within Reason: A Neo-Gewirthian Approach*. Lanham, Lexington Books (a division of Rowman and Littlefield).

Spence, E. and Quinn, A. (in press 2008) Information Ethics as a Guide for New Media. *Journal of Mass Media Ethics*.

Tiberius, V. (2008) *The Reflective Life: Living Wisely With Our Limits*. Oxford: Oxford University Press.

Varelius, J. (2004) Objective Explanations of Individual Well-Being. *Journal of Happiness Studies* 5: 73-91.

Vitek, B. and Jackson, W. (2008) *The Virtues of Ignorance: Complexity, Sustainability, and the Limits of Knowledge*. The University Press of Kentucky.

Landscapes of Ethical Issues of Emerging ICT Applications in Europe

Bernd C. Stahl and Simon Rogerson*

Centre for Computing and Social Responsibility
De Montfort University

Abstract

A central problem of the attempt to bring to bear ethical ideas to technology, in particular information and communication technology is that it tends to arrive too late. It would clearly be desirable to have a better understanding of future technological developments in order to allow ethical considerations to influence design and implementations of technology. At the same time, the unavoidable contingency of future developments provides a fundamental limit to what we can know about the future. The current paper explores how this problem can be addressed. It provides a framework of technical developments that one can reasonably expect to materialise in the medium term future (10 to 15 years) and ethical issues that are currently expected to arise. This is done by analysing cur-

* *Bernd Stahl* is Reader in Critical Research in Technology within the Centre for Computing and Social Responsibility and Teacher Fellow in the Faculty of Technology at De Montfort University, UK. His interdisciplinary interests are demonstrated by Masters-level or higher academic degrees in business, economics, engineering, philosophy, and law. He has studied or taught in Germany, France, Ireland, and the UK and exhibits a high level of cultural awareness in teaching and research. He has published three scholarly books (two single authored, one edited) and more than 100 papers in refereed journals, books, or conference proceedings. He is the editor-in-chief of the International Journal of Technology and Human Interaction (www.ijthi.net) and associate editor of the European Journal for Information Systems and the International Journal of Doctoral Studies

** *Simon Rogerson* is Director of the Centre for Computing and Social Responsibility at De Montfort University, UK. He is Europe's first Professor in Computer Ethics. He received the 2000 IFIP Namur Award for outstanding contribution to the creation of awareness of the social implications of ICT. In 2005 he received the prestigious SIGCAS Making a Difference Award by the ACM. Following a successful industrial career where he held managerial posts in the computer field, he now combines research, lecturing and consultancy in the management, organisational and ethical aspects of information and communication technologies. His current research focuses on technological assessment, ethical systems development and qualitative stakeholder analysis. He is co-editor of the Journal of Information, Communication and Ethics and Society. He conceived and co-directs the ETHICOMP conference series on the ethical impacts of ICT and created the world's leading portal on computer ethics.

rent European research funding documents with a view to exploring the trends, purposes, applications, artefacts, ethical issues, and governance structures that the European Commission foresees. The overall aim of the paper is to provide the conceptual basis of this framework.

Keywords: ICT ethics, emerging technologies, applications, research framework, Europe, FP7

Introduction

A central problem of the ethics of technology is that it tends to arrive too late. In many cases ethical issues are only recognised when the technology is already on the market and problems arise during its wide-spread use. Ethics can then become a tool to clean up a mess that might have been avoidable. It is probably not contentious to say that it would be desirable to have ethical input at the earlier stages of technology design and development. Indeed, there are ethical theories and approaches that explicitly aim at an early integration of ethics into the technology life cycle (van den Hoven, 2008). One central problem of this type of approach is that the future is unknown. By definition we do not know with certainty what will happen in the future and an ethics that relies on future development needs to be able to answer the question how it decides which technological developments to pursue. Ethics has traditionally not been well equipped to deal with issues of uncertainty (Sollie, 2007) and in particular future uncertainty.

The present paper aims to contribute to this discussion. Its approach is to identify likely scenarios of future ICT developments that are grounded in empirical facts. The idea is thus to strike a balance between unavoidable speculation when talking about the future and factual grounding necessary for academic research. This paper should be understood as a first step in identifying future developments in ICT. The chosen approach is to concentrate on an identifiable and relevant regional and policy area, namely the European Union. It aims to give a high level overview of the European landscape of emerging information and communication technologies. Its purpose is to come to an understanding of the ICTs that are likely to develop in the next 10 to 15 years with a view to understanding which ethical issues we can expect and how we may best prepare to meet them. This will lead to policy recommendations for the European Union as well as advice for individuals and organisations involved in technology development. The paper is meant to provide the grounding necessary to develop empirical work. It will develop categories of ICTs and ethical issues which will be used to investigate specific ICT research projects in order to assess whether and how ethical issues are currently taken into consideration and how policies need to be developed.

The paper will start by giving an introduction to the European research structure with a particular emphasis on the 7th Framework Programme. It will then analyse how the EU views the development of ICTs by analysing trends, applications, artefacts, ethical issues, and governance approaches as they can be identified from current EU publications.

Conceptual Basis

Before we move to a detailed discussion of the European landscapes of technology, it is important to outline briefly the context of this paper. This will start with a description of the 7th Framework Programme. We then outline our concept of ICT ethics. Finally, we discuss some of the policy background that informs the European Union's view of technology, and gives reasons for the plans and resource allocations that are meant to shape the development and use of technology. The section finishes with some considerations concerning the methodology of this paper.

The 7th Framework Programme for Research and Technological Development (FP7)

The European Union has a long tradition of fostering research and development through so-called framework programmes. The current 7th framework programme (FP7), which runs from 2007 to 2013, has a total budget of over €50 billion. The majority of this money is and will be spent on research grants, predominantly in Europe. Research to be co-sponsored by such grants is chosen on the grounds of calls for proposals and following a competitive peer review process. Given that there are national research funding mechanisms in many European countries, the European framework funding has the additional characteristic of being centred on international collaboration. The European Commission names two main aims of the framework programme (European Commission, 2007, p.7): «to strengthen the scientific and technological base of European industry [and] to encourage its international competitiveness, while promoting research that supports EU policies.» One aspect of FP7 is that it is meant to contribute to the European Research Area (ERA) (http://cordis.europa.eu/era/concept_en.html), which aims to overcome the weaknesses of European research caused by its fragmented and dispersed nature.

In order to meet the broad objectives of FP7, the programme has been divided into four categories: Cooperation, Ideas, People and Capacities. Each of these is then sub-divided into further categories and sub-programmes. The core of FP7 is the Cooperation programme, which is used to fund collaborative projects involving partners from at least three European Member States. This programme has been further divided into ten key thematic areas:

- health;
- food, agriculture and fisheries, and biotechnology;
- information and communication technologies;
- nanosciences, nanotechnologies, materials and new production technologies;
- energy;
- environment (including climate change);
- transport (including aeronautics);
- socio-economic sciences and the humanities;
- space;
- security.

The Ideas programme aims to support «frontier research», and funding is based on scientific excellence, without the need for cross-border collaboration. The People programme supports researcher mobility across Europe and the Capacities programme aims to strengthen research capacities of Europe. This paper will concentrate on the Cooperation programme and, more specifically, on its Information and Communication Technologies sub-programme. This is justified by the particular emphasis on ICTs of the project. While it stands to reason that ICTs will be developed in other areas of FP7, the ICT work programme is specifically focused on them. In addition, the ICT programme is the largest of all sub-programmes with a budget of over €9 billion over the lifetime of FP7 (http://cordis.europa.eu/fp7/budget_en.html).

A final word of justification of the choice of concentrating on European ICT research programme needs to be said given that this paper aims to investigate the global phenomenon of ethics in ICTs. One might ask why a paper published in a special issue on software engineering in the digital world should concentrate on a particular regional funding scheme. In addition to the practicalities of this paper being a result of a European research project, one can also easily argue that the EU is one of the most important economic and political entities and that its research policy has the potential of shaping future technical and economic standards. With a population of around 500 million and a gross domestic product (GDP) that represents about one third of the world's GDP, it has significant international power. The European view of ICT is important because it is developed in intercultural discourses with scientists and researchers worldwide. It shows the ways that policy makers perceive the role of ICT. At the same time, it has the potential to shape future developments. This refers to the funding available via FP7

but maybe more importantly to the many ways in which the European Union has to set policy that can shape the way technology is designed or used. While this paper and the underlying project are thus concentrating on a particular region, we believe that our findings should be of interest more generally and are likely to be transferable at least to a considerable degree.

ICT Ethics

Ethics can then be defined as the philosophical study or reflection of morality (Adam, 2005; Weil, 1969). In everyday language and even in much academic writing this distinction is not always observed (Forester, 1994; Weckert, 1997). However, the distinction between social norms and their reflection is important to observe if one wants to come to a measured understanding of normative issues and their ethical evaluation. Ethics as the reflection of morality can have different tasks. There is a distinction between descriptive ethics, normative ethics and metaethics (Marturano, 2002).

In this project the term «ICT ethics» will be used to denote ethical issues that arise from or in conjunction with ICT. Work in ICT ethics can be distinguished along the lines of the earlier distinction of ethics in general, namely in descriptive, normative, and metaethical. The different types of investigation are often undertaken by scholars from different disciplines. Descriptive ICT ethics work is typically done by researchers with a technical, social science or information systems leaning (Moores and Chang, 2006). Normative and in particular metaethical work is frequently undertaken by scholars with a background in philosophy (Bynum, 2006; Floridi, 2006; Introna, 2002; van den Hoven, 1997).

Research in ICT ethics is often multidisciplinary and attempts to come to a broad understanding of the subject at hand. Much research is focused on specific issues and problems. Among the most prominent ones one can find issues such as privacy (Brown, 2000; Introna, 2003), intellectual property (Burk, 2001; Syme and Camp, 2002), access and digital divides (Rooksby and Weckert, 2006), data quality (George, 2002), but there are many others. It often overlaps with related discourses in neighbouring disciplines, e.g. computer law (Poullet, 2004).

Much work in ICT ethics engages with the normative question how normative problems can be addressed in an ethically sound way. A typical approach that tends to be adopted is the adoption of some sort of behavioural guideline, policy or code (Siau, Nah, and Teng, 2002). Some of the most important professional bodies have gone down the route of a code of ethics, for example the British Computer Society (<http://www.bcs.org/server.php?show=conWebDoc.1587>) or the Association for Computer Machinery (<http://www.acm.org/about/code-of-ethics>). Codes of ethics can raise as many problems than they solve (Fairweather,

2000; Ladd, 1985). Alternative forms of governance have therefore been discussed.

The aim of this paper is not to champion any of the applications or approaches but to develop a framework that will allow us to capture work currently going on with a view to providing a more holistic understanding of research questions and expected future developments.

Policy Aims

Current public policies, in particular EU regulations, are pertinent to issues of ICT ethics and influence the outcomes of our paper. Normative perceptions and their ethical evaluation strongly influence what democratic governments perceive as issues to regulate. In current EU policy there are several areas where normative and ethical issues of ICT are addressed. ICT research has been identified as one of the three pillars of the «i2010 – A European Information Society for growth and employment» initiative of the European Commission. i2010 is renewing the Lisbon agenda and relies heavily on ICT to realise efficiency and economic gains. (http://ec.europa.eu/information_society/eeurope/i2010/introduction/index_en.htm)

The EU furthermore views ICT as a main tool in addressing its demographic challenges. In its Green Paper «Confronting demographic change: a new solidarity between the generations» (European Commission, 2005), the Commission has outlined the challenges the Union is facing. The demographic development continues to be a main area of concern for the EU (cf. «The demographic future of Europe – from challenge to opportunity», European Commission, 2006). Three general trends combine to create the problem of decreasing population: continuing increases in longevity, continuing growth of the number of workers over 60, and continuing lower birth rates. The EU intends to address the resulting problems with a variety of strategies. Among them there is the aim to use ICT to allow older people to remain an active part of society but also to allow them to remain independent in their homes. This has economic implications for health and social care but, more importantly, it is a matter of the quality of life for EU citizens.

The aims of the European ICT policy are broad and arguably contradictory. The aim of increasing competitive advantage, for example, can very easily lead to the use of ICT to replace traditional workplaces. Wiring companies and creating digital infrastructure can have the unintended result of facilitating outsourcing, thus further limiting the stated European aim of creating employment. To some degree the question of the net effect of technical development on the labour market is an empirical question. However, the ICT policies can also be contradictory in other aspects. The inclusion of disadvantaged groups in social processes is a

highly ethically relevant aspiration. At the same time, evidence from the digital divides literature suggests that the provision of technology can exacerbate existing barriers to social participation. The EU is aware of this and digital inclusion with all its implications is high on its list of priorities. An interesting question remains, however, whether general policy aims and the ICT research agenda that is investigated in the present paper are consistent.

The first section of the ICT work programme 2009 summarises the policy aims behind the EU ICT research initiatives (European Commission, 2008a, p. 4) as:

«Improving the competitiveness of European industry and enabling Europe to master and shape future developments in ICT so that the demands of its society and economy are met ICT is at the very core of the knowledge-based society. Activities will continue to strengthen Europe's scientific and technology base and ensure its global leadership in ICT, help drive and stimulate product, service and process innovation and creativity through ICT use and value creation in Europe, and ensure that ICT progress is rapidly transformed into benefits for Europe's citizens, businesses, industry and governments. These activities will also help reduce the digital divide and social exclusion. «

Methodological Considerations

While this paper is fundamentally of a conceptual nature and explores possible and likely futures to allow the development of more detailed research agendas, it nevertheless needs to be grounded in a shared social reality to gain acceptance of the variety of audiences who have an interest in ethical issues of emerging ICTs. In order to provide a transparent and shared account of likely developments, the empirical basis of the paper was based on a content analysis of a range of sources. Primary among these were documents created by the European Union with regards to policy planning, in particular of the 7th Framework Programme. In order to supplement and contextualise these, other sources on ICT, its future developments and ethical issues were considered. The content analysis was conducted by reading the documents with a view to the following items: applications of future technology, artefacts, ethical issues, governance structures, and others. The findings of the analysis were stored in a mindmap for easier reproducibility and then used for summarising the findings below.

European Landscapes

This section shows the major areas of technological development in ICT as well as ethical and governance aspects related to it. It is broken down according to the main items used for the data analysis: trends, applications, artefacts, ethical issues, and governance structures. The first attempts to provide an overall view

of where ICT is going. The two sections on applications and artefacts relies heavily on the most recent call for ICT projects at the time of writing this document, which is the FP7 ICT Call 4, published on 19 November 2008 with a submission deadline of 01 April 2009 (European Commission, 2008a). This is the document that explains in most detail the aims and objectives of the ICT work programme and thereby gives an exact view of what European policy makers believe to be desirable and realistic. Further documents are drawn upon where necessary.

ICT Trends

Those who have tried to forecast the next technological advances are usually incorrect. ICT has a track record of unpredictability in the specific nature and consequent impact of these future advances. The only certain thing is that there will be always be significant advances and these will always impact upon society and its people. However, several general ICT trends can be seen even though the specifics are unpredictable. It is such trends which influence the overall strategic approach, for example, to national and European research funding and to societal acceptance or rejection of technology. Vaughn (2006 pp8-14) suggests that there are four key ICT trends.

- ICT trend 1: Ever-increasing computational power plus decreasing size and cost

The move towards more computational power, with decreased size and cost, can make possible improved and entirely new types of technology and new application opportunities.

- ICT trend 2: Technology advances enabling new types of interfaces

The human interface is one of the most important determinants of whether a technology product can be used by people regardless of their skill, experience, affliction or disability. For example, advances in interface technology are creating new opportunities for better assistive technologies, more accessible mainstream technologies, and entirely new concepts for controlling both. Some of the more innovative interfaces include augmented reality, hands-free operation, voice control and direct control from the brain.

- ICT trend 3: Ability to be connected anywhere, anytime with services on demand

The latest innovations such as wireless electronics, location awareness, wearable technology and implantable technology point towards a society with widespread connectivity. This allows people to think about communication, control and presence in entirely new ways.

- ICT trend 4: Creation of virtual places, service providers and products

Web technologies have provided people with new ways of doing things hitherto not thought of or not possible. Such technologies have fostered the development of entirely new social, commercial, and educational concepts.

The evolution of ICT through such trends could impact upon everyone both positively and negatively. This is explored in detail by both Roe (2007) who uses a SWOT analysis and Vaughn (2006) who considers opportunities and barriers. Floridi (2007) considers such trends at a meta level and argues that 'in information societies, the threshold between online and offline will soon disappear, and that once there won't be any difference, we shall become not cyborgs but rather inforgs, i.e. connected informational organisms.' If this is the case then the ethical dimension of ICT becomes the ethical dimension of society per se.

Applications and Challenges

In this paper we distinguish between applications and artefacts that may give rise to ethical issues. This distinction is not reflected in the ICT Call 4 (European Commission, 2008a) but it is plausible to make this distinction. By applications we mean areas where ICTs can lead to solutions or applications. Artefacts, on the other hand, are software, hardware, or related items that can be used for particular purposes. There is often a relationship between applications and artefacts and in many cases artefacts are built specifically for particular applications. However, artefacts can usually be used in different ways and different artefacts can be used for the same applications. Since ethical issues can arise in a number of ways, including the non-intended use of artefacts, we believe that the analytical distinction between applications and artefacts is helpful to our overall aim of identifying ethical issues of emerging ICT applications.

The applications that the EU views as relevant for the next 10 to 15 years are reflected in the main challenges. These challenges are divided into two groups: «overcoming technology roadblocks and reinforcing Europe's industrial strengths» and «seizing new opportunities and applying ICT to address Europe's socio-economic challenges». The first group contains those challenges which can be seen as technological in nature, which seems to imply that their social and economic context is less important or maybe unproblematic. The first one of the three technical challenges has to do with «pervasive and trustworthy network and services infrastructure.» Its content is based on the perception that current network infrastructures, in particular the Internet, is problematic and needs to be replaced soon. The second technical challenge aims at context-aware and easy to use technologies. These are perceived to be a key technology that can further policy objectives in a number of ways. The work programme therefore calls its second main challenge that of «cognitive systems, robotics and interaction.» The

third and final technical challenge has to do with «electronic components and systems.» These are seen to be crucial for the development of the next generation of technologies and therefore as a central basis for further innovation in products and services. It is interesting to note that for all of the three technical challenges the work programme is silent on their expected consequences and link to policy goals. This implies a pervasive belief that technological progress is desirable because of its knock-on effects, the hope that it will lead to successful products, higher competitiveness and thereby to well-being and employment.

The second group of challenges, the socio-economic ones, are more immediately and more visibly linked to the European policy goals. The four challenges meant to address «Europe's socio-economic challenges» are aimed at specific areas where technology is perceived to have a crucial role. The first one is the area of «libraries and digital content». Under this heading, one can find research aimed at digitising libraries and cultural heritage. It also includes a section on technology-enhanced learning and one on intelligent information management. The second challenge addresses issues in relation to sustainable and personalised healthcare. This one is linked to the increasing costs of sophisticated healthcare that are set to further spiral because of the changing European demographics. The challenge is split in three main groups, one on personal health systems, one on patient safety and one on virtual physiological humans, which covers simulations of humans for training and research purposes. The third challenge centres on ICT for mobility, environmental sustainability and energy efficiency. Among the aims here, one can find a range of aims related to efficiency, mobility, environmental protection and distribution of energy. The fourth and final challenge on «ICT for independent living, inclusion and governance specifically» aims at developing applications for ICT related to ageing, accessible and assistive ICT, as well as ICT for governance and policy modelling.

Together these seven challenges represent the applications that the European Commission sees as central to advance its policy agenda. They set the boundaries for the type of research that will be funded under the seventh framework programme. They are therefore likely to have an influence on the technologies that will become viable and wide-spread in the next decade. It is clear that this is not an exclusive list and that there are other development agendas from private organisations, nation states, NGOs e.tc. that are similarly worth exploring. For our purposes, however, the EU policy is of central interest and we will therefore concentrate on these applications.

Artefacts

While the applications and challenges are relatively easy to identify and list, the same cannot be said for the artefacts envisaged to attain the policy goals. A detailed reading of the Call 4 document displays a range of artefacts that are considered possible solutions to a variety of problems. In addition to physical artefacts, there is a strong emphasis on processes and procedures that may lead to products or services. Rather than try to identify all of the artefacts, this paper will briefly discuss some of the more speculative ones or ones that reoccur as specific artefacts to be emphasised.

The probably most notable such artefact is related to the future of networks and in particular to the Internet. This is the next generation of Internet Protocols, called Internet Protocol version 6 (IPv6). Mentioning of IPv6 reoccurs throughout the document. More importantly, the promotion of IPv6 is named as one of the strategic priorities of European ICT research policy.

In addition to IPv6, the call document goes on to enumerate a number of ICT artefacts that are currently of a rather speculative nature but that are seen as bearers of great potential that deserve to be developed. Interestingly, these artefacts are not linked to the challenges discussed in the preceding section and are therefore not clearly identifiable as contributors to the policy aims. Instead, they form a separate part of the call document, which is listed under the heading of «future and emerging technologies».

Given that the aim of the present paper is to provide a framework for the investigation of ethical issues of emerging technologies, these emerging technologies as outlined by the call document are of particular interest. As they are at a more exploratory stage, their actual conceptual and physical form are currently still uncertain, but the technologies suggested render it clear which way the development is expected to take. The first set of such emerging technologies has to do with high speed data processing and it is listed under «Concurrent Tera-Device Computing.» The next set of technologies is based on «quantum information foundations and technologies.» «Molecular-scale devices and systems» are suggested as a further important research area. Another predominantly technical area is that of «bio-chemistry-based information technology.» The attempt to use cross-disciplinary research in order to improve ICTs is furthermore developed in the «brain-inspired ICT.»

In addition to these technical challenges, there are also application-driven emerging technologies. The first one is «human-computer confluence» which explores new modalities for individual and group perception, actions and experience in augmented, virtual spaces. There is also an area of self-awareness in autonomic systems, which aims at an improvement of the interaction between computing

artefacts and their environment. Environmental concerns are reflected in the research towards zero-power ICT.

To some degree one can see reflected the distinction between purely technical considerations, which at this stage are not yet application-oriented and those that are specific to particular issues. An interesting question that will guide our further research is whether this more or less specific outcome focus of the artefacts raises particular ethical issues.

Ethical issues

The Seventh Framework Programme (Decision N°1982/2006/EC), Article 6 (1§) states that «All the research activities carried out under the Seventh Framework Programme shall be carried out in compliance with fundamental ethical principles.» The same Decision also states later on that «the opinions of the European Group on Ethics in Science and New Technologies [EGE] are and will be taken into account» in research under the Seventh Framework Programme. The emphasis on ethics is based on the recognition of the potential impact of ICT on human rights as established by the European Convention on Human Rights (<http://conventions.coe.int/Treaty/en/Treaties/Html/005.htm>) and developed by the Charter of Fundamental Rights of the European Union (http://www.europarl.europa.eu/charter/pdf/text_en.pdf). Such general considerations are complemented by other more specific statements, notably the extensive guidelines on addressing ethics included in the guides for applicants for FP7.

Despite this high level recognition of the relevance of ethics to ICT, it is worth exploring in more depth what is meant by ethics in the context of the EU ICT research programme and how it is to be addressed. It is easy to follow the EU policy assumptions that ICT has important ethical aspects and promises solutions to pressing social and ethical issues. At the same time, ICT can raise a host of new ethical questions.

The interesting question for the present paper is how these general ethical concerns are operationalised, and whether there is any guidance on the type of ethical problems that should be considered. There are several documents that offer guidance on how to recognise and address ethical issues. A helpful distinction to categorise different ethical issues is the distinction between ethical issues as arising out of the research process and ethical questions arising from research content. In its «ethical guidelines for undertaking ICT research in FP7» (European Commission, 2008b) the Commission lists a number of substantive issues that may result from emerging ICT. The first problem identified concerns the autonomy and privacy of potential users. Researchers are reminded that a responsible approach is required and that compliance with European and national legislation

is required. Further substantive issues discussed are those connected to specific technologies, such as implants and wearable computing, which have been elaborated by the European Group on Ethics. E-health is seen as a further area worthy of specific warnings as it poses particular problems to privacy and security. The same is true for nano and bio-electronics.

The same concerns that are included in the ethical guidelines are reflected in the Annex 4 of the guidance for applicants, which also forms a part of the proposal form. This annex has the form of a check list that covers informed consent, privacy, and ICT implants. Additional issues that are not further explained are research on human embryos/foetuses, research on animals, research involving developing countries, and dual use of ICT for military or terrorist purposes. The points on this list are further elaborated on in the «Ethics for Researchers» document (<ftp://ftp.cordis.europa.eu/pub/fp7/docs/ethics-for-researchers.pdf>).

Questions of the social consequences of wide-spread use of particular technologies, which in areas such as e-health could have foreseeable consequences, are not elaborated in much depth. The documentation is quiet, for example, on how one can recognise terrorist applications and how to address such issues. One could argue that such substantive ethical issues of emerging technologies are beyond the scope of concrete current guidance and that this type of issues should therefore be covered by procedures that allow researchers to be alerted to ethical questions. A look at the procedural guidance shows, however, that it is not geared to capturing such issues either. The two main aspects of procedural guidelines are to ensure informed consent and to comply with legislation. Both are well-established ways of dealing with issues arising from the process of doing research. Informed consent in particular is the cornerstone of ethical conduct of medical research. It is open to question, however, whether it is sufficient to deal with ethical issues arising from emerging technologies. It is interesting to note that the guidelines do acknowledge that there are likely to be, hitherto unrecognised, emergent ethical issues resulting from advances in ICT research. Due to the apparent reliance on procedural ethics, it is important to ask which procedures are envisaged in the governance structures of projects.

Governance structures

The most immediately visible aspect of governance has to do with ethics review of projects. Ethical review is described as one aspect undertaken by the panel of experts that undertake the scientific evaluation of a project. The panel of experts will identify a project as requiring special attention if «projects raise sensitive ethical issues or when applicants failed to address ethical issues in an appropriate way.» (Ethical Review Procedure, <http://ec.europa.eu/research/science-society/>

index.cfm?fuseaction=public.topic&id=130). All projects thus identified as well as all projects dealing with human intervention or human embryonic stem cells will undergo ethical review. This ethical review will be conducted by a panel of experts and it aims to determine whether the project follows the standards of ethics of FP7. Projects found to be in violation of such fundamental ethical issues are then excluded from funding.

Such external governance of projects is described in some depth, but there is little guidance on internal governance of research projects. There must be some explicit ethical governance, for example in order to ensure that the procedural human research issues, in particular informed consent, are addressed according to standards. This will presumably require some sort of ethics committee but the exact form of such a committee is not clear. Specifically with regards to dual use, the «Ethics for Researchers» document recommends the recruitment of an advisory board, which can advise the project consortium on issues societal, political, and legal aspects of potential applications, on exploitation and dissemination strategies. These ideas are probably valid for other ethical issues. In addition to such external guidance, the «Ethical Guidelines for undertaking ICT research in FP7» state that «activities may, if appropriate, include specific tasks or a specific work package that explicitly addresses ethical concerns (in terms of the research, its conduct and outcomes) and outlines how ethical issues raised by the proposed research will be handled.» Further guidance on how such work packages are to be defined, which membership is desirable or how they are to be integrated in the project is not given.

Summary

It is easy to imagine that there are further general categories of relevance to be explored for a better understanding of the ethical issues related to emerging ICTs. For our current purposes of charting a landscape of such ethical issues, the outlined categories offer enough of a differentiation to allow for a detailed picture of issues that can reasonably be expected to develop in the medium term future of 10 to 15 years. Table 1 below summarises the issues enumerated in this section. It is easy to see that there are numerous possible combinations of trends, applications, artefacts, and ethical issues, which allow questions of how they relate to policy aims or which type of governance structure would be likely to be able to address them. This table provides a high level summary of the landscape of emerging ICT ethics and can be used as a basis for further research.

Conclusion

This paper aims to provide an overview of current social, political, and technical developments with a view to provide a framework for further research. It has identified current EU policy with regards to ICT research, which will have mani-

fest effects when current research and development projects come to the stage of market entrance, ten to fifteen years from now. To be useful, the framework itself needs to be applied as the basis of further research. However, there are limitations to the current approach. Both points will now be discussed.

Further Research

The framework as summarised in Table 1 will be used as the conceptual basis of empirical research on the realities of current research projects. The conceptual landscape of emerging ICTs in Europe that was developed above is very much a reflection of what policy makers and scholars in leading roles imagine to be desirable and probable futures. A different question is whether and how these aspirations are reflected in current research projects. We will initially concentrate on current ICT research projects of the sixth and seventh framework programme. It is clear that these will reflect the aspirations in their published aims and objectives (otherwise they would not be funded) but it remains to be seen how these are implemented and which consequences are expected. Of particular interest will be whether and how future developments and their ethical evaluation will be considered. This empirical analysis of extant project will then allow us to come to an evaluation of the relevance and accuracy of the above review of the European landscape with a view to developing guidelines for European policy makers on how ethical issues in ICT can and should be addressed.

Limitations

This paper has said nothing about the political processes behind the development of the policy documents. This is academically problematic because a more detailed understanding of these processes would allow for a more comprehensive contextualisation of the issues raised in the different documents. Relying on published documents also presents a more monolithic view of policy objectives and the European ICT research landscape than is likely to be realistic. Policy documents are by their nature contested and express the results of multiple power struggles. This in itself is an interesting issue as power relationships tend to be of ethical relevance. In this case, technologies may be used to promote particular agendas and an uncritical acceptance of existing documents clouds the view of what alternative agendas might have been. This is ethically problematic and also questionable from an epistemological viewpoint as the positions that prevailed in the policy development process may not be the ones that best predict future development and use of ICT.

Trends	Purpose/ Policy Aim	Applications/ Challenges	Artefacts	Ethical Issues	Governance Structure
ever-increasing computational power plus decreasing size and cost	economic growth	network infra- structure	physical artefacts	research process	external gov- ernance
		cognitive sys- tems, robotics	processes and procedures	informed consent legal compliance	
	employment	components, systems, engi- neering	IPv6	research content	ethical review as part of sci- entific review
			concurrent terra- device computing	privacy/data protection	
advances enabling new types of inter- faces	demographic challenges solutions	digital libraries and content	quantum information foundations and technologies	ICT implants/ wearable com- puting	external advi- sory board
				e-health related issues	
ability to be connected anywhere, anytime with services on demand	social/politi- cal inclusion	healthcare	bio-chemistry- based informa- tion technology	nano- and bio- electronics	internal gov- ernance
				research on animals	
			human-computer confluence	research involv- ing developing countries	
creation of virtual places, service providers and products			sustainability	self-awareness in autonomous systems	research involv- ing human embryos or foetuses
		inclusion	molecular-scale devices and systems	dual use (mili- tary or terrorist applications)	informed consent proce- dures
			brain-inspired ICT		

Table 1: Summary of Emerging EU ICT Research Landscapes.

Despite these shortcomings of the chosen approach, we believe that it can be justified. There are practical reasons for our choice, namely that a more detailed investigation of the political processes behind the published documents would have gone beyond the scope of the research and also of this paper. More importantly, we argue that possible inadequacies of the picture derived at by relying on the documents we used will be borne out during the next stage of the research. It is important to keep in mind that the purpose of this paper is not to give a comprehensive account of what will happen but to develop a conceptual basis

for empirical research into existing and planned ICT research projects. The conceptual basis as represented in Table 1 will be tested with regards to its reliability and possible inadequacies can be catered for during the next stages of the project. Finally, the European policy framework explicitly states that it focuses on a «limited set of challenges» (European Commission, 2008a, p. 5), namely those which are best suited to fulfil policy aims. By implication this means that there are other challenges that are not considered but which the present research project needs to remain open to if it wants to achieve its goal of understanding emerging ethical issues.

Despite these shortcomings, we believe that the present paper makes a substantial contribution to our understanding of the relationship of ethics and ICT. It provides a framework that can be used to identify a range of different issues whose understanding is crucial for any approach to ethics of ICT that wants to be able to contribute to the solution of expectable problems. We are therefore convinced that this paper is of interest to a diverse audience ranging from software developers and project managers to organisational and political policy makers. While not all aspects of our framework may be completely transferable to other regional and organisational environments, we believe that many of the aspects discussed are likely to be pertinent in emerging technologies and that further research from a variety of backgrounds will be able to build on our findings.

REFERENCES

- Adam, A.** (2005). *Gender, Ethics and Information Technology*. Basingstoke: Palgrave Macmillan.
- Brown, W. S.** (2000). Ontological Security, Existential Anxiety and Workplace Privacy. *Journal of Business Ethics*, 23(1), 61-65.
- Burk, D. L.** (2001). Copyrightable functions and patentable speech. *Communications of the ACM*, 44(2), 69-75.
- Bynum, T. W.** (2006). Flourishing Ethics. *Ethics and Information Technology*, 8(4), 157-173.
- European Commission. (2005). COM(2005) 94 final: Green Paper "Confronting demographic change: a new solidarity between the generations". Retrieved December 10, 2008, from http://ec.europa.eu/employment_social/news/2005/mar/comm2005-94_en.pdf.
- European Commission. (2006). COM(2006)571 final: The demographic future of Europe – from challenge to opportunity. Retrieved December 10, 2008, from http://ec.europa.eu/employment_social/news/2006/oct/demography_en.pdf.

European Commission. (2007). FP7 in Brief: How to get involved in the EU 7th Framework Programme for Research. Retrieved November 26, 2008, from http://ec.europa.eu/research/fp7/pdf/fp7-inbrief_en.pdf.

European Commission. (2008a). Work Programme 2009; Cooperation; Theme 3, Information and Communication Technologies (European Commission C(2008)6827 of 17 November 2008). Retrieved November 26, 2008, from http://cordis.europa.eu/fp7/dc/index.cfm?fuseaction=UserSite.FP7DetailsCallPage&call_id=185.

European Commission. (2008b). Ethical Guidelines for undertaking ICT research in FP7 (Annex 5 to ICT).

Fairweather, N. B. (2000). No, PAPA: Why Incomplete Codes of Ethics are Worse Than None At All. In G. Collste (Ed.), *Ethics in the Age of Information Technology*. Centre for Applied Ethics, Linköpings Universitet, Linköping, Sweden.

Floridi, L. (2006). Information ethics, its nature and scope. *ACM SIGCAS Computers and Society*, 36(3), 21-36.

Floridi, L. (2007). A look into the Future Impact of ICT on Our Lives, *The Information Society*, 2007, 23(1), 59-64.

Forester, T. (1994). *Computer Ethics: Cautionary Tales and Ethical Dilemmas in Computing* (2nd ed., p. 304). M. I. T. P.

George, R. T. D. (2002). *The Ethics of Information Technology and Business* (p. 304). WileyBlackwell.

Van den Hoven, J. (1997). Computer Ethics and Moral Methodology. *Metaphilosophy*, 28(3), 234-248.

Van den Hoven, J. (2008). Moral Methodology and Information Technology. In K. Himma and H. Tavani (Eds.), *The Handbook of Information and Computer Ethics* (pp. 49-68). WileyBlackwell.

Introna, L. D. (2002). The (im)possibility of ethics in the information age. *Information and Organization*, 12(2), 71-84.

Introna, L. D. (2003). Workplace Surveillance 'is' Unethical and Unfair [Opinion]. *Surveillance & Society*, 1(2), 210-216.

Ladd, J. (1985). The quest for a code of professional ethics: an intellectual and moral confusion. In D. Johnson and J. Snapper (Eds.), *Ethical issues in the use of computers* (pp. 8-13). Wadsworth Publ. Co. Retrieved December 14, 2008, from <http://portal.acm.org/citation.cfm?id=2569.2570>.

Marturano, A. (2002). The role of metaethics and the future of computer ethics. *Ethics and Information Technology*, 4(1), 71-78.

Moore, T. T., and Chang, J. C. (2006). Ethical Decision Making in Software Piracy: Initial Development and a Test of a Four-Component Model. *MIS Quarterly*, 30(1), 167-180. doi: Article.

Poullet, Y. (2004). The fight against crime and/or the protection of privacy: a thorny debate! *International Review of Law, Computers & Technology*, 18(2), 251-273.

Roe, P. R. (ed) (2007) Towards an inclusive future: Impact and wider potential of information and communication technologies. a COST 219ter report. COST, Brussels.

Rooksby, E., and Weckert, J. (2006). *Information Technology and Social Justice*. Information Science Publishing.

Siau, K., Nah, F. F. H., and Teng, L. (2002). Acceptable internet use policy. *Communications of the ACM*, 45(1), 75-79.

Sollie, P. (2007). Ethics, technology development and uncertainty: an outline for any future ethics of technology. *Journal of Information, Communication and Ethics in Society*, 5(4), 293-306.

Syme, S., and Camp, L. J. (2002). The Governance of Code: Open Land vs. UCITA Land. *ACM SIGCAS Computers and Society*, 32(3).

Vaughn, J. R. (2006) Over the Horizon: Potential Impact of Emerging Trends in Information and Communication Technology on Disability Policy and Practice. National Council on Disability. information and communication technologies. a COST 219ter report. COST, Brussels

Weckert, J. (1997). *Computer and Information Ethics. Contributions to the study of computer science* (p. 175). Westport, Conn: Greenwood Press.

Weil, E. (1969). *Philosophie morale*. Librairie Philosophique J. Vrin.

Note

This paper is part of the research project, Ethical Issues of Emerging ICT Applications (ETICA) funded by the European FP7 Science in Society strand.

Security, Privacy and Technophobia in the School of the Future

Giannis Stamatellos*

Doukas School
International Baccalaureate

Abstract

Education is significantly influenced by the development of Information and Communication Technology (ICT). In future schools software applications, multimedia trends and the Internet have been effectively introduced. In smart classrooms the lesson is conducted with laptops, smart boards and electronic sources. School security is also improved with computerised measures of surveillance. The aim of this paper is to discuss the ethical and social issues that arise from the use of ICT in future schools with special reference to the problem of technophobia and the moral balance between security and privacy. It is supported that whereas technophobic concerns should not be considered as obstacles in student's education and learning potentials, a threat to students' privacy may result loss of autonomy, freedom and self-institution.

Keywords: Future School, Smart Classroom, Education, Technophobia, Security, Privacy, Autonomy

Introduction

Sigmund Freud once defined education - along with psychoanalysis and government - as an "impossible profession" in which, as he claimed, "one can be sure beforehand of achieving unsatisfactory results" (Freud 1961). Freud's unenthusiastic or even pessimistic comment on education's results should be reconsidered today from the perspective of information age. The rapid development of information and communication technology (ICT) renovates the state of affairs

* *Giannis Stamatellos* studied Computer Science and Analysis in Athens, Greece. He then turned to Philosophy and Classics where he gained a Master of Arts and a PhD at the University of Wales, Lampeter. Currently, he is professor of philosophy at the New York College in Greece and an IB teacher at Doukas School specialized in Ancient Greek Philosophy, Epistemology and Ethics. He is the author of the books *Plotinus and the Presocratics* (SUNY, 2007), *Computer Ethics: A Global Perspective* (Jones and Bartlett, 2007) and of various articles and translations on philosophy published in English and Greek.

of modern citizenship and introduces information society in cases where traditional moral questions are reexamined and new ethical problems emerge. The centralization, distribution and control of information result in a different kind of discrimination between the 'poor in information' and the 'rich in information'. Human relationships are affected by new parameters of distant and impersonal communication, while malfunction and misuse of ICT bring serious concerns to individuals and groups on computer reliability. These issues are inter-related and multi-divergent underlined by the special nature of digital information and social networking. Hence the impact of ICT extends in various areas such as politics and government, economics, health, science, education and entertainment.

Based on the above considerations, Freud's aforementioned claim should be restated: to what extent information society makes education's results more unpredictable and unsatisfactory? Today the area of education is not merely determined by human factors and the traditional inter-relationship between teachers and students but it is also significantly influenced by the development of advanced technologies [Baase (2003) pp. 295-307, Edgar (2002) pp. 160-161, Rosenberg (1997) ch. 5, Stamatellos (2007) ch. 9] realised in the so called school of the future [Brent and West-Burnham (2003) ch. 61, Bottino 2004)].

In future schools software applications (i.e. educational simulations, slide presentations and tutorials), computer hardware (i.e. laptops, video projectors, notebooks), multimedia trends (i.e. video, DVD) and the Internet (i.e. googledocs, wikis) have been effectively introduced. In smart classrooms the lesson is conducted with high-tech tools (i.e. smart-boards and wifi notebooks), while, in some cases, lessons are recorded for future educational purposes or absent students. Security is also improved since electronic passwords and biometrics (i.e. hand and iris scanners, facial recognition systems) commonly used to authenticate the identity of the students and to record their attendance; as well as surveillance technologies (i.e. CCTV) monitor various activities inside and outside the classrooms and the campus.

A well-known international example of such a future school is the School District of Philadelphia which, along with Microsoft in 2003, aimed to create the first high school that will offer educational innovation and advanced technological development in an empowered learning global community of continuous, relevant and adaptive education (Microsoft 2009). The School of the Future opened in September, 2006, with the following objectives to: (1) serve as an educational model that nurtures student achievement through holistic reform of secondary education; (2) apply research and development to generate educational practices, creating an environment that involves all stakeholders and that inspires a passionate, personal responsibility for learning; (3) incorporate best-of-class technology

solutions in nearly every area of the learning community, including curriculum delivery, community collaboration, back-office support, content creation, and dissemination of content and assessment (Philadelphia District 2009).

Another example, stated in Europe, comes from Greece and Doukas School, which aims to design and build the School of the Future progressively, through a developing action plan: (1) register the current and future educational needs; (2) study the international experience and methodology; (3) prepare the new technological infrastructure required; (4) prepare the teachers through appropriate education in new technologies; (5) collect - create the appropriate educational digital material (Doukas and Kotsanis 2008). As Doukas School declared in its "mission statement", the School of the Future is a combination of the past, the present and the future in which "Parents participate actively in the educational family. Students change from passive recipients, they participate actively in the educational events as well as teachers become collaborators and guides, they organize, inspire and create biomatic activities". On this the «Classroom of the Future» is introduced providing students with a light-weight portable "electronic schoolbag" – contains in a tablet personal computer students' books and sheets as well as virtual labs, simulations, multimedia material, their schoolwork and all tools connected to the educational process – as well as each student is connected wirelessly to the interactive whiteboard of their classroom, the internet and the "electronic" tools of their teacher. The classroom can take many forms such as "traditional" class, "laboratory" or "collaborative" thanks to the wireless equipment and portable computers (Doukas 2009).

On the other hand, despite the observed flexibilities and advantages of future schools and smart classrooms, serious ethical questions arise from the extended application of ICT. Teachers and parents express technophobic concerns as regards the extensive use of computers in future schools; including (1) students' dependence on computers; (2) deskilling and computer addiction; (3) student's dehumanisation and alienation by the extensive use of social networking, (4) the reliability, appropriateness, copyright and plagiarism of digital sources and (5) the problem of digital divide especially in areas inaccessible to computers and computer networks. Another serious concern refers to the application of ICT security in school areas: to what extent ICT security measures cause a threat to students' and teachers' privacy?

Based on the above concerns the aim of this paper is to discuss the ethical issues that arise from the use of information and communication technologies in the school of the future with special reference to the problem of technophobia as well as the moral balance between security and privacy in education; technophobia leads to an increase of security which in turn may cause an erosion of privacy. It is

going to be supported that whereas technophobic concerns should not be regarded as serious obstacles in student's educations and learning potentials, a threat of privacy may result loss of student's autonomy, freedom and self-institution.

Technophobia

Technophobia (or technological fear) includes extreme skepticism and doubt about technology. The first incident of technophobia could be found in Plato's *Phaedrus* and the mythical story of the invention of writing. Plato describes the meeting of the Egyptian god Theuth (the inventor of writing) and the Pharaoh Thamus (274e-275b). In Plato's myth (actually a criticism to writing in favor of oral tradition) the Egyptian god is enthusiastic with the discovery of writing as an improvement of memory and the development of wisdom and civilization. However, Thamus replied with skepticism that "You have not discovered a portion of remembering, but for reminding; you provide your students with the appearance of wisdom, not with its reality. Your invention will enable them to hear many things without being properly taught, and they will imagine that they have come to know much while for the most part they will know nothing. And they will be difficult to get along with, since they will merely appear to be wise instead of really being so." (Cooper 1997) Umberto Eco in his famous lecture 'From Internet to Gutenberg' traces back to Plato's account and further criticizes the problem of technophobia (Eco (1996). For Eco, technophobia originates in Thamus' criticism on invention of writing (writing will diminish memory), continues in the Gutenberg's invention of publishing (books will diminish writing) and the information age (the Internet will diminish books). Likewise, parents and teachers are concerned that ICT tools in future schools (1) weaken student's skills, promote mental laziness and computer addiction; (2) utilize electronic sources with unknown validity and methods of electronic learning with unjustified results; (3) promote digital divide and inequality of access; (4) bring the problem of student's dehumanization and alienation.

Firstly, in the school of the future, students may be de-skilled – transformed to a lower level of skill. For example, students learn mechanical skills by pressing a series of buttons without making manual calculations or taking decisions that exercise creativity and imagination. However, deskilling in particular human faculties could be regarded as the consequence of the information paradigm shifting: the paradigm shift from the industrial age to the information age. ICT brings new skills in light since old skills are replaced; students are equally up-skilled – upgraded to a higher level of skill; for example, students learn how to use software applications and multimedia tools, improve their learning skills by participating in forums, blogs and exchanging information on a knowledge subject via emails and sms (Doukakis, Chionidou-Moskofoglou and Mangina-Phelan (in press)).

Secondly, in future schools paperless electronic media tends to replace paper publishing; electronic books and internet sources will be the educational sources of the future. However, future schools seem to consider electronic sources and electronic learning from a western standpoint where ICT is always available, beneficial and valuable in the development of student's education. Electronic learning and the use of electronic sources raise a number of issues on the reliability and appropriate utility of digital sources: to what extent electronic sources are trustworthy and scientifically justified, censored or biased? To what extent search engines and other online sources guide the user properly to the appropriate learning material? Are the e-sources secured by individuals against plagiarism or malpractice? However, these questions refer more to the *modus operandi* of electronic sources and not to their facilities. Electronic sources enable interactive communication; they are flexible in including other media sources; promote publication democracy and they are friendly to environment since they contribute to the realization of a paperless society. Electronic learning offers the advantage of availability: enhance long distance education and training in isolated and remote areas; students with special needs are able to access academic material.

Thirdly, the issue of digital divide arises: to what extent future schools contribute negatively in the spread of digital divide between ICT rich and ICT poor? (Tavani 2007) On the one hand, ICT has the potential to offer universal access to information, regardless of distance, age, race, gender or other personal characteristics or people with disabilities (Grodzinsky 2000). However, are electronic sources available to everyone? In a world where less than 15% of users on the globe have access to the Internet the problem of digital divide becomes one of most important issues of the information society (Internet World Statistics (2001-2008). Are students that have access to ICT a new kind of digital elite? Yet, the problem of digital divide should not be only related to the inequality of access in ICT but also to the knowledge and ability required in order to use the technology (Campaine 2001), Tavani (2007). As Herman Tavani correctly maintained, the problem of digital divide is not that of single divide but of multiple divides such as the divide between nations and a divide within nations; he explains "the division between information rich and information poor nations is sometimes referred to as the "global digital divide"; the technological divides within nations, on the contrary, exist between rich and poor persons, racial majority and minority groups, men and women, disabled and non-disabled persons, and so forth" [Tavani (2007) p. 296]. From this perspective the problem of cost should be discussed; for example, while electronic learning offered in future schools gives the opportunity of a wider availability of education for every individual independent of location and time, the cost of hardware, software or course fees might place future schools beyond the reach of an average individual and parent. To what extent governments

and schools promote and support electronic learning? On the latter, a number of efforts have been made in order to overcome the problem of digital divide in education especially with reduced-cost technologies such as Mobile Learning (Brown 2005) and the 100\$ PC introduced by the MIT professor Nicholas Negropontes (Siddle 2005). Conversely, it has to be mentioned that the problem of digital divide exists at a secondary socio-economic level; at a primary level social and economic divisions and/or inequalities already existed prior the information revolution independent of the use or access to ICT [Tavani (2007) pp. 298-299]. Whereas the globalization of information diminishes the importance of geographic, political, economic and cultural boundaries, the world-wide projection of information technology blurs social inequalities, ignores specific human and psychological factors and becomes a source of cultural homogeneity and less multi-cultural diversity.

Fourthly, ICT in future schools brings the problems of dehumanization and alienation in terms of the lack of personal contact between the student and the teacher. Dehumanization is the process by which a person or group of persons comes to be regarded or treated as lacking in human qualities and social contact. When future schools overemphasize in the use of computer in-class and extreme computer workload, a possible consequence is the disassociation of students and teachers. The loss of the human factor and human interaction may cause psychological problems such as depression and anti-social behavior. Nevertheless, dehumanization is not an inherent characteristic of technology; modern life-style results dehumanization as a subsequent factor of the high-speed rhythms. Students become computer addicted where dehumanization is actually the human loss of freedom in personal life. On this ground, we face a communication paradox: while computers and especially computer networks enable people to communicate together in ways that are otherwise difficult or impossible, computers are also the reason for social isolation and anonymity. Hence a school of the future should not be based only on the improvement of ICT in-class but also on the progress of educational methods and techniques that enhance learning and particularly the creative interaction between teachers and students.

Thus the aforementioned concerns seem to outweigh some of the advantages of ICT in education, however, ICT remains an active element in the acquisition of knowledge in the information society and so the main concern is to improve its humanistic perspective. If education is a human right that values in itself then ICT in future school must be equally provided and developed. Unfortunately, a number of technophobic concerns lead to an unreasonable increase of security which further leads to an erosion of school's privacy. This issue is discussed in the next section of this paper.

Security and Privacy

ICT security measures such as CCTV cameras (Close Circuit Television), biometric identifications (i.e. hand and iris scanners), satellite surveillance with GPS technology (Global Positioning System), dataveillance techniques (i.e. black databases, data mining, data matching), mobile identification (3G technology) and identification cards (i.e. smart cards, laser cards, e-passports, and Radio Frequency IDs) are used today in public areas such as airports, banks, streets, and schools to track individuals in order to discourage criminals and terrorists. The most common way of ICT surveillance is that of CCTV cameras. It is noteworthy, in Britain, between 150 and 300 million pounds (225–450 million dollars) per year is now spent on a surveillance industry that involves an estimated 300,000 cameras and covers shopping areas, housing estates, car parks, and public facilities. According to BBC News, there are 25 million CCTV cameras in operation worldwide with some analysts predicting a tenfold increase in the next years (Wakefield 2002).

The tragic incident in Virginia Tech April 16, 2007 raised serious concerns about the appropriate safety procedures in schools and there is a social tendency that ICT surveillance tools will assist the improvement of security in schools and universities. Hence many schools monitor (in 24-hour basis) all the activities in vital campus areas such the entrance of the campus (i.e. to check unauthorised visitors), the car parking (i.e. to secure cars and monitor driving behaviours), the library (i.e. to avoid theft of books), the schoolyard (i.e. to monitor inappropriate behaviour and prevent accidents), the labs (i.e. to act in case of emergency) and the classrooms (i.e. to evaluate the lesson, teaching techniques and students' behaviour) (Bishop 2005), Henry 2003), Wakefield 2002).

However, the extensive application of ICT security in schools brings controversial reactions. On the one hand, supporters maintain that school surveillance improves security (i.e. school protects the students from unwanted visitors and unpredictable incidents) and the quality of learning (i.e. record classroom activities for absent students and observe the effectiveness of teaching procedures). On the other hand, opponents worry about the privacy both of students and teachers. Who controls students' and teachers' personal data? Are these personal data secured? Teachers and students feel anxious being monitored. This fact may affect negatively and not positively the quality of learning. There is also a fear that surveillance technologies in education may uncritically introduce students into an Orwellian culture where the right of privacy will be lost. In fact, the nature of privacy in modern schools changed as it is changed for modern societies. As Tavani correctly maintained, privacy in modern societies changed due to the impact of information technology on (1) the amount of personal information that can be gathered; (2) the speed at which personal information can be transmitted; (3) the duration of time that the

information can be retained and (4) the kind of information that can be transferred (Tavani 2007, pp. 128-130).

But why privacy is considered as valuable for human societies? Originally, privacy is regarded a human value consisting of four rights: (1) solitude: the right to be alone without disturbances; (2) anonymity: the right to have no personal identity in public; (3) intimacy: the right to do something privately; (4) reserve: the right to control personal information about oneself as well as the methods of dissemination of that information (Kizza 2002 p. 156). More specifically, privacy includes (1) territorial privacy (protects domestic, professional, civil, and recreational environments and territories); (2) location privacy (privacy of an individual's location); (3) bodily privacy (respect of an individual's body); (4) personal privacy (protects an individuals' personal identity); (5) communication privacy (protects individual's personal communication); (6) information privacy (determines privacy as the ability of individuals and groups to determine for themselves when, how and to what extent information about themselves is shared with others) (Westin 1967 p. 7). Furthermore, privacy consists of rules (including legislation and industry codes) concerned with the collection, compilation, and selective dissemination of personal data, such as credit information and medical and government records (Rosenberg 1997, pp. 279-280).

On this basis, the right of privacy is regarded as fundamental to human social and civil freedom. The British philosopher John Locke, in the *Second Treatise of Government* argued that liberty and privacy have to be undeniably protected and the German philosopher Emmanuel Kant argued that privacy is intrinsically good and invaluable for a free person to be autonomous and, therefore, to act responsibly (Edgar 1997, pp. 224-227). However, to what extent privacy has intrinsic value (it is desired for its own sake) or instrumental value (it is desired as a mean to other ends)? On this question different philosophers and analysts offer different answers (Tavani 2007, pp. 132-134): (1) privacy should not be necessarily regarded as a universal value of equal importance and significance for all cultures and societies (i.e. Western and Eastern perspectives) (Westin 1967, Tavani 2007, p. 133); (2) privacy is an instrumental value that serves the intercultural core value of security (Moor 1998, Tavani 2007, pp. 133-134); (3) privacy is an intrinsic value necessary to achieve important human ends such as trust and friendship (Fried 1997, Tavani 2007 p. 133); (4) privacy is a social value that promotes democracy and social goods (Regan 1995, Tavani 2007, p. 134).

At its extreme, some critics of personal privacy maintain that privacy in the form of anonymity protects the criminal, covers the guilty, camouflages deception and fraud and so it is harmful to social stability. Their position is: if you are not guilty, what do you have to hide? Thus the common belief is that there is a frag-

ile balance between public security and personal privacy. When public security measures are increased personal privacy and freedom is violated. When privacy becomes anonymity there is a threat to people's security and civil order. In other words, the security-privacy balance principle could be briefly stated as: more security brings less privacy and more privacy brings less security. However, whereas the former case is justified, the latter seems to be unjustifiable since it presupposes that all people are potentially guilty with malicious intentions for social stability. Privacy or even anonymity in its extreme form does not consequently bring security threat. The desire for privacy does not mean that the individual has something to hide. Privacy protection is a fundamental human right and its desire should not be justified nor negotiated. The problem of public security seems to have its roots to social inequalities and imbalances and not the privacy of the individuals.

Following Fried (1997) and Regan (1995) privacy in education should be considered as an intrinsic value that promotes democracy, equity and autonomy. A 'surveillance school' may dangerously encourage social heteronomy in which students will learn to attribute their right and decision to extra-social authorities such as political powers and governments. On the other hand, a 'privacy school' promotes autonomy in an autonomous and self-institute society (Gezerlis 2001) – in which members are aware and responsible of their own behaviour and actions. As Castoriadis maintained, people must be educated in autonomy and self-justice in order to be aware of their own values and rights (Castoriadis 2000, p. 142 ff.). Following Castoriadis, education in its wider context and all its forms is able to contribute in peoples' autonomy; people will be able to decide for themselves about their laws that actually define their own lives according to their pronouncements (Castoriadis 2000 p. 129). On this direction, ICT could be seen as a positive feature in schools and not an obstacle; students could be educated in a direct decentralised self-governed school democracy, in contrast to an indirect representative, centralised and authoritative school control. Computers and particular computer networks such as the Internet could be used as the electronic agora where the students like citizens of a modern polis exchange goods and ideas and become autonomous and self-institute. Internet could promote direct electronic democracy due to its public, globalize, non-hierarchical, uncontrolled and pluralistic nature (Johnson 2001). Based on this perspective, ICT could be used in schools not as a tool for inequality, dehumanisation, misinformation and cultural homogeneity, but as the kernel to educate a new form of direct democracy which supports and encourages the respect of human values such as autonomy, privacy and equity.

Conclusion

Information technology facilitates learning in future schools and universities. However, there are serious ethical questions arising from the extensive application of ICT in education. While computers improve the reliability, efficiency, security and speed of learning processes, they can also be used as instruments of control imposing a threat of privacy and freedom both of students and teachers. There is also the fear of dependence of students on computers that probably results the problems of deskilling, addiction and dehumanisation. Nevertheless, information technology introduces next generation into the borders of a new era where the evolution of humanity and social growth are inseparable from the technological advance and scientific development and, on this basis, a school of the future should be a realization of human progress and prospect. Future school and future technologies have the potential to nurture students through the values of privacy, equity and freedom in autonomy and self-institution. If democracy means 'rule by the people' in which all have equal power and participation, students of future schools could be educated as citizens of future societies where equity, freedom, participation and interaction is both possible and desirable.

REFERENCES

- Baase S.**, (2003) *A Gift of Fire: Social, legal, and ethical issues for computers and the Internet*, second edition, Prentice-Hall.
- Bishop, Z.**, (2005) "CCTV 'spy' comes into classrooms" in South London Press at http://www.footprintsecurity.com.au/z_art_050614_cctv_spy_comes_into_classrooms.php, 14/6/200 [last accessed 7/6/2008]
- Bottino, R. M.**, (2004) The evolution of ICT-based learning environments: which perspectives for the school of the future? in *British Journal of Educational Technology*, Volume 35 Issue 5 pp 553-567.
- Brent, D. and West-Burnham, J.**, (2003) *Handbook of Educational Leadership and Management*, Pearson Education.
- Brown, T.**, (2005) Towards a Model for M-Learning in Africa, *International Journal on E-Learning*, v4 n3 p. 299-315.
- Castoriadis, C.**, (2000) *The Talks in Greece*, 2nd ed., Athens: Ypsilon (in Greek)
- Cooper, J. M.**, (1997). *Plato: Complete Works* Indianapolis/Cambridge: Hackett.
- Doukakis, S., Chionidou-Moskofoglou, M. and Mangina-Phelan, E.** (2009) Students' (Under Graduate Primary Teachers) Learning Styles and their ICT & National Mathematics Software's Use.

Doukas, I. K., and Kotsanis, Y., (2008) Conference Proceedings: The School of Future and the Future of the School, International Conference of Education in Doukas School Athens.

Doukas (2009) "School of the Future" in Doukas School at <http://www.sof.gr/>

Eco, U. (1996) From Internet to Gutenberg in The Italian Academy for Advanced Studies in America. at <http://www.hf.ntnu.no/anv/Finnbo/tekster/Eco/Internet.htm>. [last accessed 3/4/2009]

Edgar, L. S., (2002) *Morality and Machines: Perspectives on Computer Ethics*, Jones and Bartlett.

Forester, T. and Morrison, P., (1994) *Computer Ethics: Cautionary Tales and Ethical Dilemmas in Computing*, Massachusetts: MIT Press.

Freud, S., (1961) Analysis terminable and interminable in J.Strachey (Ed.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 23, pp. 209–253). London: Hogarth Press. (Original work published 1937).

Fried, C., (1997) *Privacy: A Rational Context* in Ermann, M. D., Williams, M. B., and Gutierrez, C., (eds) *Computer, Ethics and Society* New York: Oxford University Press.

Gezerlis, A., (2001) 'Castoriadis and the Project of Autonomy. A Review of The Imaginary Institution of Society.' *Democracy & Nature: The International Journal of Inclusive Democracy*, 7:3 pp 469-88.

Grodzinsky, F. S, (2000) 'Equity of Access: Adaptive Technology' in *Science and Engineering Ethics*, vol. 6, n. 2.

Henry, J., (2003) CCTV in class 'will monitor bad behaviour in News.telegraph at <http://www.telegraph.co.uk/news/main.jhtml?xml=%2Fnews%2F2003%2F08%2F03%2Fncctv03.xml>, 03/08/2003 [last accessed 7/6/2008].

Internet World Statistics (2001 – 2008) = Internet World Statistics: Usage and Population Statistics (2001-2008) at <http://www.internetworldstats.com/stats.htm> [last accessed: 7/6/2008].

Johnson, D., (2001) *Computer Ethics*, 3rd ed, Prentice-Hall.

Microsoft (2009) = "School of the Future: Understand the vision" in microsoft.com at <http://www.microsoft.com/Education/SchoolofFutureVision.mspx>.

Moor, J. H., (1998) Reason, Relativity, and Responsibility in Computer Ethics *ACM SIGCAS Computers and Society*, vol. 28, issue 1 (March 1998).

Philadelphia District (2009) = School District of Philadelphia official website at <http://www.phila.k12.pa.us/>

Regan, P. M., (1995) *Legislating Privacy: Technology, Social Values, and Public Policy* Chapel Hill: The University of North Carolina Press.

Rosenberg, R., (1997) *The Social Impact of Computers*, 2nd edition, California: Academic Press.

Siddle, J., (2005) "Digital guru floats sub-\$100 PC" in BBC News at <http://news.bbc.co.uk/2/hi/technology/4243733.stm> [last access: 7/6/2008].

Spinello R. and Tavani H., (2001) *Readings in Cyberethics*, Jones and Bartlett.

Stamatellos, G. (2007) *Computer Ethics: A Global Perspective*, Jones and Bartlett.

Tavani, T. H., (2007) *Ethics and Technology: Ethical Issues in an Age of Information and Communication Technology*, 2nd ed, John Wiley and Sons.

Wakefield, J., (2002) 'Watching your every move' in BBC News at <http://news.bbc.co.uk/1/hi/sci/tech/1789157.stm>, Thursday, 7 February, 2002 [last accessed 7/6/2008].

Westin A. (1967) *Privacy and Freedom*, New York: Atheneum Press.

The role of Internet Service Providers. Ethics, Reality and the Law: The Example of Promusicae v. Telefonica

Irini Stamatoudi*
LL.M. Ph. D

Abstract

The role of internet service providers is highly contested under community and national legal orders. This is essentially so by reason of the clash between various fundamentals rights and in particular between the right to privacy (data protection and secrecy of communications) and the right to property (which includes copyright). In other words, a balancing of rights is required which will also provide answers as to the role of Internet service providers. This article examines the impact of ethics, reality and the law in this balancing of rights taking as its starting point a recent judgment by the European Court of Justice (*Promusicae v. Telefonica*). By reason of the word constraints the law which is examined in this respect is community law.

Keywords: copyright, data protection, ethics, file sharing, intellectual property, Internet, Internet service providers, law, *Promusicae v. Telefonica*, secrecy of communications

* *Irini Stamatoudi* is the General Director of the Greek National Copyright Organization. She is teaching on the WIPO - ILO - University of Turin Masters in Intellectual Property, on the WIPO Academy specialisation courses on copyright and e-commerce and she was a lecturer in Law at the Faculty of Law of the University of Leicester in England. She holds degrees from the University of Athens in Greece (Ptychio Nomikis) and from the University of Leicester (LL.M and Ph.D). She has published six books (I. Stamatoudi and P. Torremans (eds), *Copyright in the New Digital Environment: The Need to Redesign Copyright*, Sweet & Maxwell, London, 2000; I. Stamatoudi, *Multimedia products as copyright works*, Cambridge University Press, Cambridge, 2002; I. Stamatoudi, *Community Competition Law and Intellectual Property*, Nomiki Vivliothiki, Athens, 2006 (in Greek); I. Stamatoudi, *Copyright Law. National Laws, Community Laws and International Conventions*, Nomiki Vivliothiki, Athens, 2007 (in Greek); L. Kotsiris – I. Stamatoudi, *Commentary to the Greek Copyright Act*, P. N. Sakkoulas, 2009 (in Greek); I. Stamatoudi (ed.), *Journalists and Mass Media Publishers. Copyright Law Issues*, P. N. Sakkoulas, 2009 (in Greek)) and numerous articles in law journals in Greece, UK, Germany, Switzerland, the Netherlands, the US, Poland and Turkey.

Introduction

This paper aims to explore to what extent ethics interact with reality and the law in cases where internet service providers (ISPs) host or transmit on line copyright material without the authorization or consent of the rightholders. It also aims to explore to what extent the law as it stands now incorporates sufficiently both reality and moral considerations. A recent judgment of the European Court of Justice is taken as an example of how the Community itself (by its Court) approached the subject from a legal point of view. This view, however, encompasses moral considerations, too, which are dictated by the clash between various fundamental rights in the case under consideration, such as privacy, data protection and copyright protection. Should we draw a line and where this line is to be drawn? How can we balance the interests of rightholders and those of the public/users of the Internet? What if we don't balance these interests, will this system still be workable? How far enforcement should go? Can we control it? Are some fundamental rights more fundamental than others and who shall attempt a weighing of interests?

Chapter I examines the role of Internet service providers as intermediaries in the online distribution of copyright material. The role of internet service providers differs from case to case according to whether they themselves host the material, act as intermediaries for the transmission of material coming from elsewhere or when their services are used as facilitator for peer to peer exchanges. The role of ISPs in these cases is determined in the minds of both legal advisors and the public on the basis of whether ISPs have knowledge or have gained knowledge of the distribution/communication of material through their net for which there is no authorization or consent by the rightholder. Other notions which are necessary for the discussion are also explored, such as IP addresses and file sharing.

Chapter II examines the notion of ethics/applied ethics through particular examples in order to bring it in terms with what we want to achieve/prove in this essay. Also the interaction between the emerging and fast evolving reality in conjunction with the law and the ethics is explored. We try to prove that ethics become more and more relevant in our days because laws are unworkable, conflicting and there is absence of appropriate rule-makers for cyberspace. Applied ethics can be used as a tool to determine public policy and expectations in an era where the evolution of technology is inextricably linked with an evolution of morals. Right and wrong do not present general abstract notions but they need to be applied to everyday realities and respond to technological necessities. The notion of ethics is explored in conjunction with the role of ISPs.

Chapter III examines in detail the recent judgment of the European Court of Justice in *Promusicae v. Telefonica* (C-275/06). This judgment explores in depth a

number of fundamental rights and how these rights interact (or even conflict) between themselves when it comes to the communication of personal data by a Spanish ISP (Telefonica) to the rightholders (in the case at issue to a non-profit-making organization representing Spanish music producers (Promusicae)) in order for the latter to ensure effective protection of copyright in the context of civil proceedings. Although this judgment leaves at the end in many respects the issue under consideration open, in the sense it only clearly provides that Community law in the area does not require ISPs to reveal personal data to rightholders concerning illegal distribution of their works using their connection in order to invoke their rights in the context of civil proceedings (leaving open at the same time the issue of penal proceedings as well as the possibility for Member States to expressly provide for such a provision in their national laws), it is a rather interesting case because it provides various arguments as to a weighing of rights (such as the right to privacy, personal data and copyright protection). These arguments incorporate in many respects moral and ethical considerations and try to accommodate them in the existing Community legal context.

Chapter IV attempts (on the basis of *Promusicae v. Telefonica*) an ethical approach to the role of ISPs. It explores the extent to which moral considerations were taken into account in the reasoning of the *Promusicae* case and what would be ethical in the case at issue. It explores the extent to which Greek law provides for answers in this area, what the current situation is, what the problems are, if there is a way to solve these problems, national, Community and international initiatives in the area, where is our world heading to, what the realities are and how up to date and conscious to the real need of the Internet can we be. It also explores the existing and changing business image of ISPs, their role in terms of law as well as their role in the market.

Lastly, conclusions are drawn as to what extent ethics are law driven or law is (or should be) ethics driven. In all questions of ethics it is also vital to ascertain whose morals is one taking into account and to what extent -in a market orientated society- morals are defined by the interests at stake.

Internet service providers, IP addresses and file sharing

Internet service providers (ISPs) are telecommunication companies, which offer their customers access to the Internet. ISPs can themselves host content, which is protected by copyright, act as intermediaries for the transmission of material coming from elsewhere or they can offer their services as a facilitator for peer to peer exchanges. It is especially in the first case that one could allege that ISPs should have certain control as to what is published and transmitted on the sites they host. In the other cases their role is more neutral, the possibility for control

is much more restricted. In any case one could allege that unless they gain knowledge by chance (unless they police the Internet) they can only gain knowledge of infringement when this infringement is notified to them.

The connection and communication of private computers on the Internet is done on the basis of IP addresses. IP addresses are numerical address formats, comparable to a telephone number, which enable networked devices such as web-servers, e-mail servers or private computers to communicate with one another on the Internet. IP addresses can either be static or dynamic. Static IP addresses are assigned in order to connect private users to the Internet, in similar fashion to connection to the telephone network. However, that is rather rare, since the Internet is at present still organised in such a way that each access provider has only a limited number of addresses available to it. For this reason subscribers are assigned dynamic addresses. Dynamic IP addresses are assigned to customers on an *ad hoc* basis from the limited number of addresses the ISP has. Therefore these addresses change each time a customer dials up.

The services of ISPs are also used for file sharing. File sharing is a form of exchange of files containing, for example, pieces of music or films. Users first copy the files onto their computers and then offer them to anyone who is in contact with them via the Internet and a particular program. Some collecting societies possess the technology which allows them to identify a number of IP addresses which were used at certain times for the purpose of 'file sharing' in respect of material whose exploitation rights they hold. On the basis of this information they ask ISPs to match the information they possess with the one they retain in their files after the connection has ended. In particular they ask them to find out which connection was used in each case and provide them with the details concerning to whom and when they assigned a particular IP address.

It is beyond doubt that the details of the persons that have used a particular IP address at a particular time of a particular day constitute personal data within the meaning of art. 2(a) of Directive 95/46. And that is so because this information is capable of linking an action to a subscriber. What the judgment (*Promusicae v. Telefonica*) has not confirmed is whether an IP address as such can be considered personal data. However, the theory is divided on this point.

Ethics and the role of the Internet Service Providers

Applied ethics can be used as a tool in particular situations - such as those which relate to the role of ISPs or the disclosure of personal data by them in specific cases - in order to determine public policy and expectations or else what the average person considers fair in the case at issue. Fairness and morality should not be confused with what one chooses as fair or moral because this serves his own

interests. Public policy is not determined by the answers given, for example, in order to be able to infringe the works of others on the Internet and not to be caught. There is an argument that 'ethical' is something which goes beyond morality, whilst morality is only one aspect of it. Ethics interact with law in two instances, in particular: a) when law is not drafted in such a way as to leave no discretion whatsoever to those who apply it as to its content and meaning; b) It also interacts in those situations where conflicting rules exist and a weighing of interests needs to take place. In these cases one has to ask oneself, what is ethically correct to apply in the situation at issue, how my decision will affect other situations, which may have social or other repercussions, or simpler what is right and what is wrong. Yet, this assessment is not always easy to make.

Prof. Hart in 1958 in one of his most reputed essays published in the Harvard Law Review put forward the following question: how do we understand a legal rule which forbids one to take a vehicle into a public park? This rule plainly forbids an automobile. But does it also forbid airplanes, a bicycle, roller skates and a toy automobile? This question was set to demonstrate that "there must be a core of settled meaning, but there will be, as well, a penumbra of debatable cases in which words are neither obviously applicable nor obviously ruled out Hart, 607". It is in these instances (debatable cases) that one has a choice. One has a choice according to what one understands by this rule looking at the same time in the interpretation given to other similar rules (interpretation by analogy), a choice according to what the law drafter wanted to achieve or even better the interpreter believes that it ought to be achieved in the case at issue (teleological interpretation) or even a choice on the basis of how the rule should have looked like (interpretation on the basis of natural law). As Hart rightfully notes, «the existing law only imposes limits on our choice and not the choice itself Hart, 611». And it is at this point that (amongst others) there is also an intersection between law and morals.

Referring to judges, Hart mentions that «[the judge] either takes the meaning that the word most obviously suggests in its ordinary non-legal context to ordinary men, or one which the word has been given in some other legal context, or, still worse, he thinks of a standard case and then arbitrarily identifies certain features in it – for example, in the case of a vehicle, (i) normally used on land, (2) capable of carrying a human person, (3) capable of being self-propelled – and treats these three as always necessary and always sufficient conditions for the use in all contexts of the word 'vehicle', irrespective of the social consequences of giving it its interpretation. This choice, not 'logic', would force the judge to include a toy motor car (if electrically propelled) and to exclude bicycles and the airplane. In all this there is possibly great stupidity but no more 'logic', and no less, then in cases in which the interpretation given to a general term and the consequent applica-

tion of some general rule to a particular case is consciously controlled by some identified social aim hart, 611».

Another example derived from Greek mythology, which is illustrative of the conflict/clash between two fundamental laws is shown in *Antigone* of Sophocles. I remind you that Antigone's brother (Polynices) was killed during an attempt to seize the city of Thebes against his brother (Eteocles). The King of Thebes (Creon) ordered that her brother -who he considered an enemy- should not be buried. However, Antigone decided to bury her brother because if she had followed the law of the 'State' she would have disobeyed the law of the Gods (natural law).

If we apply the aforementioned tracks of thought to the role of internet service providers, one could easily argue that national provisions (such as the Greek one) referring to intermediaries' liability could or could not apply to internet service providers according to how one understands the wording of the law and whether it is considered that this law conflicts with the ones of data protection and secrecy of communications. Article 64A (Injunction) provides that "rightsholders may apply for an injunction against intermediaries whose services are used by a third party to infringe a copyright or related right (article 8, paragraph 3 of Directive 2001/29)".

Taking the Greek provision literally trying to investigate its real meaning one should ask the following questions: Are internet service providers intermediaries? Are their services used by a third party for the infringement of copyright and related rights? Would this course of conduct contradict other provisions of law? What would ethically be the right decision to take?

A recent judgment of the European Court of Justice (*Promusicae v. Telefonica* C-275/06) has shed light on some of these issues examining the whole scope of Community law in the area. This case has reached the European Court of Justice in the form of a preliminary ruling and was reaffirmed on work points by LSGv. Tele 2 (C-557/07).

The example of *Promusicae v. Telefonica*

Historical Background

Promusicae is a Spanish non-profit-making organisation of producers and publishers of musical and audiovisual recordings. Promusicae applied for preliminary measures to the Madrid Commercial Court No 5 against Telefónica, a Spanish internet service provider, in order for the latter to disclose the identities and physical addresses of certain persons whom it provided with internet access services and whose IP address and date and time of connection were known to Promusicae. According to Promusicae, these persons were engaging in peer to

peer file sharing practices (using the KaZaA file exchange program) and provided access in shared files of personal computers to phonograms in which the members of Promusicae held the exploitation rights. Promusicae asked this information in order to be able to bring civil proceedings against the persons concerned for intellectual property infringement as well as engagement in unfair competition. The Court ordered the preliminary measures requested by Promusicae. However, Telefónica appealed against that order on the grounds that Spanish law authorised the communication of such data only in a criminal investigation or for the purpose of safeguarding public security and national defence and not in civil proceedings or as a preliminary measure relating to civil proceedings. Promusicae argued that Spanish law (i.e. Article 12 of the LSSI) must be interpreted in the light of Directives 2000/31, 2001/29 and 2004/48 and with Articles 17(2) and 47 of the Charter, according to which the provision of such data should not be limited solely to the purposes aforementioned. Accordingly, the Spanish Court decided to stay the proceedings and refer to the European Court of Justice the following question: "Does Community law, specifically Articles 15(2) and 18 of Directive 2000/31, Article 8(1) and (2) of Directive 2001/29, Article 8 of Directive 2004/48 and Articles 17(2) and 47 of the Charter ... permit Member States to limit to the context of a criminal investigation or to safeguard public security and national defence, thus excluding civil proceedings, the duty of operators of electronic communications networks and services, providers of access to telecommunications networks and providers of data storage services to retain and make available connection and traffic data generated by the communications established during the supply of an information society service?"

In other words, the European Court of Justice was asked whether Member States are under an obligation to provide for a communication of personal data by an internet service provider concerning its clients with regard to an exchange of files of protected material and therefore a copyright infringement (and potentially an unfair competition law infringement) in order for this data to be used in the course of civil proceedings.

Relevant legal provisions

In order for one to be able to draw conclusions, one has to consider all those legal provisions that are relevant to such an issue and which bind the parties. And the provisions that are relevant here are not only the ones referring to copyright but also those provisions referring to data protection and secrecy of communications. In other words, there seems to be a clash of rights, both fundamental rights enshrined in the Charter and European Convention for Human Rights, i.e. a clash between the right to property which includes intellectual property and the right to privacy which includes data protection and secrecy of communications. On top of everything the Spanish Court makes also reference to the right to an effective remedy, which is also consid-

ered a fundamental right but whose exercise depends to a large extent from the outcome of the weighing of the two aforementioned rights.

Relevant community law provisions relating to the information society and copyright are the following:

Directive 2000/31

Article 1 of the Directive provides that “1. This Directive seeks to contribute to the proper functioning of the internal market by ensuring the free movement of information society services between the Member States. 2. This Directive approximates, to the extent necessary for the achievement of the objective set out in paragraph 1, certain national provisions on information society services relating to the internal market, the establishment of service providers, commercial communications, electronic contracts, the liability of intermediaries, codes of conduct, out-of-court dispute settlements, court actions and cooperation between Member States. 3. This Directive complements Community law applicable to information society services without prejudice to the level of protection for, in particular, public health and consumer interests, as established by Community acts and national legislation implementing them in so far as this does not restrict the freedom to provide information society services. ... 5. This Directive shall not apply to: ... (b) questions relating to information society services covered by Directives 95/46/EC and 97/66/EC”.

According to Article 15 of the Directive “1. Member States shall not impose a general obligation on providers, when providing the services covered by Articles 12, 13 and 14, to monitor the information which they transmit or store, nor a general obligation actively to seek facts or circumstances indicating illegal activity. 2. Member States may establish obligations for information society service providers promptly to inform the competent public authorities of alleged illegal activities undertaken or information provided by recipients of their service or obligations to communicate to the competent authorities, at their request, information enabling the identification of recipients of their service with whom they have storage agreements”.

Article 18 of the Directive provides that “1. Member States shall ensure that court actions available under national law concerning information society services’ activities allow for the rapid adoption of measures, including interim measures, designed to terminate any alleged infringement and to prevent any further impairment of the interests involved....”.

Directive 2001/29

According to Article 1(1) of the Directive, the Directive concerns the legal protection of copyright and related rights in the framework of the internal market, with particular emphasis on the information society.

Article 8 of the Directive provides that “1. Member States shall provide appropriate sanctions and remedies in respect of infringements of the rights and obligations set out in this Directive and shall take all the measures necessary to ensure that those sanctions and remedies are applied. The sanctions thus provided for shall be effective, proportionate and dissuasive. 2. Each Member State shall take the measures necessary to ensure that rightholders whose interests are affected by an infringing activity carried out on its territory can bring an action for damages and/or apply for an injunction and, where appropriate, for the seizure of infringing material as well as of devices, products or components referred to in Article 6(2). 3. Member States shall ensure that rightholders are in a position to apply for an injunction against intermediaries whose services are used by a third party to infringe a copyright or related right”.

Article 9 of the Directive provides that “This Directive shall be without prejudice to provisions concerning in particular patent rights, trade marks, design rights, utility models, topographies of semi-conductor products, typefaces, conditional access, access to cable of broadcasting services, protection of national treasures, legal deposit requirements, laws on restrictive practices and unfair competition, trade secrets, security, confidentiality, data protection and privacy, access to public documents, the law of contract”.

Directive 2004/48

Article 1 of the Directive provides that “This Directive concerns the measures, procedures and remedies necessary to ensure the enforcement of intellectual property rights ...”.

According to Article 2(3) “This Directive shall not affect: (a) the Community provisions governing the substantive law on intellectual property, Directive 95/46/EC, Directive 1999/93/EC or Directive 2000/31/EC, in general, and Articles 12 to 15 of Directive 2000/31/EC in particular; (b) Member States’ international obligations and notably the TRIPS Agreement, including those relating to criminal procedures and penalties; (c) any national provisions in Member States relating to criminal procedures or penalties in respect of infringement of intellectual property rights”.

Article 3 of the Directive provides that “1. Member States shall provide for the measures, procedures and remedies necessary to ensure the enforcement of the intellectual property rights covered by this Directive. Those measures, procedures

and remedies shall be fair and equitable and shall not be unnecessarily complicated or costly, or entail unreasonable time-limits or unwarranted delays. 2. Those measures, procedures and remedies shall also be effective, proportionate and dissuasive and shall be applied in such a manner as to avoid the creation of barriers to legitimate trade and to provide for safeguards against their abuse”.

Article 8 of the Directive provides that “1. Member States shall ensure that, in the context of proceedings concerning an infringement of an intellectual property right and in response to a justified and proportionate request of the claimant, the competent judicial authorities may order that information on the origin and distribution networks of the goods or services which infringe an intellectual property right be provided by the infringer and/or any other person who: (a) was found in possession of the infringing goods on a commercial scale; (b) was found to be using the infringing services on a commercial scale; (c) was found to be providing on a commercial scale services used in infringing activities; or (d) was indicated by the person referred to in point (a), (b) or (c) as being involved in the production, manufacture or distribution of the goods or the provision of the services. 2. The information referred to in paragraph 1 shall, as appropriate, comprise: (a) the names and addresses of the producers, manufacturers, distributors, suppliers and other previous holders of the goods or services, as well as the intended wholesalers and retailers; (b) information on the quantities produced, manufactured, delivered, received or ordered, as well as the price obtained for the goods or services in question. 3. Paragraphs 1 and 2 shall apply without prejudice to other statutory provisions which: (a) grant the rightholder rights to receive fuller information; (b) govern the use in civil or criminal proceedings of the information communicated pursuant to this Article; (c) govern responsibility for misuse of the right of information; or (d) afford an opportunity for refusing to provide information which would force the person referred to in paragraph 1 to admit to his/her own participation or that of his/her close relatives in an infringement of an intellectual property right; or (e) govern the protection of confidentiality of information sources or the processing of personal data”.

Community legal provisions regarding the protection of personal data are the following:

Directive 95/46/EC

Article 2 of Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data (OJ 1995 L 281, p. 31) provides that “For the purposes of this Directive: (a) “personal data” shall mean any information relating to an identified or identifiable natural person (“data subject”); an identifiable person is one who can be identified, directly or

indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity; (b) “processing of personal data” (“processing”) shall mean any operation or set of operations which is performed upon personal data, whether or not by automatic means, such as collection, recording, organisation, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, blocking, erasure or destruction”.

According to Article 3 of the Directive “1. This Directive shall apply to the processing of personal data wholly or partly by automatic means, and to the processing otherwise than by automatic means of personal data which form part of a filing system or are intended to form part of a filing system”.

Article 7 of the Directive provides that “Member States shall provide that personal data may be processed only if: ... (f) processing is necessary for the purposes of the legitimate interests pursued by the controller or by the third party or parties to whom the data are disclosed, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection under Article 1(1)”.

Article 8 of the Directive provides that “1. Member States shall prohibit the processing of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and the processing of data concerning health or sex life. 2. Paragraph 1 shall not apply where:... (c) processing is necessary to protect the vital interests of the data subject or of another person where the data subject is physically or legally incapable of giving his consent ...”.

According to Article 13 “1. Member States may adopt legislative measures to restrict the scope of the obligations and rights provided for in Articles 6(1), 10, 11(1), 12 and 21 when such a restriction constitutes a necessary measure to safeguard: (a) national security; (b) defence; (c) public security; (d) the prevention, investigation, detection and prosecution of criminal offences, or of breaches of ethics for regulated professions; (e) an important economic or financial interest of a Member State or of the European Union, including monetary, budgetary and taxation matters; (f) a monitoring, inspection or regulatory function connected, even occasionally, with the exercise of official authority in cases referred to in (c), (d) and (e); (g) the protection of the data subject or of the rights and freedoms of others”.

Directive 2002/58/EC

According to Directive 2002/58 (art. 5(1)) Member States must ensure the confidentiality of communications by means of a public communications network and publicly available electronic communications services, and of the related traffic data. The same Directive provides for the cases where the processing of traffic data is allowed (i.e. for billing and marketing services and the provision of value added services). It does not concern the communication of that data to persons other than those acting under the authority of the providers of public communications networks and publicly available electronic communications services. In other words, they do not relate to disputes other than those between suppliers and users and therefore it does not also concern situations such as in the case at issue. However, Member States may provide for exceptions to the principle of confidentiality of personal data in cases of national security, defence and public security and the prosecution of criminal offences or of unauthorised use of the electronic communications system (as referred to in Article 13(1) of Directive 95/46) as long as these exceptions/restrictions constitute a necessary, appropriate and proportionate measure within a democratic society. All these cases referred to above do not concern situations that give rise to civil proceedings. However, the latter one, which makes reference to art. 13 (1) of Directive 95/46, allows Member States to take measures to restrict the obligation of confidentiality of personal data where this is necessary for the protection of the rights and freedoms of others. As the Court admitted, the right to property, and in particular to intellectual property, should be included within the rights referred to in the Directive. In other words, Member States are not prevented (by Directive 2002/58) from providing for an obligation to disclose personal data in the context of civil proceedings. They are, however, under no duty to do so (in other words, art. 15 (1) does not impose any such duty).

Conclusions

Although the three Directives (art. 1(5)(b) of Directive 2000/31, art. 9 of Directive 2001/29 and art. 8(3)(e) of Directive 2004/48) put forward by the Spanish Court have amongst others as their aim the effective protection of copyright, this protection does not supersede the protection afforded to personal data by the relevant Directives regulating its requirements. As explained above there are provisions in these Directives (for the protection of personal data) which provide that in cases of infringement of an intellectual property right and in response to a justified and proportionate request of the claimant, the competent judicial authorities may order that information on the origin and distribution networks of the goods or services which infringe an intellectual property right (art. 8(1) 2004/48). However, if these provisions are seen in conjunction with art. 8(3)(e)

2004/48 and art. 15(2) and 18 of Directive 2000/31 or that of art. 8(1) and (2) of Directive 2001/29 one comes to the conclusion that Member States are under no obligation to provide for the communication of personal data in the context of civil proceedings. The ECJ took this issue even further. It came to the conclusion (para 60 of the ECJ's Judgment) that although articles 41, 42 and 47 of the TRIPs Agreement (in the light of which Community law must be interpreted) require the effective protection of intellectual property rights, they do not oblige Member States to provide for an obligation to communicate personal data in the context of civil proceedings.

In the case at issue it seems to be a conflict between various fundamental rights: the right to property, which also includes intellectual property (art. 17 of the Charter) and the right to an effective remedy (art. 47 of the Charter) on the one hand, and the right to respect for private life (art. 7 of the Charter) and the right to protection of personal data (art. 8 of the Charter) on the other hand. It seems that the conflict created here could only be solved by a fair balancing of the rights involved in the present situation. According to the ECJ (para 68seq. of the ECJ's Judgment) the mechanism is contained first, in the relevant Directives which to a certain extent specify the procedures and the circumstances at issue and second, to the extent some of their provisions are general and leave discretion to their Member States, they result from the adoption by the Member States of national provisions transposing those directives in their national laws in conformity with other fundamental rights and general principles of Community law, such as the principle of proportionality.

In the light of the above the ECJ came to the conclusion that "Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market ('Directive on electronic commerce'), Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society, Directive 2004/48/EC of the European Parliament and of the Council of 29 April 2004 on the enforcement of intellectual property rights, and Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications) do not require the Member States to lay down, in a situation such as that in the main proceedings, an obligation to communicate personal data in order to ensure effective protection of copyright in the context of civil proceedings. However, Community law requires that, when transposing those directives, the Member States take care to rely on an interpretation of them which allows a fair balance to be struck between the various fundamental rights protected by the Community

legal order. Further, when implementing the measures transposing those directives, the authorities and courts of the Member States must not only interpret their national law in a manner consistent with those directives but also make sure that they do not rely on an interpretation of them which would be in conflict with those fundamental rights or with the other general principles of Community law, such as the principle of proportionality”.

An ‘ethical’ approach to the role of ISPs

From the above, and especially from the recent European Court of Justice’s judgment, it became clear that the protection of copyright constitutes and equally fundamental right as the one of the protection of personal data. This in practice leaves Member States with a choice: they have to choose which of the two rights they give priority to when it comes to providing information (personal data) to collective management societies in order for them to defend their rights with regard to the works they administer in the course of civil proceedings. This choice is not entirely open. Any balancing of rights has to take into account general principles of Community law and in particular the principle of proportionality as well as the margin of appreciation doctrine. That means that an action of the Community shall not go beyond what is necessary to achieve the objectives of the Treaty. If that is transposed into national law it means that a public authority may not impose obligations on a citizen except to the extent to which they are strictly necessary in the public interest to attain the purpose of the measure.

The principle of proportionality underlines the Member States’ discretion on the one hand and on the other hand allows ethics to enter the image. Member States on the basis of the margin of appreciation doctrine can make their own judgment as to what sits well with their socio-legal and economic tradition.

In this light the primary question: should ISPs be obliged to provide the details of the persons using their services (or a particular IP address) at a particular date to collective management societies, in order for them to pursue their rights to the civil courts against those persons? hides in it another more important question: do we want to make possible for rightholders to pursue their rights effectively on the Internet and does the provision of this information constitute the appropriate tool to do so?

Many arguments have been raised so far for and against any positive answer to this question. It has been argued that providing this information does not always lead you to the actual infringer but solely to the subscriber, which in many cases may not be the infringer. It has also been argued that such a provision would open the floodgates for a series of suits against youngsters in whose minds file sharing is a legitimate activity of exchanging music and films. It will also allow

rightholders to use this ability to spy on subscriber's choices and actions. It is natural that all this create fears of uncontrolled activities which may jeopardize one's right to one's private life and secrecy of communications.

Yet, one asks oneself: what is the point of being given or acknowledged a property right, which is exclusive and absolute, if this right is not in practice enforceable. On top of it, why a state cannot provide for such a mechanism that will allow others to disengage from infringing practices well before the rightholder is given the right to turn to the civil courts or allow the availability of this information once the practices involved can be characterized as crimes. Is Internet different in this respect than the real world is? And, if yes, in what respect? Does the fact that downloading and exchanging files can be done at an enormous speed, easily and without any cost, legalizes infringing activities? If we believe that a law should not be applied the solution is not to prevent it from applying but to change it. That means that the law has lost its ethical justification and no longer corresponds to public morals. Is this the case at issue here?

Do ethical considerations require one-sided answers? Is right and wrong two notions which are watershed separated such a black and white or something may also be grey. In other words can we protect private life and intellectual property rights or are they contradicting notions? The ECJ in the case at issue left the answer open to Member States. Its conclusions -if regarded from an ethical point of view- could be summarized as follows. Ethics may differ from country to country. Therefore states should look into their current prevailing ethics and decide for themselves what they consider ethical. In legal terms they need to attempt their own weighing of interests and apply the principle of proportionality on the basis of what they consider proportional which of course should not manifestly impinge on other peoples' ethics. However, since the Court did not reach a particular stance on the matter it would be difficult for any country's decision to be found disproportionate.

Conclusions: Are ethics law-driven or is law ethics-driven?

The reality of the Internet is undoubtedly one which evolves in a tremendous speed and before any actual regulation for specific problems or situations that may incur. Therefore the example of the vehicle that is forbidden into the public park and that of Antigone are probably examples that we will come across in different facets regarding the Internet. It is more likely of course to come across the first example rather than the second just because most fundamental laws in modern society have found their way into the European Convention of Human Rights.

It could be argued that ethics preexist. However, ethics may have been formed by existing laws in similar cases. As Plato alleged "one's interpretation of law

will necessarily influence the interpretation of ethics, and vice versa". In any case new laws should be drafted with ethics in mind: with what we want to achieve in a fair, democratic and balanced society. The drafter cannot bear all situations in mind. The unregulated ones however should mirror the aforementioned principles as to how society ought to be. This approach of course is highly subjective and is continuously defined by new evolving ethics. But there is some minimum which is acceptable to the majority of a society and this constitutes the ethics of our times in this particular society.

Laws regulating the Internet present one more problem in this respect. They cannot mirror the ethics of particular societies. They set by definition globalised problems.

The role of ISPs is a small part of all this discussion. And perhaps our era is an era which should avoid one-sided solutions. It is therefore difficult to accept that laws or legal mechanisms cannot be designed in such a manner which can accommodate the worries of both sides: of those who care for privacy and those who care for their intellectual property rights.

REFERENCES

Akdeniz, Y., Walke, C. and Wall, D., (2000). *The Internet, Law and Society*, Longman, Dorchester.

Arai, Y., (2002). *The Margin of Appreciation Doctrine and the Principle of Proportionality in the ECHR*, Intersentia, Antwerp.

Bowrey, K., (1998). *Ethical Boundaries and Internet Cultures*, in Bentley, L. and Maniatis, S. (eds.), *Intellectual Property and Ethics in Perspectives on Intellectual Property*, Sweet & Maxwell, London, 36.

Dworkin, R., (1986). *Law's Empire*, Harvard University Press, Harvard.

Fragkouli, A., (2008). *Are IP Addresses personal data and what are the consequences?* (in Greek), DIMEE, 198.

Fuller, L., (1958). *Positivism and Fidelity to Law: A Reply to Professor Hart*, *Harvard Law Review*, 630.

Hamelink, C. J., (2000). *The Ethics of Cyberspace*, Digitaal Fatsoen, Holland.

Hart, H. L., (1958). *Positivism and the separation of Law and Morals*, *Harvard Law Review*, 593.

Hugenholtz, B., (2000) (ed.). *Copyright and Electronic Commerce. Legal Aspects of Electronic Copyright Management in Information Law Series*, Kluwer Law International, The Hague, London, Boston.

Kitchin, H., (2008), *Research Ethics and the Internet: Negotiating Canada's Tri-Council Policy Statement*, Fernwood Publishing Co Ltd.

Langford, D., (2000). *Internet Ethics*, St Martin's Press, USA.

Posner, R., (1999). *The Problematics of Moral and Legal Theory*, Harvard University Press, Harvard.

Shany, Y., (2005). Toward a General Margin of Appreciation Doctrine in International Law? Vol. 16 *European Journal of International Law*, 907.

Strowel, A. and Ide, N., (2000/A). Liability of Internet Intermediaries: Recent Developments and the Question of Hyperlinks", *Revue Internationale du Droit d'Auteur*, vol. 186.

Strowel, A. and Ide, N., (2000/B). Liability for linking to illegal content, *Computer und Recht International*, 91.

Strowel, A. and Ide, N., (2001). Liability with Regard to Hyperlinks, *Colum.-VLA J.L. & Arts*, vol. 24, 403.

Strowel, A., (1998). Liaisons dangereuses et bonnes relations sur Internet: à propos des hyperliens, *A&M*, 4.

Strowel, A., (1999). Introduction to intellectual Property Issues in Relation to Electronic Commerce, in Kaestner, J. (ed.), *Legal Aspects of Intellectual Property Rights in Electronic Commerce*, Munich, 6.

Strowel, A., (2003). Security and Privacy on the Internet, WIPO-ESCWA Arab Regional Conference on Intellectual Property and Electronic Commerce, Beirut (7-8 May 2003).

Strowel, A., (2009) (ed.). *Peer-to-peer File Sharing and Secondary Liability in Copyright Law*, Edward Elgar Publishing, cheltenham.

Vedder, A., (2001). *Ethics and the Internet*, Intersentia Publishers, Antwerp.

Wüstenberg, D., (2006). Argumente gegen die Rechtmässigkeit der Vorratsdatenspeicherung, *RDV*, 102.

Virtually Good – What Can We Learn from the Argument from False Pleasures?

Johnny Hartz Søraker

Department of Philosophy, University of Twente

Abstract

Virtual worlds offer a way to substitute things that contribute to our well-being with virtual “surrogates”. For instance, many find it easier to pursue virtual surrogates for friendship, love, community, aesthetic beauty and so forth. This could be interpreted as something positive, since virtual worlds seemingly offer new opportunities for attaining whatever it is that provide us with well-being. However, a number of arguments have been put forward to the effect that these kinds of surrogates are somehow false or inauthentic, and that if we replace the real thing with virtual surrogates we thereby reduce the quality of our lives. In this paper, I will discuss a particular type of argument in this vein, which I will refer to as ‘the argument from false pleasures’. I will argue that the argument suffers from some serious problems, but that it nonetheless sheds light on a number of important features of virtuality. In particular, I will argue that sometimes there is no significant difference between actual and virtual and that sometimes this difference is irrelevant. I will also argue that we ought to keep a strict distinction between consideration of self-interest and other-interest, and that we ought to relate any claim about the value of virtual surrogates to individuals’ particular and contingent life conditions

Keywords: Virtual Worlds, Well-being, Epistemology, Ontology, Objectivity, the Experience machine

* *Johnny Hartz Søraker* is a PhD Research Fellow at the Department of Philosophy, University of Twente, Netherlands. Currently, he is doing research on the ontology, epistemology and ethics of virtual reality as part of an International research group, funded by the Dutch research council, on “Evaluating the Cultural Quality of New Media”. He received his MA degree (Cand. Philol.) in Philosophy from the Norwegian University of Science and Technology, then working on philosophy of mind and AI. Since then he has published and lectured extensively on topics within philosophy of computing and computer ethics.

Introduction

Virtual worlds seemingly offer the opportunity to fulfill desires that might be otherwise unattainable, be it for ethical reasons or due to physical shortcomings and obstacles. As such, the increasing use and pervasiveness of virtual worlds could be regarded as good news, since it allows more people to pursue their conception of happiness, be it by way of more easily fulfilling their desires (desire-fulfillment theory), experiencing pleasure (hedonism) or to more easily attain the kinds of intrinsically valuable things that philosophers have proposed since ancient times (objective list theories), such as friendship, love, knowledge, aesthetic beauty and so forth.

Of course, this conclusion does not follow if the virtual “surrogates” are somehow inferior to the real thing. Thus, virtual worlds have, in many ways, become a test bed of our intuitions regarding the kinds of things that makes our lives fare better. Indeed, a number of philosophers, journalists, policy makers and others have argued that virtual worlds pose a threat to well-being, because we might come to settle for less valuable virtual surrogates. For instance, Albert Borgmann makes a distinction between instrumental, commodified and final communities and argues that virtual communities can at best be instrumental or commodified, because they do not contain “the fullness of reality, the bodily presence of persons and the commanding presence of things” (Borgmann 2004, p. 63). In a similar fashion Barney (2004) sees virtual communities as inferior due to their lack of physical practices and Dreyfus is critical of what he describes as the nihilist, irresponsible and often uninformed nature of virtual communities (Dreyfus 2004). Winner has also characterized what he refers to as the cyberliberterians’ conception of community as hollow and banal, primarily because they ignore the importance of “obligations, responsibilities, constraints, and mounds of sheer work that real communities involve” (Winner, 1997, p. 17).

A general line of argument in this vein is what Fred Feldman (2004) has dubbed ‘the argument from false pleasures’, exemplified by e.g. Robert Nozick’s ‘experience machine’ (Nozick 1993, pp. 42-45) and Shelly Kagan’s ‘deceived business man’ (Kagan 1998, pp. 34-36). In short, these arguments entail that some entities and experiences are false or inauthentic, and therefore should not be taken as conducive to well-being, at least not to the same degree as actual or authentic entities and experiences. Although the arguments suffer from some serious flaws, many of which I will address below, a proper scrutiny of the strengths and weaknesses of the arguments can yield important insights into the nature of virtual worlds and their potential impact on the quality of our lives.

In particular, I will focus on five important implications. First, I will argue that some virtual states of affairs are just as true or authentic as actual states of af-

fairs – that sometimes there is no significant epistemological difference between actual and virtual. Second, I will argue that virtual states of affairs are often connected to what Nozick refers to as “deeper reality” – that sometimes there is no significant ontological difference between actual and virtual. Third, I will discuss whether epistemological and ontological differences are ‘axiologically relevant’, i.e. whether these differences in themselves are of any importance when evaluating the value of virtual worlds and entities. Finally, I will argue that claims about the inferiority of virtual states of affairs should not be constructed as universal truths, and that we need to maintain a strict distinction between theories of well-being (or self-interest) and theories of ethics (or other-interest). Before turning to these considerations, some further elaboration of the argument from false pleasures is in order.

The argument from false pleasures

The argument from false pleasures is an argument against the proposition that if *x* produces a pleasurable mental state, then *x* is ipso facto conducive to the good life – or against the general hedonist (or ‘mental statist’) premise that only pleasurable mental states directly contribute to the quality of your life. If it can be shown, all else being equal, that an *x* produces more pleasure than *y*, but any reasonable person would still choose *y* over *x*, then there must be more to the good life than pleasure – or so the argument goes. However, although the arguments are intended to be arguments against naïve hedonism, they can also be interpreted as arguments to the effect that certain entities and experiences are less ‘pleasure worthy’ because they are illusory or inauthentic on one way or another.

The most famous and relevant versions of the argument have been proposed by Robert Nozick and Shelly Kagan. In Nozick’s famous thought experiment “the experience machine” (Nozick 1993, pp. 42-45), he asks us to envision the opportunity to plug into a machine that is capable of producing any pleasure you can imagine. You are given the opportunity to program this machine according to your conception of what you find pleasurable and live the next two years of your life having all of those pleasures satisfied. He further stipulates that we should disregard any worries concerning who controls the machine, malfunction, lack of sensory richness and so forth. If we still choose not to plug into the experience machine, in spite of the premise that it will provide us with more pleasure than if we choose the actual world, then there must be more to well-being than pleasure alone.

A similar challenge against hedonism can be found in Kagan’s “deceived businessman” (Kagan 1998, pp. 34-36). Imagine a businessman who regards his life as good because he believes that his colleagues, his wife and his children love him. In reality, however, they all despise him and merely give the appearance that they

love him. Kagan argues that hedonist theories of the good life would have to conclude that the businessman has succeeded in attaining a good life, which seems to be at odds with our intuition that it is not the kind of life we would want for ourselves.

What the arguments have in common is that they draw on the intuition that any reasonable person would choose e.g. love that corresponds to true states of affairs in the world (e.g. a human being who actually loves you) over love that corresponds to false states of affairs, be it to a human being who does not actually love you (the deceived businessman), or to someone who is not an actual human being (the experience machine). In other words, pleasures that are in some way illusory or inauthentic are believed to be inferior to those that stem from objectively true or authentic states of affairs. When applied to virtual states of affairs, the arguments entail that they lack the kind of ontological objectivity afforded by actual or physical states of affairs.

Lessons learnt from the argument from false pleasures

The first three implications addressed below follow from the so-called principle of formal equality, which is a helpful guideline for evaluating the validity of certain arguments. The principle states that a difference in treatment or value between two kinds of entities can only be justified on the basis of a relevant and significant difference between the two. This principle underpins a number of arguments in applied ethics, animal ethics and bioethics in particular. To take an example, if we treat or value human beings differently from other animals this must be grounded in (1) a theoretical difference between humans and other animals, and (2) a discussion of whether this difference justifies a difference in treatment ('ethical relevance') or evaluation ('axiological relevance'). In this manner, many issues in animal ethics can be approached by first discussing the theoretical differences between humans and other animals (e.g. by pointing to some animals' lack of central nervous system, lack of linguistic ability, lack of sentience or similar), and subsequently discuss to what degree these differences are relevant to their value. In a similar manner, one way of discussing the validity or relevance of the argument from false pleasures is to discuss whether the differences the argument refers to really apply to all kinds of virtual entities and experiences. In this regard, the argument admits of two interpretations. On an epistemological interpretation the argument entails that beliefs that correspond to true states of affairs are more valuable than those that correspond to false pleasures. On an ontological interpretation, the argument entails that beliefs that correspond to physical states of affairs are more valuable than those that correspond to non-physical states of affairs. Thus, one way of evaluating the validity

of the arguments is to discuss whether there are significant epistemological and/or ontological differences between virtual and actual states of affairs.

There is not always an epistemological difference

The notion of 'virtual' is contested and notoriously difficult to define. It is often used in the same sense as 'quasi-', 'pseudo-' or 'almost the same as'; something that is almost but not quite real. This common usage invites the conclusion that virtual states of affairs are in some sense false or illusory. However, if we employ a more technical definition, it becomes clear that this does not necessarily follow. Although there is no consensus on a definition of virtual, for present purposes it is sufficient to point out that any definition must include the fact that virtual states of affairs are sustained by a computer. An underlying computer simulation is, as it were, the condition of possibility for all things virtual. This seemingly trivial characteristic entails that virtual worlds are radically different from mere products of the mind. First, the computer simulation allows a significant degree of regularity, such as permanence of objects and causal chains of events, and it allows a persistent space in which multiple users can interact – with the world and with each other. Second, this regularity, combined with a persistent three-dimensional space shared by multiple users, gives rise to intersubjectivity. When you have a certain degree of regularity combined with intersubjective availability, this opens up for all sorts of epistemologically objective facts about virtual states of affairs. It is, for instance, either true or false that my avatar is currently located on the top of a virtual mountain in the virtual world of Second Life. It is true in the sense that anyone can test whether this is in fact the case, and it can be tested because it is intersubjectively available and regulated by the underlying computer simulation. By contrast, the content of dreams, hallucinations and other products of the mind do not have the same degree of regularity and they are not intersubjectively available. As such, claims about virtual states of affairs are radically different from purely subjective, first-person experiences, and statements about virtual states of affairs can be true or false.

Still, most virtual states of affairs are different from actual states of affairs, in that they are, in part, observer-dependent, i.e. in contrast with physical entities (skepticism aside) they depend on conscious beings for their existence. If there were no conscious beings around, the gravitational force between the moon and the earth would still exist, but not the virtual gravitational force between a virtual moon and a virtual earth. According to John Searle's terminology (cf. Searle, 2001, p. 54ff), facts about physical states of affairs are both ontologically and epistemologically objective. Facts about virtual states of affairs, on the other hand, are always ontologically subjective but can still be epistemologically objective. In other words, virtual states of affairs correspond to what John Searle

refers to as social reality, i.e. entities that exist only as a result of an intersubjective recognition among conscious beings that “x counts as y in the context c” – what Searle refers to as a ‘status function’. For instance, Barack Obama counts as president in the context of the United States, and if you disagree you are simply wrong; it is an epistemologically objective fact even if neither presidents nor the USA would exist as such if there were no conscious beings. In a similar manner, if my avatar is in possession of fifty Linden dollars (the currency of Second Life) and currently sits on top of a virtual mountain wearing a virtual black robe, you would be wrong to disagree. This fact is partly constituted by the regularity and intersubjective availability that is maintained by the computer, and partly by the ability of a virtual community to collectively recognize status functions. Virtual states of affairs are observer-dependent and require an underlying computer simulation, but this does not necessarily entail an epistemological difference. Thus, although these entities and facts might not be connected to a deeper reality, they are not thereby necessarily false. For instance, if someone derives immense pleasure from having created a virtual piece of art, this is comparable to someone deriving immense pleasure from having created a physical piece of art. That the entities actually exist is an epistemologically objective fact, even if there is a significant ontological difference between them. The pleasure taken in the two achievements are both significantly different from taking pleasure in having created a piece of art if you have, as a matter of fact, done no such thing (i.e. if it is a false belief).

There is not always an ontological difference

As argued above, some virtual states of affairs are epistemologically objective. Thus, keeping the principle of formal equality in mind, the argument fails to point out a significant epistemological difference between all virtual and actual entities. However, there is clearly an ontological difference between, say, the virtual and the actual piece of art; they have a different mode of existence. This ontological difference is what prompts Nozick to conclude that one reason for no plugging into the experience machine is that it has no “actual contact with any deeper [physical] reality” (Nozick, 1993, p. 43). This is clearly the case for many kinds of virtual entities, but is it always the case?

It is telling that Nozick constructs his experience machine as a “single player” environment, where perhaps the most significant “falsehood” would be that the people (or bots) we would interact with would not be connected to conscious beings. In virtual worlds, however, the avatars we interact with are precisely controlled by conscious beings, thus connected to “deeper reality” in a non-trivial sense. To take but one example, Julian Dibbell’s (2007) account of the first virtual rape illustrates how purely ‘intravirtual’ states of affairs can have dramatic ‘ex-

travirtual' effects; that the virtual entities are indeed connected to a deeper reality. Thus, as long as we are talking about 'virtual worlds', which on my definition only includes multi-user virtual environments, we can make true or false claims about the intentional states of those we interact with. That is, Nozick's claim is correct when it comes to simulation of physical objects, but not true of avatars as mediators for conscious beings. For instance, even if virtual relationships are mediated by an avatar, you do have a relationship with an actual human being, albeit one that might still be inferior due to other limitations (see e.g. Cocking & Matthews 2000).

Avatars being mediators for conscious beings is one way in which virtual experiences are connected to a deeper reality. If we return to the kinds of institutional entities discussed in the previous section, there are other ways in which virtual states of affairs are connect to actual ones. According to Searle, one of the most important facts about institutional entities is that they often iterate and interconnect in highly complex ways. For instance, money is not something that gains its power from the status function "this piece of paper counts as money" itself, but the way in which it interlocks with innumerable other social facts, such as contracts, properties, banks, wages, stocks, taxes, inheritance, fines and so forth. Although these entities can be reduced to a set of fundamental status functions, it is often difficult and needless to do so. For a 100 dollar bill to be exchanged for renting a piece of property, we do not need to be aware of the complex history of the evolution of currency and property rights. What is important is that these interrelations can cut across any ontological difference between actual and virtual. For instance, 'virtual money' (a phrase that hardly makes sense) can be intertwined with "actual" money, in such a way that having million dollars worth in Linden Dollars is equivalent to having million dollars worth in any currency. Furthermore, these interrelated status functions are likely to become more prevalent as virtual worlds come to occupy more important roles in our lives. For instance, in Second Life there has been some (perhaps unrealistic) debate on whether or not virtual wedding ceremonies (x) ought to count as a wedding (y) not only in the context of the virtual world (c) but also in the context of the real world – with the complex rights and privileges that come with actual weddings. I shall not discuss whether this is a good idea or not, but only point out that there are no in-principle obstacles to this. Just like an Elvis impersonator at a drive-in wedding chapel in Las Vegas can have the right to pronounce you man and wife, so it is in-principle possible that a priest (who is authorized in the actual world) can perform virtual wedding ceremonies as a virtual priest in a virtual wedding chapel that lead to the same rights and obligations as a wedding performed in the actual world – there is not necessarily an ontological, in-principle difference between the two.

Do the differences matter?

Although many will find Nozick's intuition pump convincing, it seems to entail more than many people are probably willing to accept. For instance, many people regard their belief that there is a supernatural being as part of what makes their lives worth living. If Nozick is right, we would have to conclude that if there are no spiritual beings, all those who derive well-being from their religious beliefs would actually have seriously diminished quality of life (whether conscious of the falsity of their beliefs or not) – in much the same way as the deceived business man. If there were one God, all those who believe in the “wrong” god would have seriously diminished quality of life. It seems absurd that someone's quality of life could be determined by the unanswerable question of whether, as a matter of fact, there is a God or not. If we presuppose, as seems plausible, that some people do have improved quality of life because of their religious beliefs, this shows that quality of life does not have to be founded on epistemologically or ontologically objective states of affairs. It can also, descriptively speaking, be grounded in beliefs that are more or less removed from empirical testability. The same also holds for other non-evidential means of being confident in something despite lack of evidence. For instance, trust is in many cases constituted by a belief that someone will do something even if we are not fully justified in this belief. This illustrates that we should not always base our judgment of value on whether or not the things we value correspond to true states of affairs.

The considerations above do not entail that there is no relevant difference between actual and virtual. For instance, virtual worlds might not give rise to the same kind of trust that we can have in the actual world (cf. Weckert, 2005). However, the argument from false pleasure is supposed to entail that virtual states of affairs are less valuable because of the fact that they are not true or not physical alone, which is different from saying, for instance, that trust is impossible in virtual worlds due to lack of e.g. the kinds of behavioral cues we will find in non-mediated communication.

Theories of well-being are not ethical theories

The argument from false pleasure also sheds light on a number of more meta-ethical issues, for instance regarding the difference between what is good for oneself and what is good for others. It seems trivial to claim that theories of self-interest are fundamentally different from theories of other-interest, but the two are commonly seen as reciprocal (most notably, at least on some interpretations, in Plato) so I will briefly discuss how the argument from false pleasures illustrates why they ought to be kept separate. The purpose of Nozick's argument is to pump the intuition that we would not plug into the experience machine, and Nozick claims that the reason behind this intuition is that we think there is more to the good life

than just pleasure. I think Nozick is half right, but that the conclusion in terms of well-being is wrong. That is, we might very well hold that the only things that directly increase well-being are those that produce pleasure, yet at the same time acknowledge that we have certain ethical obligations that prohibit certain pleasures. In other words, we cannot always do what would be in our own self-interest, because sometimes this is at odds with the interest of others, and sometimes this entails that we should not do it – ethically speaking. Thus, the intuition that Nozick's argument draws on might for many people be an ethical one. Nozick rightly points out that one reason not to plug in is because it would be tantamount to a form of suicide, but he does not specifically mention the fact that some people might desire to commit suicide but choose not to because of their ethical commitments. This is also illustrated by the reverse form of the argument. Consider a Matrix-like scenario in which you have been in the experience machine all your pleasurable life and you are offered the opportunity to leave. This will probably be a more difficult choice for many, and I believe one of the main reasons is that leaving such an experience machine would not involve the breach of any ethical commitments.

If we consider not only our own but all possible kinds of responses to the choice of plugging into the experience machine, it becomes clear that the reasons for (not) doing so will often be an intricate blend of ethical and prudential reasons. Some might refuse it because it is not in their own self-interest, whereas others will refuse it because it would not be in the interest of others about whom they care. If we conflate the ethical and prudential notion of the 'good', we run the risk of undermining the human ability to act against our own self-interest for the good of others; to recognize the greatness of those who have sacrificed their own well-being for the benefit of others. Another reason for keeping the distinction is that this allows us to have a subjectivist theory of well-being that allows for a wide variety of individual conceptions of the good, while at the same time maintaining a strict objectivist theory of ethics. In other words, it is perfectly consistent to hold that whether something is good for you is largely a matter of personal preferences, but whether it is ethically justifiable is not a matter of personal preferences.

This distinction is also important to keep in mind when assessing the potential impact of the increasing use and pervasiveness of virtual worlds. For instance, if the mother of a young child spends the majority of her time in virtual worlds, the question of how this impacts her quality of life is largely irrelevant; the question should be whether this activity causes harm to the child, not whether it causes happiness in the mother.

'Opportunity costs' and the contingency of well-being

Nozick is probably right that many people would choose not to plug in and, for the sake of the argument, let us go along with the stipulation that this conclusion is reached on the basis of self-interest. Although Nozick might be right about many, perhaps even the majority of people, it seems plausible that some people would choose to plug in – for good reasons. For some people, and it should be easy enough to imagine examples, their well-being is so diminished and their prospect for improved well-being so small that they would easily choose the experience machine over their miserable lives – especially if they have no ethical commitments to others. This illustrates that the evaluation of something as detrimental to one's quality of life can never be a universal claim, but must be made with an eye to the individual's particular situation.

This can be illustrated through the notion of 'opportunity cost'. This is primarily a term from economics, and relates to the choice between two or more desirable but mutually exclusive actions. For instance, if a company decides to build a new parking lot, the costs and revenues are not merely those of building the parking lot, but also the revenue lost by not creating, say, a new building or a park instead. Thus, 'opportunity cost' is an apt term for describing the relation between two activities that are both desirable, but where one is less so than the other – which will often be the case with virtual surrogates. The interesting thing is that the relevant opportunity costs to take into consideration are only those that are sufficiently plausible. The company in the example above need not consider whether the potential benefits of building a space station. In a similar manner, it seems irrational and detrimental to well-being for a person with severely limited opportunities to forego the opportunity to have virtual friendships just because they might be less valuable. This is certainly controversial, because it raises the question of when it is rational to settle for virtual surrogates, but the main point is that this is a matter of threshold, not a matter of universal principle.

Concluding remarks

The purpose of this paper has been to draw out some of the implications of Nozick's experience machine argument and similar arguments from false pleasures, in particular their implication for how we ought to evaluate the impact of virtual worlds on the quality of our lives. I have argued that sometimes there is no significant difference between actual and virtual and that sometimes this difference is irrelevant. I have also argued that we ought to keep a strict distinction between considerations of self-interest and other-interest; between theories of wellbeing and ethical theories. Finally, I stressed the importance of not making universal, absolute claims about what can and what cannot contribute to well-being, given the enormous differences in our individual capabilities, opportunities and con-

ceptions of the good. We should not draw the conclusion that there is nothing lost by settling for virtual surrogates, but that there is not always an epistemological or ontological difference between virtual and actual, and even when there is such a difference, it is not always relevant. As I have tried to show, however, the arguments do still prompt a number of considerations that are necessary for any evaluation of the impact of virtuality on our lives.

REFERENCES

- Barney, D.** (2004). The Vanishing Table, or Community in a World That is No World. In A. Feenberg & D. Barney (Eds.), *Community in the Digital Age: Philosophy and Practice* (pp. 31-52). Oxford: Rowman and Littlefield.
- Borgmann, A.** (2004). Is the Internet the Solution to the Problem of Community? In A. Feenberg & D. Barney (Eds.), *Community in the Digital Age: Philosophy and Practice* (pp. 53-68). Oxford: Rowman and Littlefield.
- Chalmers, D.** (2005). The Matrix as Metaphysics. In C. Grau (Ed.), *Philosophers Explore the Matrix* (pp. 132-176). Oxford: Oxford University Press.
- Cocking, D., & Matthews, S.** (2000). Unreal friends. *Ethics and Information Technology*, 2(4), 223-231.
- Dibbell, J.** (2007). My Tine Life: Crime and Passion in a Virtual World. Retrieved January 1, 2007, from <http://www.lulu.com/content/1070691>
- Dreyfus, H. L.** (2004). Nihilism on the Information highway: Anonymity versus Commitment in the Present Age. In A. Feenberg & D. Barney (Eds.), *Community in the Digital Age: Philosophy and Practice* (pp. 69-82). Oxford: Rowman and Littlefield.
- Feldman, F.** (2004). *Pleasure and the Good Life: Concerning the Nature, Varieties, and Plausibility of Hedonism*. Oxford: Oxford University Press.
- Kagan, S.** (1998). *Normative Ethics*. Oxford: Westview Press.
- Malebranche, N.** (1688/1997). *Dialogues on Metaphysics and on Religion* (N. Jolley & D. Scott, Trans.). Cambridge: Cambridge University Press.
- Nozick, R.** (1993). *Anarchy, state and utopia*. Oxford: Blackwell.
- Searle, J.** (2001). *Rationality in action*. Cambridge, Mass.: MIT Press.
- Weckert, J.** (2005). Trust in Cyberspace. In R. J. Cavalier (Ed.), *The Impact of the Internet on our Moral Lives* (pp. 95-117). New York: SUNY Press.
- Winner, L.** (1997). Cyberlibertarian myths and the prospects for community. *SIGCAS Computers and Society*, 27(3), 14-19.

Remixing and Recoding: Revisiting the Copyright Wars

Richard A. Spinello*

Department of Operations and Strategic Management
Carroll School of Management
Boston College

Abstract

This paper examines two contentious disputes in the area of copyright law: remixing and recoding. It argues that while copyright law needs some prudent reform in order to accommodate more remixing, it should not be modified to permit all types of recoding. As Hegel and others have postulated, a creative product, whose message reflects the interior dimensions of authorial creativity, is the embodiment of an author's personality. As a result, undermining a work's integrity by recoding the intended meaning in objectionable ways impairs authorial dignity. Secondary users need limited access to creative works, but intellectual property rights should continue to protect the integrity of the work by giving the author the tools to safeguard its original meaning and message within the bounds of social fairness.

Keywords: control, copyright, fair use, Hegel, integrity, personhood theory, re-code, remix, secondary users, social meaning.

Introduction

Much of the debate about current copyright law focuses on the issue of a proper balance between protection that is too strong or too weak. Inadequate protec-

* *Richard Spinello* is Associate Research Professor and Director of Ethics Programs, Carroll School of Management, Boston College. He is the author of numerous books, such as *A defense of intellectual property rights*, Edward Elgar Publishing, 2009 (with Maria Bottis), *The Genius of John Paul II: The Great Pope's Moral Wisdom* (Lanham, MD: Sheed & Ward, 2007), *Intellectual Property Rights in a Networked World: Theory and Practice* (Editor with Herman Tavani). New Brunswick, NJ: Idea Group Publishing, 2005, *Case Studies in Information Technology Ethics* (Upper Saddle River, NJ: Prentice-Hall, 2003), *Regulating Cyberspace: The Policies and Technologies of Control* (Westport, Conn.: Quorum Books, Greenwood Press, 2002) *Readings in Cyberethics* (Editor with Herman T. Tavani) (Sudbury, MA: Jones & Bartlett, 2001; Second Edition published in 2004).

tion deprives authors of their just opportunity to control their literary work and appropriate its social and economic value. In addition, without reasonable copyright protection, the incentives to create new works will be inadequate. On the other hand, if protection is too strong, the creative efforts of future downstream authors can be impeded by the limits placed on the availability of cultural resources.

The main issues in the ongoing debate about this proper balance are crystallized in two well-known and contentious controversies surrounding remixing and recoding. Lessig (2004), a long time champion of digital creativity and “free culture,” has argued with some insistence that users should be given broader fair use rights in order to blunt the encroachment of a “permission culture.” Specifically, users should be allowed to remix digital content by recombining pieces from different cultural objects even if those objects have a copyright. Filmmakers, for example, should be allowed to construct new movies out of clips compiled from digital movies located on computer systems around the world. Such a creation, of course, is technically illegal, but Lessig (2008) maintains that the law must be changed, so that ordinary people become “producers” of culture, not just “consumers” of culture. In this way we can return to an “amateur” creative culture that supports the participation of the multitude instead of just an elite few.

The second dispute involves a more complex controversy that revolves around the issue of recoding. Some reformers and deconstructionists advocate the need for marginalized groups to recode intellectual property by radically changing the primary intended meaning of a copyrighted work. They claim that without a broad latitude to recode, cultural symbols and icons can become oppressive. The two issues intersect in some cases where remixing leads to objectionable recoding.

In this paper we will argue that amateur users require more creative freedom to work with cultural objects. Therefore, remixing within limits should be permissible, especially if the user has no commercial interests and there is no “unjust enrichment” at the expense of the copyright holder. For the most part, copyright law should not regulate this form of creativity. Recoding, on the other hand, beyond the bounds of the current fair use parameters is far more problematic. We argue that unfettered recoding is inconsistent with the moral basis of copyright protection, and turn to Hegel’s property theory to support this perspective. We begin this analysis with a brief overview of current copyright law.

U.S. Copyright Law: An Overview

Of all forms of intellectual property disputes, the ones involving copyright issues are arguably the most contentious. The historical origins of copyright protection in Anglo-American and European jurisprudence are well-documented (Spinello

and Bottis 2009). There is no doubt that the scope of copyright protection has expanded over the last several decades more than necessary, and this expansion has worked to the detriment of creativity. At the same time, it is hyperbolic to claim that creativity is “in chains” or that the law thoroughly encumbers new creators by “locking down” the culture.

In the wake of digital technology, copyright disputes have proliferated over the past decade. A pivotal point was the introduction of the MP3 format allowing for the digitization of music. This was followed by epic legal battles pitting the content industry against on-line companies like Napster, MP3.com, and Grokster. While the content industry has adapted somewhat to new technology, it continues to favor its traditional business model and still relies on the heavy hand of the law to protect its property rights. The Youtube generation, however, wants easy and free access to digital content, and so the legal wrangles persist, with the spotlight continuing to shine on copyright law.

Copyright law protects the expression of ideas but not the underlying ideas themselves. The law is designed to keep off limits the “raw materials” of creativity such as ideas, common plots, scenes à faire, and common facts, in order to avoid burdening new creators. According to O’Rourke (2001), “copyright law does not extend protection to factual information because such information is the core raw material that others need to use to further progress.”

The expression must be original, and it must be fixed in a tangible medium in order to qualify for a copyright. In the case of music, for example, two different copyrightable works are recognized: the composition itself, that is, the music and lyrics and the recorded version of a musical work.

The copyright statute grants creators of content several exclusive rights in their expressions: «An owner of a copyright has the exclusive right to reproduce, distribute, and publicly display copies of the work» (17 U.S.C. 106, 2000). Let us briefly consider each of these rights. First and foremost is the right to reproduce the copyrighted work, which protects a literary, musical, dramatic, artistic, architectural, audio, or audiovisual work from being copied or reproduced without the permission of the copyright holder. Even a temporary copy of a music or movie file stored in a computer’s random access memory (or «RAM») requires permission to be reproduced.

Copyright law also gives the copyright holders exclusive distribution rights, that is, the sole right to distribute their works to the public. Court rulings, for example, have held that unauthorized uploading of an MP3 file, that is, making it available or exposing it over the Internet is tantamount to the public distribution of that digital music or movie file, and hence it is illegal.

In addition, copyright law gives the copyright holder the right to prepare derivative works based upon the pre-existing copyrighted work along with the right to authorized public performances of the copyrighted work. This includes the right to publicly display the work. The law maintains that when a creative work is made available on a web site, that work is being «displayed.»

Based on this broad set of exclusive rights, infringement can occur in one of two ways: illegitimate access to a work (for example, making copies without permission) or the creation of a derivative work that is too similar to its original. The ideal copyright framework will “assure authors the right to their original expression, but encourage others to build upon the ideas and information conveyed by a work” (Feist Publications, Inc. v. Rural Telephone Service Co. 1991).

Finally, there are limitations built into copyright law, most notably, first sale and fair use, which are sometimes referred to as “safety valves.” The fair use doctrine is particularly apposite for this discussion and deserves some elaboration. Fair use means that every author or publisher may make limited use of another person’s copyrighted work for purposes such as criticism, comment, news, reporting, teaching, scholarship, and research. There are four factors that help the court determine fair use: (i) the purpose and character of the use [for example, commercial use weighs against the claim of fair use]; (ii) the nature of the copyrighted work [for example, creative works receive more protection than factual ones]; (iii) the «amount and substantiality of the portion used» in relation to the work as a whole; (iv) the effects of the use on the market for the work: «fair use, when properly applied, is limited to copying by others which does not materially impair the marketability of the work which is copied» (Harper & Row v. Nation Enters., 1985). All of these factors are weighed together and decisions are made on a case by case basis. Thus the fair use doctrine would enable a teacher to reproduce and distribute a few paragraphs from a book or magazine article, but it would probably not allow reproduction of the whole article or several chapters of the book. Also making private copies of certain material is considered fair use. For example in *Sony v. Universal* (1984) the U.S. Supreme Court affirmed that consumers can make a video copy of a television program for their own private use.

Another exception is the «first sale» doctrine. The first sale provision allows the purchaser of a copyrighted work to sell or lend that copy to someone else without the copyright holder’s permission. Both of these limits on copyright law are designed to balance the rights of the copyright holder with the public’s interest in the broad availability of books and other artistic works.

The paradigmatic legal dispute over copyright in the nascent digital era has centered on file sharing over peer-to-peer (p2p) networks (Spinello 2008). Defenders of the unrestricted use of these p2p networks to share copyrighted music ar-

gued that even if sharing digital copies over a network is equivalent to making an unauthorized reproduction of a copyrighted work, that action comes under the fair use exception. Some scholars contend that it is far from evident under current law whether individual users are liable for copyright infringement if they engage in «personal copying.» On the other hand, most legal scholars admit that «it is generally an infringement to download large amounts of copyrighted material without permission; even if you already own the corresponding CD, the case could be made that a network-derived copy is infringing» (Gantz and Rochester 2005). There is broadening consensus that noncommercial file sharing among ordinary users should be decriminalized. Nonetheless, the courts have been sympathetic to the content industry (MGM v. Grokster 2005).

Remixing

In addition to the ongoing dispute over file sharing, a second and related controversy has emerged in recent years. Digital technology has made it increasingly possible for users to exercise their creative freedom through remixing. Remix “artists” recombine different pieces of content, such as images, music, and text into a new creative whole. Even amateurs can begin to “write” and create using a *mélange* of music, video, and images. For example, a popular piece of music known as Night Ripper remixes 200 samples from 167 music artists. Remixing, however, is “presumptively illegal” (Lessig 2008). If a young woman remixes photos of her vacation with a catchy song from Cyndi Lauper (“Girls Just Want to Have Fun”), she has technically violated the copyright law by using this copyrighted song without permission. The copyright is infringed even though this remix is an amateur and noncommercial work. In one case Universal Music Studio sued Stephanie Lenz for using a 20 second clip of a song by Prince (“Let’s Go Crazy”) to accompany a video of her eight month old baby, who was dancing to this music. Despite the fact that there was no market for this amateur and poor quality production and no potential loss of revenues for the company, Universal vigorously pursued this case, claiming that it was deliberate copyright infringement.

Like the examples we have cited, most forms of remixing do no harm to copyright holders. These derivative works are unprofessional and non-commercial, have no effect on future market sales, and respect the original integrity of those cultural objects that have been recombined. At the same time, remixing should be encouraged. It’s a growing phenomenon that signals the transition from a read only (RO) culture to a read write (RW) culture (Lessig 2008). If people are so inclined, they should not be passive absorbers of culture but co-creators of sorts. The benefits of these derivative forms of creativity are limited, but they help build community

and surely have some educational value. Hence a tenable case can be made that remixing activities enhance social welfare.

Given that most forms of remixing are innocuous and not intended to adversely affect the marketability of a copyrighted product, revision of the copyright law seems in order so that the law will no longer inhibit these new forms of creativity. A careful reform of copyright law could be implemented so that it would be easier for users to engage in amateur remixing activities. Unlike professional remix, this assumes that there is no commercial distribution and the remix author stands nothing to gain for his efforts. According to Lessig (2008), we need “to restore a copyright law that leaves ‘amateur creativity’ free from regulation.” The specifics of how this revision would be crafted are beyond the scope of this paper. Let it suffice to say, however, that the noncommercial boundary must be firmly established. If an amateur remix (such as the Cyndi Lauper song and the vacation photos) ends up being broadcast on a commercial website such as cbs.com, copyright restrictions would then apply.

The Recoding Controversy

Lessig’s proposal to facilitate amateur remixing is one of many recommendations that would alter the landscape of copyright law, and thereby adjust the balance of protection in favor of the user. How extensively the landscape would change depends on many factors such as how to differentiate professionals from amateurs and how to deal with amateur distribution over popular web sites such as youtube.com.

Of course, calls for reform of copyright jurisprudence are not confined to demands to decriminalize file sharing or exempt amateur remixing from the long reach of the law. Some legal scholars and activists seek far more drastic reform. They are keen to dismantle the current asymmetry and imbalance between producers of culture and consumers of culture. While these reformers would applaud the reforms advocated by Lessig, their goal is to expand the creative freedom of users in a far more substantial fashion.

These reformers criticize the current copyright law for its tendency to protect original authors “while chilling downstream users” and stifling various forms of cultural exchange (Craig 2007). They argue that the copyright system is predicated on a “moral divide” between the independent creative author and downstream derivative expression. This system, heavily biased in favor of authors versus copiers and original creators versus imitators, is an anachronism in a post-modern world that calls into question fundamental notions of authorship, originality, and independent creativity (Craig 2007).

Indeed, at the foundation of copyright law lies a conventional notion of authorship, the independent creator who deserves some sort of legal protection for his or her original work. Without a clear and unambiguous sense of authorship it becomes difficult to assign a copyright. In a frequently cited U.S. case, copyright protection was first granted to photographers based on “the nature of authorship and of originality, intellectual creation, and right to protection” (*Burrows-Giles Lithographic Co. v. Sarony* 1884). Inspired by various trends in post-modernity, however, legal scholars have sought to dismantle this romanticized notion of the author. According to Halbert (1999), the autonomous “proprietary author” is a “myth” of the copyright system. Individual authorship is socially constructed idea according to a Romantic ideal. According to this line of reasoning, we must be especially wary of associating a creative work exclusively with a discrete, individual author. Despite the author’s efforts, his or her creative output is not a product of this one individual but of various social forces and multiple cultural inputs that constitute the author’s finished work. Similarly, it makes little sense to talk about a fixed creative work with a determinate meaning.. Roland Barthes (1977) proclaimed the death of the author along with the end of the literary work as some sort of fixed entity with a stable meaning. Interpretation of texts is completely “intertextual,’ inter-related to other texts, fluid, and boundless (Craig 2007). In summary, the notion of autonomous authorship “is theoretically suspect, texts are unstable and determined in large part by their readers, and originality is more often a pose or a pretense rather than a definable aspect of a work” (Vaidhyathan 2001).

Once the theoretical underpinnings of copyright are destabilized we have far more latitude to re-conceive the legal parameters for copyright protection. To remedy the problem of asymmetry between creators and consumers of culture, copyright reformists want to amend the law so that consumers can reconfigure and reshape the original meaning of literary and artistic works. Along these lines, they argue for a broad right to recode. What is the precise meaning of the term “recoding”? According to Coombe (1991), recoding is “the consumption of commodified representational forms in a productive activity in which people engage in meaning-making to adapt signs, texts, and images to their own agendas.” More simply, recoding refers to the modification of a work’s primary intended meaning by “secondary” users. In other words, the consumer of cultural works is also a producer who can take those works and adapt them by changing the underlying meaning even in ways that are radically transformative. Recoding typically applies to trademarked works such as a brand or an icon or to copyrighted works such as music, movies, novels and plays.

Sometimes secondary users want to substantially recreate an existing copyrighted work in order to advance a particular social or political agenda. Consider the

proposal several years ago (refused by the copyright holder) for the cross-gender casting of the movie, *Steel Magnolias*. In a similar case, a theater group in Europe sought to put on an all female production of Becket's play "Waiting for Godot." (Lindon contra La Compagnie Brut de Buton 1992). Many intellectual property critics and legal scholars argue passionately for radical recoding freedom that would have allowed both of these productions. They support the user's right to reconstitute the meaning of a text or an image without fear of reprisal from the copyright holder.

Not everyone in the legal community accepts the arguments for recoding rights. La Voi (2004) distinguishes between the "high protectionists" who want to maintain the traditional copyright regime and the "deconstructionists" who argue for much thinner protection that will accommodate the increased freedom to recode. The deconstructionists contend that secondary users should be allowed to 'stamp their unique meaning on another's intellectual property before passing it along' (La Voi 2004). They further contend that limiting the freedom to recode often works to the disadvantage of marginalized groups who may want to use a valuable persona or trademark to convey their minority viewpoint. Coombe (1996), for example, contends that authorial works function in a way that silences the "other," who struggles against "forces of alterity that operate as dangerous supplements to the integrity of the author/work relationship." Also, without the right to recode (or at least, broad fair use rights), copyright can become an "instrument of suppression" (Gordon 2002). Gordon cites the famous example of Alice Randall's parody of Margaret Mitchell's famous civil war novel, *Gone with the Wind*, known as *The Wind Done Gone*. Mitchell's estate sued and the initial result was an injunction against Randall's novel. Mitchell's novel was seen as an affront to Blacks in the south and Randall's response was this elaborate parody. According to Gordon (2002), works of this nature should be encouraged: "predecessor authors should not be entitled to harm us and then use copyright to prevent us from having redress." Free riding, in Gordon's view, is simply a manifestation of the acute need for people to rely on each other to make and remake the culture.

Other arguments in favor of broader recoding rights focus on the utilitarian benefit of thin copyright protection, which will enhance and expand the marketplace of ideas. Advocates for loosening copyright restrictions contend that reconfiguring the law will create the opportunity for a more diverse and robust social discourse. Thick copyright laws, on the other hand, tend to promote a more homogeneous discourse. In Aoki's (1996) view, the exclusive controls currently granted to authors inhibits cultural innovation since cultural objects are off limits to future creators: "increasing aspects of our culture imaginary are being fenced 'off-limits' as intellectual property, marked with the equivalent of 'no-trespassing signs.'"

What are we to make of these demands for expanded recoding freedom? Should the law be altered to accommodate recoding as well as amateur remixing? It might seem so since these cases bear some similarity. In both situations commercial and economic interests are typically not at stake. Both cases revolve around the issue of creative freedom. Those who advocate for more recoding freedom are simply seeking a greater voice in the marketplace of ideas – they are not trying to free ride or make money off another person's creative efforts. Deconstructionists often link the right to recode with free speech rights and argue that the liberties guaranteed by the First Amendment should take priority over intellectual property rights. Also, if the theoretical underpinnings of intellectual property law have been destabilized, can we continue to base policies on faulty assumptions about individual authorship or originality?

There is a big difference, however, between recoding and remixing. Remixers usually want to preserve the original meaning of a work or alter the meaning in a marginal and benign ways by associating this work with other cultural objects. Recoders, on the other hand, typically seek to radically transform and re-shape the social meaning of an artistic work to reflect their own social viewpoint. Often they want to subvert the original meaning of the work and so the recoded product becomes objectionable to the copyright holder. Hence, the key question at the center of the recoding debate is who should control the social meaning of creative works? Should such control be the exclusive prerogative of the author/creator or should the audience who consumes the work have more input into its meaning? Should the audience be given much more space for “talking back” to creative authors? (La Voi 2004).

To a great extent, how one answers these provocative questions from a moral perspective will depend on two issues: (i) how stable are the concepts and presuppositions underlying copyright protection and (ii) how should we most properly understand the grounding of intellectual property rights. Are those rights justified purely on economic grounds or is there a valid non-economic justification that should be operative in making a determination about this demand for the increased freedom to recode

Space constraints forbid us from addressing the first question in any depth. Elsewhere I have explained that the assumptions of the post-modern critics are highly questionable and their announcement about the death of the author is a bit premature (Spinello and Bottis 2009). It is worth considering one issue, however, because of its relevance to this debate. If meaning is so fluid and indeterminate, perhaps the question of recoding is just moot. Is Barthes correct when he asserts that a text has no definite, determinate meaning? This idea sounds attractive, but if all textual meaning is variable and open to unlimited interpretation,

what about the meaning of Barthes' own claims found in his seminal text on the death of the author? If we express his preeminent claim about meaning as a simple proposition it might look something like this: "all textual meaning is variable and subjective." But what about the meaning of this proposition itself? The proposition cannot survive: either that proposition is also open to interpretation, in which case, it can mean something else or it means what it says, in which case there is fixed and determinate meaning and the proposition is self-refuting. Those who want to relativize the meaning of all textual objects have to deal with the epistemological problem that their own written or verbal assertions are also relative.

Second, what is the theoretical basis for resisting the proposal in favor of recoding rights? A number of persuasive theories have been put forth to justify these exclusive rights, but those with the greatest intellectual resonance can be found in the philosophical writings of John Locke and G.W. F. Hegel and in the philosophy of utilitarianism. Locke is credited with providing the philosophical underpinnings of the labor theory of ownership, while aspects of Hegel's elaborate philosophical system form the basis for the so-called 'personality theory.' Utilitarianism contains a more pragmatic philosophical approach that has been most appealing to economists and legal scholars. All three theories are applicable to intellectual as well as physical property. Elsewhere I have demonstrated the plausibility of the non-economic rationale for copyright protection such as the Hegelian approach, which has particular salience for this controversy (Spinello and Bottis 2009).

Consequently, our primary ethical argument for this position against an expanded recoding prerogative is based on Hegel's theory of property which firmly grounds property rights in the personhood interests of authors. The Hegelian paradigm has emphasized that people base aspects of their personality on property, which is essential for the realization of freedom. Let us review the essentials of Hegel's position.

Hegel's political philosophy centers on the notion of ethical life (*Sittlichkeit*). Ethical life is actualized when it is incarnated in institutions established by the rational state, which overcome individuality and aim at the universal. The person becomes free in the state which represents "concrete freedom." Hegel's ideas about freedom and the state differ from liberal doctrine; since membership in the state is not optional but substantial. The state and its institutions represent "our participation in the universal," and that participation is necessary for authentic human freedom which must take the form of rational self-determination (Hegel 1955).

One of the most basic institutions that constitutes 'ethical life' is private property. Hegel was committed to property rights from his earliest writings but his mature

thinking on property is expressed in the *Philosophy of Right* (1952). In that work Hegel explains that a person must be able to control objects in his environment, otherwise the world will remain external and alien to him. "Personality," writes Hegel (1952), "is that which struggles to lift itself above this [subjective] restriction and to give itself reality, or in other words to claim that external world as its own" (§ 39). In order to overcome this restriction the human subject requires "the right of putting his will into any and every being and making it his property" (§ 44). Hegel argued that a person must be allowed to "translate his freedom into an external sphere" and that "property is the first embodiment of freedom and so is in itself a substantive end" (§ 45). When external things such as a house embody the person's will and express his personality, they must belong to that person.

For Hegel, the person cannot be free without property, since property allows one to overcome the opposition between self and world and to freely put one's personality into external objects that exist beyond the inner self. Since the first embodiment of freedom is property, the right to property belongs to the essence of the human person. Property, therefore, is a "substantive end" because it is essential for the realization of human freedom. This property must be privately owned and controlled because common property "violates the right of personality" (§ 46). When external things embody the person's will and express his or her personality, they must belong to that person. If these personal items are exclusive to the person he or she can rely on these goods which create the scope for the person's future actions.

Hegel's theory is particularly appropriate for intellectual objects and creative works, since authors and artists typically have a strong "personality stake" in their respective works. If physical property is the "embodiment of personality," then the same can surely be said for intellectual property. As human beings freely externalize their will in various intellectual objects such as novels, works of art, or poetry, they create "property" which they are entitled to control because those intellectual products are a manifestation or incarnation of their personality. The created work becomes an extension of the author's personhood, reflecting his thoughts, ideals, and motivations. Radin (1982) sums up Hegel's property doctrine quite clearly: "the more closely connected with personhood, the stronger the entitlement" to an exclusive property right.

Some might question whether a creative work is an expression of one's personality or merely a re-assembly of existing works. If the latter were true, the Hegelian argument would lose its force. The nature of the creative act is somewhat mysterious, and we can treat it only superficially in this paper. The relevant question is whether or not creativity reflects the person's inner self and personality in some way. Zemer (2006) appears to think this is not so, claiming that "meaning-mak-

ing is not an internal process.” Rather, the creator simply absorbs and modifies external objects. Others see the creative process quite differently as one in which the author or creator “gives birth” to a work in a way that engenders a unique bond between the creator and her work (Kwall 2006). This sentiment is nicely expressed by the author of the Air Pirates cartoon: “Taking this comic strip away would be like losing my arms and legs” (Amabile 1996). Kwall (2006) has made a persuasive case that the author’s creation “is an embodiment of his intrinsically motivated message.” Many authors have talked about their deep inspiration for creating a certain work or the prolonged “gestational period” that occurs before they begin to express their ideas. Therefore, according to this framework, the external work embodies the author’s “intrinsic creative process,” which often involves key aspects of the author’s interior self such as his motivations or inspiration (Kwall 2006).

Sensitivity to the dynamics of the creative process is essential for structuring policies based on the author’s needs and ensuring his or her presumptive rights. Also, if we concentrate on this intimate relationship between the author and her work, Hegel’s conception of property assumes a special relevance. Based on the Hegelian paradigm, a compelling case can be made against significantly expanding the right to recode (i.e., beyond what is currently allowed by the fair use provision of the copyright law) and against a broad “free rider” prerogative. The expansion of recoding freedom violates the author’s right to ensure the stability of social meaning of his or her work, at least for a limited amount of time so that the work’s underlying message can be established in the mind and psyche of the public. Artistic works, for example, have a particular truth and meaning of their own. As Heidegger (1950) writes, “in a work truth is at work. . . .” At a minimum, the work bears testimony to an artist’s inspirational vision stabilized in a particular form or format. The author deserves the right to conserve the “truth-at-work” in his or creative work, at least for a limited amount of time. In some contexts, this conservation may be as essential as the creation itself. To allow someone else to misappropriate a creative work for the sake of giving it a radically new meaning in the name of free speech is to undercut authorial dignity and artistic integrity.

Authors, therefore, deserve the right to control the social meaning of their work, especially since this work is a manifestation of their creativity and identity. This right should not be perpetual, since at some point the “marketplace of ideas” should have its say in the evolution of an author’s original ideas. We may want to shorten copyright protection or broaden fair use parameters after several decades of strong protection. As Hughes (2003) points out, the eventual expansion of the right to create derivative works and the liberalization of other fair use rights should not “endanger those initial decades when the author is trying to secure a place in the culture for her work as she intended it.” At the same time, the

right to protect the integrity of a work should not be so broad that it precludes various forms of fair use such as parody. While fair use and limited term are not mandated by a strict following of the Hegelian theory, policy makers must also take into account utilitarian considerations. Society's interests in access to creative works must be weighed against the exclusive rights of authors. Similarly, the author's rights must be balanced against free speech rights of other users. As Kwall (2006) points out, "an unbounded right of integrity that would enable authors to prevent all perceived mutilations, unwarranted criticisms, and objectionable contextual uses" would raise serious free speech concerns. On the contrary, there must be prudently tailored authorial moral rights that respect free speech rights while providing authors with the opportunity to propagate and preserve the original meaning of their work for a limited period of time. While the current fair use mechanism in the United States needs some adjustment, it serves as a decent model for achieving this sort of balance.

In addition to the deontic Hegelian argument the position that authors must be allowed to control social meaning can be supported on utilitarian grounds. Stability of meaning is valuable not only to owners but also to non-owners of intellectual property. Hughes (1999) has argued that the audience benefits when a cultural object is allowed to maintain its stable meaning. Many people derive psychological and social benefits from certain, firmly established cultural icons. Indeed, the ability to recode usually depends on the "underlying stability" of the recoded object (Hughes 1999). Therefore, moderate copyright protection, which limits recoding according to the current fair use parameters, has societal benefits that cannot be overlooked.

Finally, those who favor recoding rights often fail to perceive that the current legal infrastructure already allows for a significant level of recoding. Deconstructionists and copyright critics often exaggerate the level of control over copyrighted works. Works are not hermetically sealed off completely off limits for future creators. Even if they so desired, "an intellectual property owner cannot possibly appropriate all of the information (and thus social value) generated by her creation" (Wagner 2003). Thus, thanks to the fair use doctrine delineated above and the porous nature of intellectual objects, users can borrow from existing works, engage in parody, appropriate general content such as scenes à faire, plot lines, or factual knowledge. Trademarks are also subject to exceptions for fair use and parody.

Sometimes there are border line cases such as Randall's book *The Wind done Gone*, which allegedly infringed the copyright of Margaret Mitchell's *Gone with the Wind*. Despite the allegation of infringement and the costly litigation, in the end the Appellate and Eleventh Circuit Courts got it right. Both courts concluded that *The Wind done Gone* qualified as fair use because of its "parodic character" (*Suntrust Bank v. Houghton Mifflin* 2001). Also in its favor was the transformative character of Randall's work, "for the more transformative the new work, the less

will be the significance of other factors, such as commercialism, that may weigh against a finding of fair use" (*Suntrust Bank v. Houghton Mifflin*, citing *Campbell v. Acuff-Rose Music* 1994). As the outcome of this case demonstrates, thanks to the safety valves built in to copyright law, there are already ample opportunities available for "recoding" by downstream secondary users.

Conclusion

In summary, we are sympathetic to those arguing for a change in copyright law to make remixing easier. The time has surely come to adjust the fair use provisions in order to enable more forms of amateur remixing. On the other hand, we contend that there is no basis for a broad recoding or free ridership right. Our argument has been based on the notion that property rights are necessary to protect the author's intimated personality stake embodied in his or her creative work. From this highly plausible personality theory follows a moral imperative to protect the integrity of the creator's works. This precludes most forms of recoding aside from what is allowable under copyright safety valves such as fair use. Audiences should be allowed to remix more freely but not re-code with impunity as some legal scholars would clearly prefer.

Finally what about cases where amateur remixing involves recoding? What should be done if a neo-Nazi web site mixes various offensive images to the tune of a popular song? In this case, the artist's commercial interests are unaffected, but the integrity of the work has undoubtedly been impaired by this objectionable association. Unfortunately, Lessig dismisses these cases in his zeal to promote remixing. He maintains that by loosening controls the author doesn't have to worry about misuse: "because the law allows the copyright owner to veto use, the copyright owner must worry about misuse, [but] if the owner can't control the use, then the misuse is not the owner's responsibility" (Lessig 2008). This argument, however, misses the point. Even if an author can deflect responsibility for the misuse of his work, he is deeply affected by objectionable uses that undermine the work's integrity just as he is psychologically affected by a burglar who violates his personal physical property. Thus, in my estimation, the law must be tailored in such a way that blatant and objectionable forms of recoding, even by amateur remixers, should not be allowed to stand if the authors choose to seek legal, injunctive action. Given the distinctive and intimate bond between the author and her work, the right of integrity must be given priority in these admittedly difficult cases.

REFERENCES

- Amabile, T.** (1996). *Creativity in Context*, Westview Press, Boulder, CO.
- Aoki, K.** (1996). Intellectual Property and Sovereignty: Notes toward a Cultural Geography of Authorship, 48 *Stanford Law Review* 1293.
- Barthes, R.** (1977). The Death of the Author, in *Image-Music-Text*, ed. S. Heath, Hill & Wang, New York, 142-148.
- Burrows-Giles Lithographic Co. v. Sarony* (1884). 111 U.S. 53.
- Campbell v. Acuff-Rose Music** (1994). 510 U.S. 569.
- Coombe, R.** (1996). Authorial Cartographies: Mapping Proprietary Borders in a Less-Than-Brave New World, 48 *Stanford Law Review* 1357.
- Coombe, R.** (1991). Objects of Property and Subjects to Politics: Intellectual Property Laws and Democratic Dialogue, 69 *Texas Law Review* 1853.
- Craig, C.** (2007). Reconstructing the Author-Self: Some Feminist Lessons for Copyright Law, 15 *American University Journal of Gender, Social Policy, & the Law* 207.
- Feist Publications, Inc. v. Rural Telephone Service Co.* (1991). 499 U.S. 340.
- Gantz, J. and Rochester, J.B.** (2005). *Pirates of the Digital Millennium*. Financial Times Prentice-Hall, Upper Saddle River, NJ.
- Gordon, W.** (2002). Authors, Publishers, and Public Goods: Trading Gold for Dross, 36 *Loyola of Los Angeles Law Review* 159.
- Halbert, D.** (1999). *Intellectual Property in the Information Age*, Quorum Books, Westport, CN.
- Harper & Row Publishers, Inc. v. Nation Enterprises.* (1985). 471 U.S., 539.
- Hegel, G. W. F.** (1955). *Die Vernunft in der Geschichte* ed. J. Hoffmeister, Niemeyer, Hamburg.
- Hegel, G.W.F.** (1952). *Philosophy of Right*, trans. T. Knox., Oxford University Press, Oxford.
- Heidegger, M.** (1950). *Holzwege*, Klostermann, Frankfurt..
- Hughes, J.** (2003). Fair Use across Time, 50 *UCLA Law Review* 775.
- Hughes, J.** (1999). Recoding, Intellectual Property and Overlooked Audience Interests, 77 *Texas Law Review* 923.

Kwall R. (2006). Inspiration and Innovation: The Intrinsic Dimension of the Creative Work, 81 Notre Dame Law Review 1945.

La Voi, S. (2004). The Value of Recoding within Reason, 14 DePaul University Journal of Art and Entertainment Law 171.

Lessig, L. (2008). Remix: Making Art and Commerce Thrive in the Hybrid Economy, Penguin Press, New York.

Lessig, L. (2004). Free Culture. Penguin Press, New York.

Lindon contra La Compagnie Brut de Buton (1992). 155 Revue Int'l du Droit D'Auteur (T.G.I. Paris 3e ch.).

Metro-Goldwyn-Mayer Studios v. Grokster, Ltd (2005). 125 U.S. 2764.

O' Rourke, M. (2001). Property Rights and Competition on the Internet, 16 Berkeley Technology Law Journal 561.

Radin, M. (1982). Property and Personhood, 34 Stanford Law Journal 57.

Sony Corp of America v. Universal City Studios, Inc. (1984) 464 U.S. 417.

Spinello, R. and Bottis, M. (2009). A Defense of Intellectual Property Rights, Edward Elgar Publishing, Cheltenham, UK.

Spinello, R. (2008). Intellectual Property: Legal and Moral Challenges of On-line File Sharing, in The Handbook of Information and Computer Ethics, eds. K. Himma and H. Tavani, Wiley, Hoboken, NJ, 553-570.

Suntrust Bank v. Houghton Mifflin (2001) 252 F.3d 1165, [11th Cir].

Vaidhyanathan, S. (2001). Copyrights and Copywrongs: The Rise of Intellectual Property and How it Threatens Creativity, New York University Press, New York.

Wagner, R. P. (2003). Information wants to be Free: Intellectual Property and the Mythologies of Control, 103 Columbia Law Review 995.

Zemer, L. (2006). The Copyright Moment, 43 San Diego Law Review 247.

The Digital Divide among Under-Age Individuals: An Economic and Legal Approach

Stiakakis Emmanouil*

Department of Applied Informatics,
University of Macedonia

Alexandropoulou-Egyptiadou Evgenia**

Department of Applied Informatics,
University of Macedonia

Abstract

This study aims to provide insights for a better understanding of the digital divide. The factors which contribute to the widening of the digital divide are presented and age is shown to be significantly associated with the use of Information and Communication Technologies (ICTs). The existence of a digital divide among under-age individuals is proved based on the findings of a secondary research. These findings indicate the considerable differences among Northern, Central and Southern Europe, with regard to the use of the Internet by minors. The regulations and policies to bridge the digital divide among minors are an additional aspect of this study.

Keywords: Digital Divide, Digital Inequality, Information Society, ICT

* Dr. *Stiakakis Emmanouil* is Lecturer in Digital Economics at the Department of Applied Informatics, University of Macedonia – Thessaloniki – Greece. He holds a BSc in Mechanical Engineering from the Aristotle University of Thessaloniki, an MSc in Manufacturing Systems Engineering from Cranfield University – UK, and a PhD in Applied Informatics from the University of Macedonia. His research interests include production and operations management, Total Quality Management, e-business, and digital economy. His research has been published in international journals and conference proceedings.

** Dr. *Alexandropoulou-Egyptiadou Evgenia* is Associate Professor in Computer Law at the Department of Applied Informatics, University of Macedonia – Thessaloniki – Greece. She is also attorney at law (Bar of Thessaloniki). She has been the Head of the legal department of a Greek bank (Egnatia Bank) and member of the editorial committee of the Law Review *Harmenopoulos* (edited by the Bar of Thessaloniki). She has written many scientific articles and books in the area of Civil, Banking, Labour, International and IT Law. In the last few years her interests focus on personal data protection, human rights and intellectual software property.

Introduction

Participation in the information society is considered a fundamental human right, constitutionally guaranteed in several countries, including Greece (Alexandropoulou-Egyptiadou, 2007; Iglezakis, 2008). Necessary instruments for the function of the information society are the Internet and all the other ICTs. The ability to access and use ICTs appears to be an indispensable prerequisite for participation in the information society. But on the other hand, the digital divide is a reality that exists both in the developed and developing countries. In a recent report, the World Economic Forum indicated that 88% of all Internet users are from industrialized countries that comprise only 15% of the world's population (Pick and Azari, 2008).

The digital divide is a new form of socioeconomic inequality. It is defined by the Organization for Economic Co-operation and Development (OECD) as the "gap between individuals, households, businesses, and geographical areas at different socioeconomic levels with regard both to their opportunities to access ICTs and their use of the Internet for a wide variety of activities" (2001). Debate about the digital divide started in the early 1990s with the distinction between ICT "haves" and "have-nots" (Rallet and Rochelandet 2007; Wilhelm and Thierer 2000). This topic was initially restricted to OECD countries, but now involves all countries by creating a new type of international inequality. There has been much discussion about the determination and analysis of the digital divide's components (DiMaggio et al. 2004; Dewan and Riggins 2005). While in the 90s the focus was mainly on infrastructural access, nowadays the focus is moving to the users and the factors that generate digital inequality. This term refers to socio-economic disparities inside the "online population", such as the quality and the cost of the connection to the Internet, the skills and the knowledge to find the required information, e.tc. The primary issue nowadays is not whether there is an Internet access but what people are able to do when they have access to the Internet. DiMaggio and Hargittai (2001) disassociated the inequality of access from digital inequality, while Attewell (2001) refers to this distinction as the "first-level" and "second-level" digital divide. Emphasis in both approaches is given on single factor analysis rather than integrative frameworks and measurements (Barzilai-Nahon 2006). Despite the increasing attention and the plethora of academic literature on the digital divide, there has been no consensus on its causes and how to properly bridge it. In fact, little research has been performed on causation of technological differences among countries, and what empirical lessons can be learned to support leaders and policy makers in reducing the digital divide.

Factors Contributing to the Digital Divide

Measuring the digital divide is complicated since it is apparently composed of multiple dimensions. Let us mention some of these dimensions. According to James (2005), the traditional measurement of the digital divide in terms of direct access to ICTs makes sense only in the developed countries, but not in low-income countries, where people can have indirect access to the Internet through various intermediaries (e.g. listening to local radio stations which get information via the Internet). Guillen and Suarez (2005) examine the relationship of democracy and the Internet and argue that the more democratic the polity, the greater the Internet use. Quibria et al. (2003) point out that income, education and infrastructure are the most important variables with regard to the use of six kinds of ICTs. Liu and San (2006) found that the factors associated with the rapid diffusion of the Internet at the turn of the 21st century are lower cost of access, greater societal openness, political stability, literacy, urbanization, and utilization of TV sets. Noteworthy is the approach of Dasgupta et al. (2005), which emphasizes that long-term infrastructure may be more important than access to ICTs, especially for developing nations. An earlier finding, of great interest to our work, is that primary education has higher returns on investment than secondary and higher education, and this finding is even more valid for developing countries (Psacharopoulos, 1994). However, specific ICT factors were not included in this study.

So far there has not been a worldwide research concerning the factors which contribute to the digital divide. Besides, the factors widening the digital divide are not the same for each country in the world. This is why we adopted in our analysis the factors identified in the U.S.A. as the most important for the expansion of the digital divide. These factors concerning individuals, since they are different for businesses, are the income, age, education, race, household type, and geographical location (Sipior et al., 2002). Each of these is discussed in the following lines.

- **Income:** ICT use rises as income rises. Households with incomes of over \$75,000 (early 2000s) are more than 20 times more likely to have access to the Internet than those at the lowest income level. Only 23% of low-income children have access to the Internet at home compared to 58% of children in high income families (Wilhelm and Thierer, 2000).
- **Age:** the most sensitive age groups with regard to online penetration are children up to twelve and seniors aged over 65 years. On the contrary, the age group with the highest online penetration comprises individuals between 35 and 44 years old. Internet usage among children rises with a considerably

increasing rate; however, there are still significant differences among the European countries, which are actually the subject of this paper.

- **Education:** the differences in the level of education are in agreement with income disparities because most people who are poor are uneducated. Individuals with a university or college degree are more than 9 times more likely to use the Internet than those with an elementary school education. The dramatic difference in ICT use is due to literacy since the vast majority of online content is text-based.
- **Race:** an important factor widening the digital divide, although it is not as important as the previous three factors. White individuals use the Internet much more than Blacks or Hispanics. According to the data from U.S.A., the gap in Internet access between White and Black households and White and Hispanic households actually increased in the early 2000s compared to 1990s.
- **Household type:** married couples with children less than eighteen years old have the highest Internet penetration, while female-headed households with children the lowest. This factor is associated with an increase in income, since it makes sense to suppose that married couples have higher income levels, taking for granted that both parents work.
- **Geographical location:** the location of residence within a country and the location of a country within a geographical region also affect ICT adoption. Urban areas tend to have better infrastructure and lower costs of ICT use in contrast with rural ones. Providing high-speed Internet access through traditional means to areas with a small population base is still expensive.

Another factor that has been examined a lot is gender. The role of gender in accessing and using ICTs is often examined in parallel with age. However, most studies show that contrary to what happens with age, gender contributes less to the digital divide widening. The differences in ICT use, regarding gender, are mainly illustrated between older men (over 50) and younger women (up to 45). Men over 50 years old are more likely to use the Internet than women of the same age, while women up to 40s make a greater use when compared with men of the same age (National Telecommunication and Information Administration, 2002).

Under-Age Individuals and ICTs

It is unquestionable that the digital divide has to be confronted. Bridging of the digital divide must be achieved by enabling the conditions for everyone to take part in the information society, i.e. e-inclusion (Commission of the European Communities, 2007). As we already noted, the examination of the digital di-

vide nowadays revolves around the types of users. Following this, our analysis focuses on a very vulnerable social group, under-age individuals. This is due to the following reasons: minors are the prospective users of the next generation of digital technologies, as well as they constitute a social group which is affected a lot by many external forces, such as family, school environment, friendships, e.tc. It is true that most parents did not grow up using ICTs, so they do not know how to teach their children to use computers responsibly and effectively. In addition, in most countries, schools do not contribute to a deep understanding of new technologies. Consequently, a number of minors are able to exploit ICTs at a high rate, whereas others use them in an inappropriate way (stealing, for instance, music, movies or software through peer-to-peer networks or launching a destructive Internet virus), and finally some others fear the new technologies. E-inclusion is especially significant for minors since they will be the citizens of the future worldwide information society. The appropriate education of minors (and their parents and teachers as well) in order to access and appropriately use ICTs, and the adequate legal environment are some of the necessary means towards e-inclusion.

As previously noted, the scope of this paper is to demonstrate the digital divide among minors, as well as the steps that have been taken towards the direction of encouraging their e-inclusion. We present data with regard to the use of ICTs by under-age individuals and examine the most important factors that contribute to the different usage rates. Analyzing these factors is a prerequisite to making suggestions on how to bridge this digital divide. These suggestions will include specific measures in order to fight e-exclusion of minors, taking into consideration the necessary precautions for the secure accessibility and use of ICTs.

The approach followed in this paper is an adaptation of the model presented by Korupp and Szydlik (2005), taking into account the particular needs of under-age individuals. Three main factors determine the digital divide among under-age individuals: human capital, the family context and the social context. Human capital is associated with the general education and the individual performance on specific computer courses. There is a positive relation between a person's human capital and his or her private use of computers or the Internet. Getting acquainted with computers at school increases an under-age individual's likelihood of using a computer or the Internet for his or her private needs. The family context also plays a significant role concerning ICT penetration to minors. It is a fact that, the home environment where people are raised determines their habits towards new technologies. The family composition, income or principles are only some of the many components of the family context, which specify the limits of ICT usage by minors. Parents sometimes do not make the right decisions, as for example when they forbid their children to have access to the Internet trying to protect

them from unwanted information. The best way of doing so would be to control the access to unwanted sites. The third factor explaining the use of a computer or the Internet by an under-age individual is the social context. Friendships, groups where minors belong to, as well as the school environment are basic determinants of the social context. We should point out that, as school environment we mean the social contacts and any kind of interaction at school, since the level of education is associated with the human capital. Giving a broader meaning to the social context, determinants such as the generation, gender, ethnicity, and region should also be included. While generation is not so important in our analysis, since all minors have more or less the same characteristics, each one of the other three determinants has its own part in the widening of the digital divide among minors.

Computer use for an under-age individual mainly involves playing games interactively or alone, and with regard to the Internet searching for information. Recent surveys on the use of ICTs by children between 6 and 13 years old show that 70% play games on computers, 50% use computers for their homework, 44% use them for learning programs, and 33% draw and write on the computer or surf on the Internet (Medienpädagogischer Forschungsverbund Südwest, 2003). Respective surveys on teenagers (12-18 years old) reveal that 41% play computer games, 44% use the Internet for e-mail, 26% use it for instant messaging, 22% for educational and occupational purposes, 31% seek for information, 23% download news or music, and 36% in order to do their schoolwork (Medienpädagogischer Forschungsverbund Südwest, 2004). Comparing the results of these surveys, which were conducted approximately the same time period, it is apparent that younger children devote mostly their time in entertainment, while teenagers have several common online activities with adults.

Regulations and Policies to Bridge the Digital Divide Among Minors

At EU level, the importance of e-Inclusion was recognised in i2010 and Member States, the European Commission, industry, and NGOs representing users have undertaken several actions to advance e-Inclusion. A landmark was the 2006 Ministerial "Riga Declaration" on ICT for an inclusive information society, which set concrete targets for Internet usage and availability, digital literacy, and accessibility of ICT by 2010, preparing the ground for a legislative framework, common all over the European Union (EU). However, despite all these valuable initiatives, progress is still lacking and most of Riga targets may not be achieved. Fragmentation of efforts and lack of collaboration continue to persist (Commission of the European Communities, 2007). An e-Inclusion Ministerial Conference took place from 30 November to 2 December 2008 in Vienna, in order to acceler-

ate progress towards digital inclusion and provide a forum to discuss policies to achieve an inclusive, barrier-free, information society and the related social and economic benefits.

Participation in the information society is a human right, constitutionally guaranteed in several countries, including Greece. The right to e-participation includes the aspect of e-inclusion. The imperative to promote e-inclusion is enshrined in Greek Constitution (Revision of 2001), which in the article 5A par. 2 provides for a social right to e-participation. The above mentioned article states that "All persons are entitled to participate in the Information Society. Facilitation of access to electronically circulating information, as well as of production, exchange and diffusion thereof constitutes an obligation of the State".

According to the above mentioned article, the Greek State is obliged to take positive actions in order to make equal and active access to the Information Society possible for all (Iglezakis, 2008; Mitrou, 2006; Papakonstantinou, 2006; Kontiadis, 2002) and the benefits from e-Government and e-Democracy. Greece has not developed a particular strategy for digital inclusion, but it introduced specific related projects, mainly in telemedicine, homecare e.tc. It also launched initiatives concerning the development of ADSL and Wi-Fi networks, the subsidy of acquirement of technological equipment by professionals, students e.tc. (Iglezakis, 2007). Although there is no specific rule related to e-inclusion, the article 5A of Greek Constitution constitutes a legal basis for policies against digital divide.

E-Inclusion is especially significant for minors, since they will be the citizens of the future worldwide Information Society. The efforts of the legislator should focus on the appropriate information and education of minors, in order to access and use ICTs. Special interest is given to those living in rural areas, having disabilities or belonging in ethnic minorities. Abilities to approach ICT technologies are also offered to their parents and teachers through special training programs. Additional means towards e-Inclusion include programs to subsidise electronic equipment and access to the Internet for minors. Significant factor to the achievement of e-inclusion guarantees safe access and use of the Internet to minors. In this particular area, EU and its member states have done noteworthy steps.

At EU level, the European Schoolnet (2009) is an organization, established in 1996, which aims to the support of teaching and learning in European schools, mostly through ICTs. This organization aims to: a) build one rich, multi-language European community for the innovation and cooperation in educational politics and practice, b) constitute a European gate amongst national and regional educational networks and the facilities they offer, c) encourage technical innovation, interoperability and common standards and, finally, d) create a strong and effi-

cient European organization that offers synergy and added value to a networked educational world.

In Greece, the Hellenic Ministry of National Education and Religious Affairs (MoE) has launched a coordinated effort for the utilization of ICTs and their incorporation into the daily educational procedure. This effort is implemented in the fields of the third Community Support Framework mainly from the Operational Program of Information Society, under the direct supervision of the General Secretary of the Hellenic MoE and with the support of the Hellenic MoE Information Society office and the «Strategy for ICTs in Education» Committee. It is constructed onto four lines of action: a) installation and support of network and computational equipment, b) development of software and digital content for educational and administrative purposes (educational software, information systems, Internet content e.tc.), c) training of the educational community on ICTs, targeting to the utilization of the above areas, d) modernization of administration areas. This initiative aims to: a) the incorporation of ICTs in the teaching process, b) the support of the Informatics lesson taught in high schools, senior high schools and technical schools, c) the support of every cognitive area through the use of ICTs, d) the elimination of digital illiteracy and variations on ICT skills (Ministry of National Education and Religious Affairs, 2009). In addition, the Information Society Office of the MoE deals with the exploitation of ICT supportive technologies in special needs education, in particular for pupils with hearing, visual, kinetic or intellectual disabilities.

Other significant policies of the Greek government towards e-inclusion are the following Actions with the contribution of EU: a) the Action “parents.gr” (2009), which offers to parents of pupils (secondary education) free computing lessons, as well as subsidised connection to the Internet for two months. By the end of March 2009, 43.000 parents have participated to the program. b) the Action “see it digitally” (2006-2008). This offers subsidies for laptops amounting to 80% of their market value to students entering Universities and Technological Institutions achieving high grades, or suffering from serious illnesses. By the end of March 2009, 15.871 students have participated to the program, c) according to a ministerial decision (153.604/ΨΣ11182–B/7.10.2008), subsidized training programs in the area of ICTs are offered to students entering Universities and Technological Institutions achieving high grades, or suffering from serious illnesses.

Research findings

A secondary research was undertaken to gain insight into the digital divide among under-age individuals within the EU and more importantly, to contribute to narrowing the divide. Concerning the methodological part of this research, the EU was divided into three groups: Northern Europe, Central Europe and Southern Europe. The variable measured was the percentage of under-age individuals who used the Internet on a regular basis (in order to avoid the random use of the medium) in the years 2006, 2007, and 2008. For the purposes of our research, the most recent data available at the time of writing have been chosen, published by Eurostat (2009). It should be mentioned that, there were no data available for United Kingdom, Germany, France, Belgium, and Luxembourg. The research findings are presented in Figures 1, 2, and 3. In each figure, the last two columns represent the Internet use in the EU of 15 countries (EU-15) and the EU of 27 countries (EU-27), allowing the readers to compare the values of each region's countries with the average values of EU. The mean value in EU-15 for the three years period is 89.6% and in EU-27 87.6%.

It can be seen that there are significant differences among the three regions. In Northern Europe, most percentages of Internet use reach approximately 100%. The only exception in this finding is Ireland. It is true that, people in north European countries are much more familiar with the use of ICTs. This is mostly due to the high level of education and training, the development of telecommunications infrastructure, as well as the mentality, particularly of younger people, to fully exploit the tremendous possibilities of the Internet. In Central Europe, the percentages are lower compared to Northern Europe, since the highest values range about 90%. Slovakia seems to have the highest Internet use, while Romania the lowest. The situation in Southern Europe is completely different. According to the findings, under-age individuals in south European countries use the Internet to a much lower extent. This is particularly obvious in Bulgaria, Cyprus, Greece and Italy, while Malta, Portugal and Spain have higher percentages. The mean value in Greece for this time period is 78%. There are many causes for this inequality in Southern Europe. Noteworthy are the higher cost of Internet access and the much lower family income, although the finding for Cyprus was quite surprising.

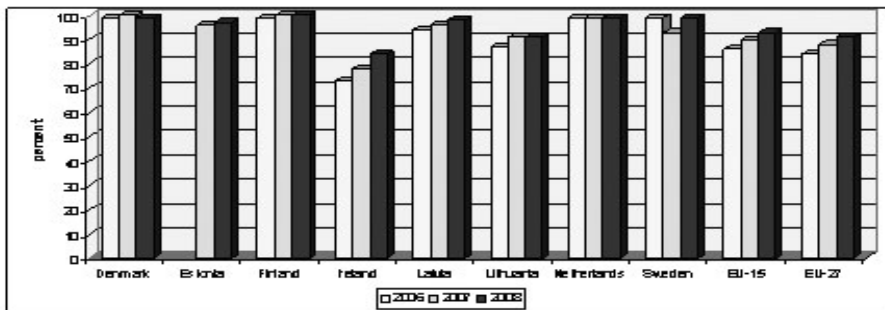


Figure 1: Internet use by under-age individuals in Northern Europe

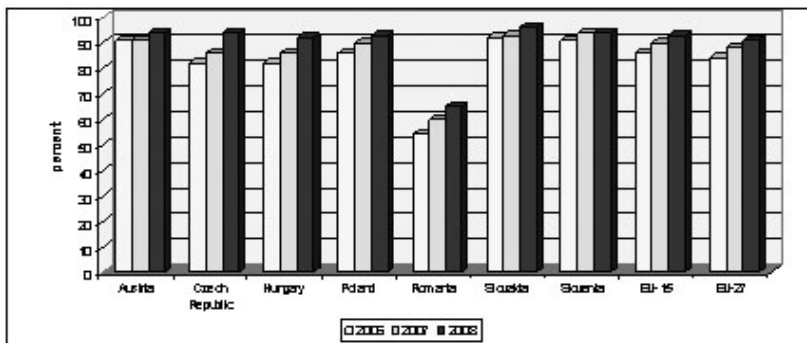


Figure 2: Internet use by under-age individuals in Central Europe

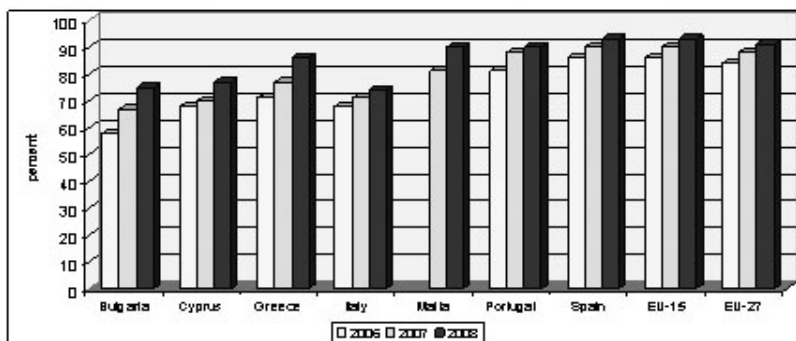


Figure 3: Internet use by under-age individuals in Southern Europe

Conclusions

According to the findings of the secondary research (described in the previous section) there are significant differences among Northern, Central and Southern Europe, with regard to the use of the Internet by minors. These findings give us evidence of the existence of a digital divide among minors. Under-age individuals use the Internet to a great extent, much more than seniors, however this extent is contingent upon certain factors, such as the geographical region of the country where they live. Determining and analyzing such factors could be a suggestion for further research, since this study deals only with the digital divide among minors, based on the geographical location of their countries.

It could be said that the digital divide is a generational phenomenon that will disappear in time. If so, age is not a relevant variable. The next generation will have spent more of their lives surrounded by computers, improving skills and gaining confidence in the use of ICTs. However, we need to be cautious in this perspective. There is evidence that the digital divide changes to a different kind of inequality, where the primary issue will be not whether there is an Internet access but what people are able to do when they have access to the Internet. Differences in the ways people use ICTs will remain and may increase, as the rate of emergence of innovations and new applications of digital technologies increases.

It is obvious that there is no single solution to bridge the digital divide. It is not just a problem of individuals' choices of having or not having connections to the Internet, nor is it the economic affordability of Internet services. It is an issue of how central the Internet is in individuals' everyday lives and particularly the habits and interests of young people. It is an issue of whether business and government leaders will recognize the importance of bringing everyone and particularly minors onto the information society. It is also an issue of whether and to what extent the society and every one of us will deal with the problem of differences in ICT usage based on age, as well as ethnicity, gender, geography or cultural preferences.

REFERENCES

- Alexandropoulou-Egyptiadou, E.**, (2007). Personal Data. Sakkoulas, Athens-Komotini.
- Attewell, P. A.**, (2001). The first and second digital divides, *Sociology of Education*, 74, 252 – 259.
- Barzilai-Nahon, K.**, (2006). Gaps and bits: Conceptualizing measurements for digital divide/s, *The Information Society*, Vol.22, No. 5, 269 – 278.
- Commission of the European Communities, (2007). Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. European i2010 initiative on e-Inclusion, Brussels.
- Dasgupta, S., Lall, S. and Wheeler, D.**, (2005). Policy reform, economic growth and the digital divide, *Oxford Development Studies*, Vol.33, No. 2, 229 – 243.
- Dewan, S. and Riggins, F. J.**, (2005). The digital divide: Current and future research directions, *Journal of Association for Information Systems*, Vol.6, No. 2, 298 – 337.
- DiMaggio, P., Hargittai, E., Celeste, C. and Shafer, S.**, (2004). Digital inequality: From unequal access to differentiated use, in *Social Inequality* (ed. K. Neckerman), Russell Sage Foundation, New York.
- DiMaggio, P. and Hargittai, E.**, (2001). From the 'digital divide' to 'digital inequality': Studying Internet use as penetration increases, Center for Arts and Cultural Policy Studies, Working Paper #15, Princeton University.
- European Schoolnet, (2009). <http://www.eun.org/portal/index.htm>. Eurostat, (2009). <http://epp.eurostat.ec.europa.eu/portal/page/portal/sdi/indicators/theme3>.
- Guillen, M. E. and Suarez, S. I.**, (2005). Explaining the global digital divide: Economic, political and sociological drivers of cross-national Internet use, *Social Forces*, Vol.84, No. 2, 681 – 708.
- Iglezakis, I.**, (2008). The development of e-Governance and the issue of digital inclusion in Greece with particular regard to the constitutional right of e-Participation, in *e-Government and Digital Inclusion* (ed. C. Maioli), Prensas Universitarias de Zaragoza, Zaragoza.
- James, J.**, (2005). The global digital divide in the Internet: Developed countries constructs and Third World realities, *Journal of Information Science*, Vol.31, No. 2, 114 – 123.

Kontiadis, X., (2002). The new constitutionalism and fundamental rights after the Constitutional Revision of 2001. *Ant.N. Sakkoulas*, Athens-Komotini.

Korupp, S. E. and Szydlík, M., (2005). Causes and trends of the digital divide, *European Sociological Review*, Vol.21, No. 4, 409 – 422.

Liu, M. C. and San, G., (2006). Social learning and digital divides: A case study of Internet technology diffusion, *Kyklos*, Vol.59, No. 2, 307 – 321.

Medienpädagogischer Forschungsverbund Südwest, (2003). Kinder und medien, computer und Internet, in *Causes and trends of the digital divide* (S. E. Korupp and M. Szydlík), *European Sociological Review*, Vol.21, No. 4, 409 – 422.

Medienpädagogischer Forschungsverbund Südwest, (2004). Jugend, information und (multi-) media, in *Causes and trends of the digital divide* (S. E. Korupp and M. Szydlík), *European Sociological Review*, Vol.21, No. 4, 409 – 422.

Ministry of National Education and Religious Affairs, (2009). http://www.ypepth.gr/ktp/en_ktp_sne.htm.

Mitrou, L., (2006). The right of participation in the Information Society. *Sakkoulas*, Athens.

National Telecommunication and Information Administration, (2002). A nation online: How Americans are expanding their use of the Internet, U.S. Department of Commerce, Washington, D.C. OECD, (2001). Understanding the Digital Divide, <http://www.oecd.org/dataoecd/38/57/1888451.pdf>.

Papakonstantinou, A., (2006). The constitutional right of participation in the Information Society, *Revue of Public and Administrative Law*, 2, 233 – 242.

Pick, J. B. and Azari, R., (2008). Global digital divide: Influence of socioeconomic, governmental, and accessibility factors on information technology, *Information Technology for Development*, Vol.14, No. 2, 91 – 115.

Psacharopoulos, G., (1994). Returns to education: A global update, *World Development*, Vol.22, No. 9, 1325 – 1343.

Quibria, M. G., Ahmed, S. N., Tschang, T. and Reyes-Macasaquit, M. L., (2003). Digital divide: Determinants and policies with special reference to Asia, *Journal of Asian Economics*, 13, 811 – 825.

Rallet, A. and Rochelandet, F., (2007). ICTs and inequalities: The digital divide, in *Internet and Digital Economics: Principles, Methods and Applications* (eds. E. Brousseau and N. Curien), Cambridge University Press, New York.

Sipior, J. C., Ward, B. T. and Marzec, J. Z., (2002). An initiative to narrow the digital divide: Preliminary results, <http://is2.lse.ac.uk/asp/aspecis/20020055.pdf>.

Wilhelm, A. G. and Thierer, A. D., (2000). Should Americans be concerned about the digital divide?, *Insight on the News*, in the *Washington Times*, Vol.16, No. 33, 40 – 43.

Situational Crime Prevention and Insider Threat: Countermeasures and Ethical Considerations

Marianthi Theoharidou* & Dimitris Gritzalis**

Information Security and Critical Infrastructure
Protection Research Group
Dept. of Informatics, Athens University
of Economics & Business, Greece

Abstract

Insider threat is a threat posed to information systems. In this paper we examine how the theory of Situational Crime Prevention can contribute in mitigating computer abuse by insiders. We present the theory, which focuses in removing opportunities for criminal activity, and examine whether and how its techniques can be applied on an Information System or Critical Infrastructure. Thus, we create a set of security countermeasures that correspond to the techniques of the theory, so as to achieve its goals, namely: (a) to increase the effort required to attempt an

* *Marianthi Theoharidou* received a B.Sc. in Informatics, in 2002, and a M.Sc. in Information Systems, in 2004, both acquired from Athens University of Economics and Business, Greece. She is currently performing her Ph.D. research in the Informatics Department of Athens University of Economics and Business, Greece. Her published work includes several scientific articles in refereed journals and conferences. Her research interests are currently information systems security management, risk assessment and management, Information Security Education and Critical Infrastructure Protection. She has been involved in several R&D and consulting projects funded by the European Commission, public administration and private institutions.

* *Dr. Dimitris Gritzalis* (dgrit@aueb.gr) is a Professor of Security and the Director of the Information Security and Critical Infrastructure Protection Research (www.cis.aueb.gr) with the Dept. of Informatics of the Athens University of Economics and Business, Greece. He holds a B.Sc. (Mathematics, Univ. of Patras), a M.Sc. (Computer Science, City University of New York), and a Ph.D. (Critical Information Systems and Infrastructures Security, Univ. of the Aegean). He has served as Associate Commissioner of the Greek Data Protection Commission, and as the President of the Greek Computer Society. He is the representative of Greece to IFIP TC-11. For more than twenty years he has been participating in numerous R&D and consulting projects funded by the European Commission, international organizations (EUROPOL, CEN), public administration (Ministries of: the State, Government, Economy, Healthcare, Public Order, Education, Social Security, etc.), private institutions (Banking, Transportation, Telecommunication, Health, Energy, etc.). His technical publications include seven books and more than 120 papers. His current research interests focus on security in ambient intelligence, new security paradigms, critical infrastructure protection, and strategies for security-critical infrastructures.

abuse, and (b) the risk of computer abuse, (c) to reduce the expected reward, and (d) possible provocations, as well as (e) to remove excuses. We, then, discuss the social and ethical impact of these security countermeasures, highlight points of interest and suggest how to deal with these issues.

Keywords: Insider Threat, Situational Crime Prevention, Information Security, Ethics.

Insider threat

Recent reports, such as the CSI Computer Crime & Security Survey for 2008 (Richardson, 2008), indicate that the insider threat does not rely only on the volume of incidents, but also on the fact that some insiders are particularly well-placed to cause significant impact on an organization. Insider crime cases, such as leaking customer data, often remain undetected, which means that no direct costs can be associated with the theft. This fact is also supported by the 2008 Information Security Breaches Survey (BERR, 2008), which surveys UK organizations. Nearly 2/3 of the worst incidents appear to have an internal cause in this particular survey. Both surveys rank insider abuse or misuse as one of the most frequent incidents, especially in very large companies. This issue is receiving attention not only by businesses, but at a governmental level, as well, when Critical Infrastructures are at stake (Noonan and Archuleta, 2008).

Analysis of the literature on insider threat has shown that the tools and methods applied draw upon Criminology Theories. Concepts such as computer crime, computer abuse/misuse, deterrence, motives, opportunity and so on are widely used in research of insider threat. Examples of theories that are explored in the computer field are, among others, General Deterrence Theory, Social Bond Theory, Social Learning Theory, Theory of Planned Behavior or Situational and Crime Prevention (Theoharidou et al., 2005).

One modern Criminology Theory, which has been applied in numerous fields, is Situational Crime Prevention (Clarke, 1980, 1997, 2005). In this paper, we briefly present its main concepts and examine how can be applied on an Information System. We expand the countermeasures set proposed by Willison (2006a), and enrich its analysis, with techniques proposed by the CRAMM Risk Analysis and Management Methodology (Section 3). The theory has received critiques over the years on social and ethical issues. In Section 4, we present the most common arguments regarding Ethics, and discuss how applying such a theory in the computer world may raise ethical issues. Then, we provide the reader with an insight on how it would be wise to deal with these issues. We conclude with our findings and some ideas for further research.

Situational crime prevention

The theory of Situational Crime Prevention (Clarke, 1980) is based on the hypothesis that to commit a crime, a person must have both motive and opportunity. It views both factors as equally important, but it differentiates from other theories (e.g. Social Learning Theory, Social Bond Theory (Theoharidou et al, 2005)), as it does not focus solely on motives, but tries to affect the environment of an offender and reduce the available opportunities which are necessary for a crime to take place. The theory is based on the principles of Routine Activity Theory, Rational Choice Theory and Crime Pattern Theory (Clarke, 2005). Routine Activity Theory (Clarke, 1980; Willison, 2001, 2004) focuses on the characteristics of the crime and not on the actual person committing the criminal act; it explores the mechanism through which social changes in the numbers of 'suitable targets' for a crime, or in the numbers of 'capable guardians' against crime can increase or reduce a crime rate. The Rational Choice Perspective (Clarke, 1980), on the other hand, tries to explain crime from the perspective of the offender. It focuses on the thinking process of an offender, how (s)he evaluates criminal opportunities and how (s)he reaches the decision of committing a crime or not. Finally, Crime Pattern Theory seeks to explain how offenders seek or stumble across opportunities for crime in the course of their everyday life (Clarke, 2005).

The theory of Situational Crime Prevention utilizes the above ideas and forms a set of countermeasures for each case, which aims to reduce the criminal opportunities in a certain context. These measures address issues concerning the formation, management or change of the environment and their goal is five-fold: (a) making a criminal act appear more difficult, by increasing the effort required, (b) making a criminal act appear more risky, by increasing the possibility of detection, (c) reducing the expected benefits of an act, (d) reducing provocations that may trigger an offender, and (e) removing the excuses a person can make in order to justify the criminal act. The theory has been applied in many cases (Clarke, 1997) and various contexts. Examples include application of the theory in shops (e.g. CCTV, RFID), private housing (e.g. burglar alarms), street offences (e.g. city guards, neighborhood watch, e.tc.). To comprehend the concepts of the theory and its application, examples of measures reducing opportunity are presented in Table 1.

Table 2: 25 Techniques of Situational Crime Prevention (Clarke, 2005)

Increase Effort	Increase Risks	Reduce Rewards	Reduce Provocation	Remove Excuses
1. Target harden: <ul style="list-style-type: none"> • Steering column locks and immobilizers • Anti-robbery screens • Tamper-proof packaging 	6. Extend guardianship: <ul style="list-style-type: none"> • Take routine precautions: go out in group at night, leave signs of occupancy, carry phone • “Cocoon” neighborhood watch 	11. Conceal targets: <ul style="list-style-type: none"> • Gender-neutral phone directories • Unmarked bullion trucks 	16. Reduce frustrations and stress: <ul style="list-style-type: none"> • Efficient queues and polite service • Expanded seating 	21. Set rules: <ul style="list-style-type: none"> • Rental agreements • Harassment codes • Hotel registration
2. Control access to facilities: <ul style="list-style-type: none"> • Entry phones • Electronic card access • Baggage screening 	7. Assist natural surveillance: <ul style="list-style-type: none"> • Improved street lighting • Defensible space design • Support whistleblowers 	12. Remove targets: <ul style="list-style-type: none"> • Removable car radio • Women’s refuges • Pre-paid cards for pay phone 	17. Avoid disputes: <ul style="list-style-type: none"> • Separate enclosures for rival soccer fans • Reduce crowding in pubs • Fixed cab fares 	22. Post instructions: <ul style="list-style-type: none"> • “No Parking” • “Private Property” • “Extinguish camp fires”
3. Screen exits: <ul style="list-style-type: none"> • Ticket needed for exit • Export documents • Electronic merchandise tags 	8. Reduce anonymity: <ul style="list-style-type: none"> • Taxi driver IDs • “How’s my driving?” decals • School uniforms 	13. Identify property: <ul style="list-style-type: none"> • Property making • Vehicle licensing and parts marking • Cattle branding 	18. Reduce emotional arousal: <ul style="list-style-type: none"> • Controls on violent pornography • Enforce good behavior on soccer field 	23. Alert conscience: <ul style="list-style-type: none"> • Roadside speed display boards • Signatures for customs declarations
4. Deflect offenders: <ul style="list-style-type: none"> • Street closures • Separate bathrooms for women • Disperse pubs 	9. Utilize place managers: <ul style="list-style-type: none"> • CCTV for double-deck buses • Two clerks for convenience stores • Reward vigilance 	14. Disrupt markets: <ul style="list-style-type: none"> • Monitor pawn shops • Controls on classified ads • License street vendors 	19. Neutralize peer pressure: <ul style="list-style-type: none"> • Idiots drink and drive” • “It’s ok to say No” • Disperse troublemakers at school 	24. Assist compliance: <ul style="list-style-type: none"> • Easy library checkout • Public lavatories • Litter bins

5. Control tools/weapons: <ul style="list-style-type: none"> • “Smart” guns • Disabling stolen cell phones • Restrict spray paint sales to juveniles 	10. Strengthen formal surveillance: <ul style="list-style-type: none"> • Red light cameras • Burglar alarms • Security guards 	15. Deny benefits: <ul style="list-style-type: none"> • Ink merchandise tags • Graffiti cleaning • Speed humps 	20. Discourage imitation: <ul style="list-style-type: none"> • Rapid repair of vandalism • V-chips in TVs • Censor details of modus operandi 	25. Control drugs and alcohol: <ul style="list-style-type: none"> • Breath analyzers in pubs • Servers intervention • Alcohol-free events
---	--	---	---	--

One can observe that the techniques proposed are selected in order to mitigate various types of crimes (e.g. vandalism, robbery, theft, e.tc.), and in various contexts (e.g. pubs, football fields, stores, houses, e.tc.). The thing that is uniform for all the techniques, regardless of the goal, is that their aim is focused on changing the physical or social surroundings of an offender. Also, the fourth category is one of the more recent additions of the theory (Cornish and Clarke, 2003; Clarke, 2005), as it does not appear in the initial publications.

Computer abuse by insiders

Willison (2001, 2004) suggests that the theory of Situational Crime Prevention and its core concepts can be applied in the Information System Security Management field, providing a theoretical basis for understanding and addressing the issue of computer abuse within organizations. He observes that some of the techniques are already implicitly used in the organizational context (Willison, 2006b). Examples include property marking (to identify property), clear desk/screen policies (to remove targets), anti-virus detection (to target harden), and firewalls (to screen exits). He presents an initial attempt to classify some of the best practice safeguards cited in the ISO/IEC 27002:2005 standard (ISO, 2005) according to the 25 techniques of the theory (Willison, 2006a).

We create an analogy to the techniques used in other fields through three steps: (a) we re-examined the proposed classification by (Willison, 2006a) and made adjustments and corrections, (b) we enriched it with ideas and concepts found while examining the countermeasure database of the CRAMM methodology, which is a widely used risk analysis and management methodology, and (c) we tried to add measures that were missing and are analogous to the measures that have been applied in other settings. Finally, we made sure that all countermeasures are related to insider abuse, in its wider form, which includes on the insider set, anyone that has access, meaning both employees and visitors. The resulting merged set of security measures is presented in Tables 2-6:

Table 2: Techniques that increase effort for computer abuse

Goal	Technique	Security Measures
Increase Effort	Target harden	<ul style="list-style-type: none"> • Malicious software protection • Physical locks and restrains for critical equipment and media • I/O Controls • Sensitive System Isolation
	Control access	<p><u>Physical:</u></p> <ul style="list-style-type: none"> • Card/token for access • Physical locks for doors • Reception desk and security guards (at entry) • Visitor tags/cards <p><u>Logical:</u></p> <ul style="list-style-type: none"> • Authentication techniques (Password, smart card, token) • Intrusion detection systems • Strong remote authentication
	Screen exits	<p><u>Physical:</u></p> <ul style="list-style-type: none"> • Security guards and reception desks • Visitor tags/cards • Accountability for assets that exit the premises <p><u>Logical:</u></p> <ul style="list-style-type: none"> • Firewalls
	Deflect of-fenders	<ul style="list-style-type: none"> • Segregation of duties • Personnel screening • Key splitting (two or more individuals need to collaborate for access)
	Control Tools	<ul style="list-style-type: none"> • Authentication systems • Download control and Mobile Code Protection • Web access controls • Access removal for ex-employees • Removal of administrative rights • Restricted use of devices (i.e. USB tokens, wireless access) • Need-to-know access to information

Table 3: Techniques that increase the risk of computer abuse

Goal	Technique	Security Measures
Increase Risks	Extend guardianship	<ul style="list-style-type: none"> • Escorting of visitors. • Supervision of staff in secure areas • Guardianship of mobile facilities outside offices
	Assist natural surveillance	<ul style="list-style-type: none"> • Open plan offices • Incident reporting mechanism
	Reduce anonymity	<ul style="list-style-type: none"> • ID tags for staff and visitors • Audit trails • Event logging
	Utilize place managers	<ul style="list-style-type: none"> • Management supervision • Two person sign-off • Monitoring by system administrators
	Strengthen formal surveillance	<ul style="list-style-type: none"> • Intrusion detection systems • Security guards • CCTV in areas with sensitive equipment or information • Alarms (both physical and logical)

Table 4: Techniques that reduce the reward of computer abuse

Goal	Technique	Security Measures
Reduce Rewards	Conceal targets	<ul style="list-style-type: none"> • Minimize information about location of critical equipment or offices • Conceal use of PCs when travelling • Reduce website details • Minimize information on login application screens • Use of logical decoys • Remove any rank or status information on authentication IDs
	Remove targets	<ul style="list-style-type: none"> • Clear desk policy • Workstation Time-out/Password Protected Screen Savers • Paper shredders • Secure disposal of old PCs and Media • Regulate use of USB devices or other media • Thresholds on access to resources
	Identify property	<ul style="list-style-type: none"> • Property marking • Digital signatures • Copyright protection • Data Labeling
	Disrupt markets	<ul style="list-style-type: none"> • Intellectual Right Protection • Freeware, open source programs
	Deny benefits	<ul style="list-style-type: none"> • Encryption • Property marking • Software dongles • Use of multiple hardware or storage media (backup) • Business Continuity Planning • Insurance • Effective/timely incident handling, crisis management

Table 5: Techniques that reduce provocation towards computer abuse

Goal	Technique	Security Measures
Reduce Provocation	Reduce frustrations and stress	<ul style="list-style-type: none"> • Pleasant working environment • Recreational activities • Breaks
	Avoid disputes	N/A
	Reduce emotional arousal	N/A
	Neutralize peer pressure	<ul style="list-style-type: none"> • Disciplinary processes
	Discourage imitation	<ul style="list-style-type: none"> • Rapid repair for web defacement • Prompt software patching • Enforcing security policy on incidents or disciplinary procedures

Table 6: Techniques that remove excuse for the computer abuser

Goal	Technique	Security Measures
Remove Excuses	Set rules	<ul style="list-style-type: none"> • Information System security policy • Disciplinary procedures • Conflicts of interest guidelines • Confidentiality agreements • Training /Awareness Program
	Post instructions	<ul style="list-style-type: none"> • Email disclaimers • Security Policy • Access labels for critical areas
	Alert conscience	<ul style="list-style-type: none"> • Use of messages, i.e. copyright protection, privacy protection e.tc.
	Assist compliance	<ul style="list-style-type: none"> • Security education for staff • Single sign-on • Point of reference for security issues
	Control drugs and alcohol	<ul style="list-style-type: none"> • Screening

Many of the traditional countermeasures, which are already used in typical organizational context, aim at reducing opportunity. We have found analogies that make sense for 22 of 25 techniques. We have not found logical or applicable as-

sociations with only three techniques. The concept of reducing provocation is less applicable in the Information Security field, as we could not find applicable measures for avoiding disputes or reducing emotional arousal. These techniques have been applied in more violent environments, like bars or football fields. The same fact applies for techniques of controlling drugs and alcohol, which is not a threat associated with computer abuse. It would apply in a form of screening, only to law enforcement employees or to some extent to security employees for critical infrastructures.

Social and ethical considerations

Situational crime prevention theory has been critiqued about the social and ethical issues that arise by its application (Clarke, 2000). In the following section, we examine which specific arguments have been posed against the theory. We examine how the 25 techniques have been adjusted to the computer abuse field and which are more likely pausing social and ethical considerations, thus, requiring more attention during their application. Since the theory focuses on removing opportunities for crime, many argue that removing a type of opportunity may cause displacement (Clarke, 2000). This means that it may lead to different means or ways to commit the crime or to different types of crime, thus, potentially creating escalated results. One typical example of the computer field is the one of email spam, where more sophisticated ways of sending email spam are constantly developed, but the phenomenon increases despite the growing anti-spam market and the plethora of countermeasures. The same idea can apply to many of the above countermeasures. For example, if one focuses on protecting sensitive data mainly in their electronic form, one is more likely to follow the easiest path of acquiring the data in their printed form. This means that one needs to be prepared for possible side effects and reexamine the results of their application, a task usually neglected after completing a standard or regulatory compliance or a risk management process.

If we observe some measures, like the target hardening ones or the natural and physical surveillance ones, they may appear extreme and may not be accepted by the employees of the organization. They can produce the image of a «fortress society» and create associations to the Orwellian «Big Brother» (Clarke, 2000). Characteristic examples are physical constraints like CCTV, alarms, barriers or locks, as well as logical, equivalent measures like monitoring, audit and event logging. These measures are usually applied in areas where critical data or equipment are stored, but not throughout the organization. For example, in Critical Infrastructures, such measures are commonly used, due to the potential high impact of computer abuse in such systems, combined with the culture of the employees, who are accustomed to screening processes. However, one must not assume per se that employees of a business organization will not get affected by such measures. Special attention is necessary to ensure

that civil rights are not threatened in the effort to mitigate computer abuse by insiders. Some measures, like the one that includes monitoring or screening of employee activity raise ethical or legal issues. Balancing security and privacy is not a trivial task (Mitrou and Karyda, 2006), as both culture and legal frameworks vary significantly among countries and organizations.

Another factor that one needs to balance, is the possible inconvenience caused by the measures. An example is the use of sophisticated or complicated authentication techniques, which can tire or frustrate the employees or visitors of an organization. A critique to the theory is that it forces inconvenience on the 'law-abiding' ones, who are the ones that it wishes to protect in the first place (Clarke, 2000).

In order to handle such social and ethical concerns, the following recommendations can be made. One must select measures that are appropriate for the risk analysis and management method used. This means the security officer needs to suggest measures that will cover the framework of situational crime prevention, but will also apply them when and where the risk is high enough to justify it, following the principles of proportionality and necessity. One must also ensure that the selected measures are not in contrast with the organizational culture, which is not an easy task to determine and estimate, as it may vary among employees, organizations and countries.

The proposed set of measures must also comply with the legal and regulatory framework of the organization. It is recommended to select the least obtrusive measures possible, which will not be that noticeable or inconvenient to the employees. However, the measures that are applied need to be disclosed and transparent to the employees, especially those that may affect their personal privacy. Another idea is to maintain existing measures as much as possible, since the employees are accustomed to them. For any additional measures that may be required, the above preconditions need to be fulfilled.

Conclusions and future work

In this paper we discussed how the Situational Crime Prevention theory could be adjusted in order to provide insight and be applied in an Information Security context. More specifically, we examined the theory in terms of computer abuse by insiders. We adopted the five goals and the respective 25 techniques that the theory suggests and adjusted them for the particular context of business or critical infrastructure and the crime of computer abuse by insiders.

We formed a set of countermeasures that, more or less, are already used by best practices and methods, and examined how these relate to the techniques of the theory. This way we examined if the concepts of the theory can be applied to the

field. Since the theory affects the environment of an offender and some measures can be obstructive, we discussed potential social and ethical issues that rise by applying such a method for mitigating computer abuse.

This is a first view on how the method could be translated for the Information Security field. The 25 techniques could be enriched furthermore by other sources, like other standards or risk management methods. Future steps would include designing case studies in order to test the theory in this context, which is not a trivial task and would require a joint computer and criminology research team. One of the most challenging tasks is to design case studies, and receive statistical data, as insider abuse presents the difficulty that many crimes remain undetected and losses and crime rates are not easily quantifiable.

REFERENCES

PriceWaterhouseCoopers, Enterprise and Regulatory Reform (2008) 2008 Information Security Breaches Survey - Technical Report.

Clarke R. (1980) Situational crime prevention: theory and practice. *British Journal of Criminology* 20, 136-137.

Clarke R. (1997) Situational crime prevention: successful case studies. NY: Harrow and Heston.

Clarke R.V. (2000) Situational prevention - Social values and social viewpoints. In: von Hirsch, A., Garland, D., Wakefield, A. (Eds.) *Ethical and Social Perspectives on Situational Crime Prevention*, Oxford: Hart Publishing, 97-112.

Clarke R.V. (2005) Seven misconceptions of situational crime prevention. In: *Handbook of Crime Prevention and Community Safety*, N. Tilley (Ed.), Willan Publishing.

Cornish, D., Clarke, R. (2003) Opportunities, Precipitators and Criminal Decisions: A Reply to Wortley's Critique of Situational Crime Prevention. In M. Smith, D. Cornish (Eds.), *Theory for Practice in Situational Crime Prevention*. *Crime Prevention Studies*, 16, 151-196, NY: Criminal Justice Press.

ISO/IEC 27002:2005 (2005), *Information technology - Security techniques - Code of practice for information security management*, ISO.

Mitrou E. & Karyda M. (2006) Employees' Privacy vs. Employers' Security: Can they be balanced?, *Telematics and Informatics Journal*, Elsevier, 23 (3), 164-178.

Noonan T., Archuleta E. (2008) Insider threat to critical Infrastructures. The National Infrastructure Advisory Council's Final Report and Recommendations.

Richardson R. (2008) 2008 CSI Computer Crime & Security Survey.

Theoharidou M., Kokolakis S., Karyda M., Kiountouzis E. (2005) The insider threat to Information Systems and the effectiveness of ISO 17799. *Computers & Security*, 24 (6), 472-484.

Willison R. (2001) Understanding and addressing criminal opportunity: the application of situational crime prevention to IS security. Working Paper Series 100. Dept. of Information Systems, London School of Economics and Political Science.

Willison R. (2004) Understanding the offender/environment dynamic for computer crimes: Assessing the feasibility of applying criminological theory to the IS security context. Proc. of the 37th Hawaii International Conference on System Sciences.

Willison R. (2006a) Understanding the perpetration of employee computer crime in the organizational context. Working paper no. 4-2006, Copenhagen Business School.

Willison R. (2006b) Understanding the perpetration of employee computer crime in the organizational context. *Information & Organization*, 16 (4), 304-324.

Everybody wins: challenges and promises of knowledge discovery through volunteer computing

Lauri Tuovinen and Juha Röning*

Intelligent Systems Group, Department of Electrical
and Information Engineering

Abstract

Knowledge discovery technology provides methods for extracting useful knowledge from large quantities of data. The knowledge, hidden in the distribution and internal structure of the data, is beyond the reach of traditional analysis methods, but with data mining algorithms it can be found. Effective knowledge discovery often requires access to substantial computing resources, and one way to acquire this access is to employ volunteer computing. In volunteer computing the data is processed by computers administrated by individuals or groups who are willing to take part chiefly for the pleasure of contributing to interesting research. The computers are mainly used for other things, but a share of their processor time is devoted to solving assignments downloaded over the Internet from a project server. The challenges involved are not all technical but social as well; for example, establishing trust between researchers and volunteers is of major importance. It is therefore evident that issues with significant ethical connotations may need to be solved in knowledge discovery, yet such issues are completely disregarded by currently accepted, technically oriented process models. We examine volunteer computing, and knowledge discovery in general, to identify the social factors that should be addressed in order to ensure that a knowledge discovery effort satisfies the legitimate expectations of all stakeholders.

* Lauri Tuovinen (M.Sc.) is a researcher and postgraduate student at the Department of Electrical and Information Engineering, University of Oulu, Finland. He has been a member of the department's Intelligent Systems Group (ISG) since 2002, working on various aspects of knowledge discovery and representation.

Juha Röning (Dr.Tech.) is Professor of Embedded Systems, head of the Department of Electrical and Information Engineering and principal investigator of ISG at the University of Oulu. He has been a visiting research scientist at the University of Cincinnati and held a Young Researcher Position at the Academy of Finland. His main research interests are intelligent systems, machine vision and software security.

Keywords: Knowledge discovery, data mining, volunteer computing, stakeholder analysis, trust, fairness, persuasion, privacy.

Introduction

Knowledge discovery through data mining is a branch of computer science that studies methods and technologies for extracting knowledge from quantities of data so vast that special algorithms are required to reduce the data into a compact representation that can be understood and used. The knowledge comes in the form of interesting patterns that are hidden to a superficial inspection but can be made visible by fitting a computational model to the data.

Research on knowledge discovery largely focuses on technology: methods for preparing the data, mining it for knowledge and visualising the results. This work has been fruitful and resulted in many successful applications in all areas of human activity where data is generated and gathered. Just to name a few prominent application domains, knowledge discovery has proved useful in industrial quality control (Chen et al., 2004; Zhang et al., 2003), medical diagnostics (Li et al., 2004; Zaffalon et al., 2003) and marketing (Chou et al., 2000; Gersten et al., 2000).

More recently there has been increasing interest in the process through which knowledge is extracted from data (Laurinen, 2006). Improved understanding of the process, together with the accumulation of effective algorithms for data mining tasks, has in turn made it possible to create new technology in the form of data mining frameworks that combine a library of algorithms with a graphical application builder (Berthold et al., 2006; Mierswa et al., 2006). Several of today's most widely used database management systems are also now available bundled with a data mining toolkit, illustrating the acceptance of knowledge discovery technology as a powerful business instrument.

A notable deficiency in the way the knowledge discovery process is currently viewed is that while it covers the transformation steps required to get from data to knowledge and the interactions between the steps, it largely neglects the people who participate in the process and the interactions between them. This is a significant oversight because people and the relationships they share play a central part in any knowledge discovery effort. These relationships can sometimes be quite intricate, as demonstrated by volunteer computing.

Volunteer computing is a form of distributed computing where the processing power needed to solve a problem is provided by volunteer-administrated computers working part-time on small fragments of the problem. Volunteering usually takes place by downloading and installing a client application, which announces itself to a project server. The server splits the data to be processed into independ-

ent work units, distributes the work units among the volunteers and collects the results. Volunteers are free to leave the project at any point and generally receive no material compensation for their contribution.

In the best case, volunteer computing is an arrangement where everybody wins. The people running the project, who are usually researchers, gain access to computing resources at a fraction of the cost of buying the equivalent in new hardware, and the volunteers get pleasure from contributing to a worthy cause and from belonging to a community of people with a shared interest. Society benefits from the new knowledge acquired and from the more efficient use of existing resources.

None of these benefits come automatically: they are contingent on the establishment of a reciprocal relationship between the researchers and the volunteers. Such a relationship can not exist unless certain fundamental conditions are satisfied. The availability of enabling technology is an important condition, but there are other conditions, social rather than technical, with considerable ethical implications. Particularly notable is the necessity of trust, because both parties are making themselves vulnerable in the arrangement.

The diverse technological and social issues surrounding volunteer computing are illustrative of the complex interplay of people, knowledge, technology and ethics that is characteristic of knowledge discovery in general. People collaborate with technology and with other people to generate new knowledge, which is often used to create new technology. The social dimension of the discovery process raises ethical concerns related to the expectations and responsibilities of various stakeholders, and the knowledge itself may have ethical implications depending on who has the power to apply it and how.

In this paper we explore the social component of the knowledge discovery process and especially how it manifests itself in volunteer computing. We find concrete examples of social and ethical issues that need to be settled to ensure the success of a volunteer computing effort, thus demonstrating the inadequacy of the current technology-centred process model of knowledge discovery. From volunteer computing we expand the discussion to present a view that covers all possible stakeholders and their legitimate concerns. The objectives set for a knowledge discovery project should be based on the complete list of relevant stakeholder concerns.

The next section examines the concept of knowledge in knowledge discovery and its interactions with technology and ethics. A more detailed introduction to volunteer computing is then given, with several examples of active projects. This is followed by a discussion of the stakeholders of volunteer computing and the ethi-

cal issues associated with their relationship. Finally, we expand this perspective to the stakeholders of knowledge discovery in general.

Knowledge, technology and ethics

There is a two-way relationship between knowledge and technology. The development of new technology is based on existing knowledge, and conversely, the search for new knowledge is aided by existing technology. Data mining illustrates this cycle neatly: technology (measuring instruments, databases, algorithms) is used to generate knowledge (predictive or explanatory models), which is then incorporated in technology (software applications).

Besides knowledge discovery, another important concept bridging the gap between knowledge and technology is knowledge representation. This refers to creating data structures specially designed for storing not only data but also the semantics of the data. Databases based on such data structures are known as knowledge bases.

The notion that knowledge can be acquired and possessed by technology has the potential to stir up philosophical controversy. This has nothing to do with storing knowledge per se – that has been done since the invention of writing, and even today some of the systems called knowledge bases are actually barely more than digital reference books. What makes knowledge discovery and knowledge bases philosophically interesting is that they appear to challenge the exclusive claim of humans to knowing things. The dependence of computers on humans diminishes when computers no longer only store and retrieve knowledge but can also generate more knowledge and use it to solve problems without human intervention. The more computers come to resemble independent actors instead of tools, the more tempting it is to say that they indeed possess knowledge.

Whether it can truly be said that computers have the ability to know things is a question beyond the scope of this paper, but the nature of the relationship between humans and computers in knowledge acquisition is not. This relationship is becoming less and less straightforward as computing technology evolves and is able to tackle more and more complex tasks. The key to successful knowledge discovery lies in seamlessly combining the strength of computers, arithmetic prowess, with the strength of humans, creativity. Humans are, in effect, partially outsourcing their thought just like they have partially outsourced their memory by making records of things they know.

The implication of the preceding paragraphs is that while it is debatable whether machines can independently discover and possess knowledge, it is well established that they have entered a symbiotic relationship with humans in the ac-

quisition and storage of knowledge. Computers without human initiative and guidance would be useless, but on the other hand, humans without the aid of computers would be severely crippled in their attempts to understand the huge quantities of data characteristic of empirical research today.

Both knowledge and technology are sources of considerable power in the sense that someone in possession of them can accomplish things that would otherwise be beyond their reach. Power can be wielded benevolently or malevolently, so knowledge and technology are inevitably linked with ethics. Not every piece of knowledge or every item of technology is ethically controversial, but the debates generated by those that are can get very hot. Examples are abundant among the scientific discoveries of the modern era – nuclear energy, stem cells and cloning, just to name a few.

Especially interesting from the perspective of this paper are the intersections of knowledge discovery with ethics. There are many such intersections, as demonstrated by the survey made in (Tuovinen & Rönning, 2005). The most interesting of these are the applications where knowledge discovery techniques are applied to data regarding human individuals, because they give us yet another role in which humans may appear in the knowledge discovery process: as objects of study. As such they are entitled to be treated with certain respect by the people doing the studying. In particular, they have the right to expect the data concerning them to be handled in a way that respects their privacy.

When discussing privacy preservation in knowledge discovery, we can make a distinction between two cases based on who is gathering the data. We shall consider first the case where the data is collected by a company, a research institute or some other organisation with no authority status. In this case collecting the data generally requires the consent of each individual, and the collecting organisation has an obligation to state clearly what the data may be used for and by whom. Violating these standards may lead not only to moral outrage but to litigation as well.

The situation is somewhat different when the data is collected by a public authority. Organisations such as police departments and tax offices hold a legal mandate to gather information about the people within their jurisdiction and to use the information to enforce laws and regulations. Problems would soon ensue if acquiring the information relied entirely upon the goodwill of the people, so in such special cases the common good is considered to override the right of individuals to control information about them.

Fundamentally, however, the two cases are not that different in a society that recognises its citizens' right to privacy. The law grants the government the abil-

ity to breach this privacy, but it must be to protect some interest considered even more important and the scope of the breach must not be out of proportion. By this principle, privacy violations are sometimes necessary, regardless of whether knowledge discovery technology is involved, but they should be rare and limited in extent, again whether or not knowledge discovery is involved. We will therefore treat such breaches as exceptions and drop the distinction between our two cases for the rest of our discussion.

Assuming, then, that sensitive data can only be gathered with the consent of the individuals with whom it is associated, it might seem that there is nothing to discuss. However, the reason why privacy issues in knowledge discovery have sparked such interest is that even if the data in itself is not sensitive, the conclusions may be, because data mining techniques can identify surprising connections between data items that, considered separately, appear unrelated and uninteresting (Tavani, 1999). To counter the unwanted effects of this ability, privacy preserving data mining techniques have been studied (Verykios et al., 2004). The goal of these is to allow useful knowledge regarding a population to be derived while making it impractically difficult to connect any of the knowledge with specific individuals in the population.

The existence of sensitive data, or data from which sensitive conclusions may be drawn, in a database is not necessarily in itself a problem, if the organisation that administers the database is committed to using the data responsibly. However, there is always the possibility that the data is stolen by someone with no intention to honour such commitments. Therefore an important part of the responsibility of the database owner is to make sure that the security of its information systems is up to modern standards.

The requirement of proper information security also means that the data must be protected against corruption. If the data is bad, it is not realistic to expect that mining it will produce good knowledge. On the other hand, bad results may also come from good data if it is mined using inappropriate methods. In either case, the upshot may just be that some resources are wasted on a knowledge discovery effort that produces nothing useful, but if the results are to be used in a way that affects the people whom the data concerns – in a medical application, for instance – then incorrect results, or the delay from finding them incorrect and re-visiting the data, can be harmful to them. In such cases it is especially important to look after the integrity of the data and the knowledge derived from it.

Volunteer computing

Volunteer computing can be defined as scientific computing by means of an ad-hoc distributed system dependent on computing resources provided of their own

volition by individuals with no formal affiliation with the project. Instead of being formally employed, the participants choose to contribute because they find the work interesting and rewarding. The resources they contribute are typically desktop PCs with comparably modest Internet connections, available to the project only part-time and at unpredictable times. (Anderson, 2004)

The usual way of participating in a volunteer computing project is to download and install a client application distributed at the project website. The client connects to the project server, which assigns a work unit to the client. The client processes the work unit, connects to the server again, sends its result and receives another work unit (see Figure 1). This cycle continues until the project ends or the volunteer ceases to contribute. The volunteer can also choose the amount of resources allocated for the client; for instance, the volunteer may wish the client to be active only when the computer is running idle so that it will not consume resources when the machine is being used for something else.

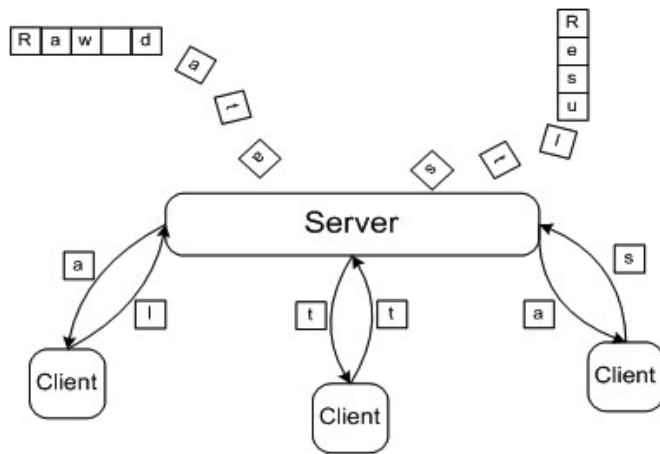


Figure 1. A volunteer computing server splits the data to be processed into work units and sends each work unit to a client. From the partial results returned by the clients the server pieces together the total result.

Since the volunteers may make their computers only sporadically available and are free to drop out any time they wish, volunteer computing faces technical challenges besides those that are always encountered in distributed computing. The project server can not contact the clients but has to wait for them to contact it, so the server is in a constant state of uncertainty regarding the state of the clients. The server therefore has to make decisions such as whether to keep waiting

for a given client to return its result based on incomplete information. Volunteer computing platforms like BOINC (Anderson, 2004) and SLINC (Baldassari et al., 2006) have been created to solve this and other technical difficulties, allowing researchers to concentrate on formulating the task of the clients instead of trying to tackle computer engineering problems that have nothing to do with their research

Volunteer computing has its roots in the GIMPS (<http://www.mersenne.org>) and distributed.net (<http://www.distributed.net>) projects. Both projects are still active; GIMPS searches for prime numbers belonging to a class known as Mersenne primes, whereas distributed.net has worked on several problems in cryptography and abstract mathematics. It is worth noting that the computational methods used in these projects are straightforward brute-force techniques – for instance, the RC5 cryptography subprojects at distributed.net rely on simply trying every possible key until the right one is found – so not all volunteer computing projects use data mining.

The project that really brought volunteer computing to public attention is SETI@home (<http://setiathome.ssl.berkeley.edu>), launched in 1999. SETI, an acronym for the Search for Extraterrestrial Intelligence, refers to efforts aimed at finding evidence of the existence of extraterrestrial civilisations from data recorded electronically by astronomers surveying the sky. The SETI@home data consists of radio signals collected using the 305-metre dish of the Arecibo Observatory in Puerto Rico. In 2004, the computers contributing to SETI@home were estimated to provide a sustained processing rate of over 70 TFLOPS (7×10^{13} floating-point operations per second), twice the rate of the largest conventional supercomputer at the time (Anderson, 2004). This figure gives an indication of the amount of processing power potentially available for a volunteer computing project to harness.

SETI@home, like many other volunteer computing projects, now runs on the BOINC platform. At the BOINC website (<http://boinc.berkeley.edu>) there is a list of projects that gives an idea of the variety of problem domains where volunteer computing can be a useful tool. There is also a list of scientific publications produced by BOINC projects; some of these appear in highly prestigious journals such as *Nature* (Murphy et al., 2004; Stainforth et al., 2005) and *Astrophysical Journal* (Cole et al., 2008). The accumulation of published results shows that it is possible to do serious scientific research while relying on volunteers to provide the computing resources needed for data analysis.

There are variants of the standard model of volunteer computing where the work units are processed by the volunteers themselves instead of their computers. One of these is Galaxy Zoo (<https://www.galaxyzoo.org>), which recruits volunteers

for visual classification of galaxy images. For each image, the volunteers are requested to answer a series of questions on various features of the galaxy. This approach is comparably simple to implement because no client application is needed – the volunteers simply use a web browser to log on to the Galaxy Zoo website, which presents the images and questions to them.

The Foldit project (<http://fold.it>) adopts a more elaborate approach. There is a client application that the volunteers download and install, but instead of performing computations in the background it allows the volunteers to play a puzzle game. The objective of the game is to fold a protein into the most stable state it can assume, given its amino acid structure. This information is used to predict the shape the protein will naturally fold into, which in turn is crucial knowledge for understanding the function of the protein.

The list of volunteer computing projects given above is not exhaustive, but it is illustrative of the wealth of research areas where volunteer computing can be applied and the variety of technical approaches that can be adopted. Technical implementation is not everything, however; there are also social and ethical matters to be considered, although these may not be immediately obvious. We explore these issues in the next section.

Ethical issues in volunteer computing

As seen in the previous section, there are many technical challenges in volunteer computing, but several projects have been able to overcome these and succeed in producing new knowledge. Thanks to free platforms such as BOINC, a researcher can start a new volunteer computing project without extensive computer skills or a great investment of money or time. This has brought volunteer computing closer to being a universally accessible tool for knowledge discovery.

BOINC and similar systems open the necessary channel of communication between researchers and volunteers, but this is just the beginning. The computing platform provides the tools for collaboration, but it is up to the people participating in the project to use those tools productively. It turns out that some of the most important issues in volunteer computing are not technical but ethical. Three major themes can be singled out: trust, fairness and persuasion. Below we discuss each of these separately, but we find that they are intertwined in many ways.

Trust

A research project that depends on volunteer computing will not get anywhere unless the researchers are able to muster a large enough group of people willing to contribute to the research. Reaching people who find the project interesting is

a matter of properly designed and executed advertising, but interest alone is not enough to make one a potential volunteer; trust is also a crucial requirement.

Trust is important because downloading and installing software without knowing exactly what it does is always a risk. Contributing to a volunteer computing project means installing a client application and allowing it to use the processing capacity and network connection of the client machine, and it takes considerably higher than average computer skills to be sure that the application will not have any undesirable effects. The researchers running the project therefore need to get the public to trust that the software they are distributing is not malicious or dangerously defective.

On the other hand, the researchers are also taking a chance in trusting the volunteers. They hope that those who install the client are motivated by a sincere wish to contribute to the research, and mostly this is a safe assumption, but the possibility of somebody joining the project with the intention of sabotaging it can not be ruled out. For example, a rival research team might want to hinder the work by trying to feed false results to the project server.

As Anderson (2004) points out, the volunteers in volunteer computing are in an asymmetric relationship with the researchers. One of the implications of this asymmetry is that volunteers and researchers must rely on different means to gain the trust of the other party. In fact, it would be more accurate to say that the establishment of trust is entirely in the hands of the researchers – it is up to them to make sure both that the volunteers can trust them and that they can trust the volunteers.

With potentially hundreds of thousands of aspiring volunteers, individually evaluating each one to prevent potential saboteurs from joining is not a feasible task. Instead, the researchers must accept, and adjust to, the fact that some percentage of the volunteers may be trying to sabotage the effort. In practice this means using technology that minimises the effects of malicious behaviour on the efficiency and results of the computation.

The traditional technique for detecting and discarding falsified results is redundant computing, where each work unit is sent to several clients and the clients vote to determine the canonical result for the unit (Anderson, 2004). When the number of clients whose results agree reaches some threshold M , their result is chosen as the canonical one. Increasing M makes the system more resistant to bad results, but it also increases the time it takes to finish the computation. Sarmenta (2001) has proposed an open framework where voting can be combined with other techniques such as spot-checking and blacklisting to calculate a credibility measure for each result. Using a combination of methods reduces the slowdown

incurred by a given error-tolerance level compared to the case where only voting is used.

So, when it comes to finding trustworthy volunteers, the best the researchers can do is give everyone a chance and try to weed out the dishonest ones based on their output. Not much here is of philosophical interest: there is nothing ethically ambiguous about sabotage – it is clearly wrong, at least assuming that the research itself is not ethically questionable – and improving the error tolerance of computations is a technical issue. Blacklisting clients without human discretion is somewhat problematic because an erroneous result is not necessarily intentionally falsified, but this is a tangential issue, and to avoid a diversion from the main thread of discussion, we shall not treat it in any detail.

It is worth noting here that although implicitly trusting every aspiring volunteer has its risks, it also helps bring in honest volunteers. People wishing to join a volunteer computing project appreciate it if joining is made easy for them, whether or not they consciously think of it as a sign of trust. For example, as Sarmenta (2001) points out, requiring volunteers to provide a stronger form of identification than an email address would make it more difficult for blacklisted saboteurs to rejoin, but it would probably also deter many good volunteers. We shall return to this topic in the subsection on persuasion.

Given that the researchers do not know who the volunteers are, except in the rather trivial sense that each volunteer is required to register, it could be argued that the way the researchers relate to the volunteers is not really a matter of trust, except in the sense that they show a trusting attitude when they presume that the majority of registered volunteers are honest. If it can be safely assumed that most participants are not out to sabotage the project, tolerance mechanisms built into the computing server can be used to cancel out the effect of those who are. We see a different picture, however, when we look at the way the volunteers relate to the researchers.

One aspect of the asymmetric nature of volunteer computing is that the volunteers are in a better position to gather information about the other party and to decide whether they want to have any dealings with them. This seems appropriate, since it is the volunteers who are being asked to give someone else partial control of their property. It does, however, raise the question of what the researchers can do (and also what they should not do) to convince the volunteers that they can be trusted with what they are asking for.

Also here technology may help the researchers, although in a more indirect manner. Since BOINC, for instance, is used by dozens of projects and the client has been installed on hundreds of thousands of computers, it is likely that if there

were some feature or defect in the software that may damage client systems, somebody would soon discover and report it. If there are no such reports to be found, it is an indication to potential volunteers that it is relatively safe to run the software. Still, the researchers are ultimately responsible for the software they use and should act accordingly to protect it against attempts to sneak malicious code into it.

The role of word-of-mouth in assessing the trustworthiness of software is a specific example of the more general truth that reputation is an important factor in establishing trust. One's expectation of someone's trustworthiness is influenced by the recommendations of others, and the strength and direction of the influence are in turn affected by the perceived trustworthiness of the recommenders. Several authors have undertaken to describe the dynamics of trust and reputation in networked computer systems as formal models; see e.g. (Abdul-Rahman and Hailes, 2000).

In the future it may be possible for organisations and individuals to rely on such models to automatically verify the trustworthiness of the organisations and individuals they interact with, but currently, barring a few special cases, there are no such shortcuts available. An organisation wishing to recruit computing volunteers for a research project can therefore mostly just hope that it has a positive public image and try to make sure that potential volunteers associate the image with the project.

A research group has a good chance of accomplishing this by being thoroughly open about itself: which scientific institution it is affiliated with, which institutions it has collaborated with, which research topics and projects it has worked on, what it is trying to achieve now, who are currently working in the group, who have worked there in the past. This gives a potential volunteer, sympathetic to science in general, a chance to connect the group to something or someone he or she already knows and thinks positively of. Such connections lend the group credence by association.

About openness it is worth noting that when extended to other aspects of volunteer computing, it leads to a trade-off between how much the researchers can trust the volunteers and vice versa. In particular, the more details are disclosed about the computing software, the more confidence the volunteers can have that the software is not harmful, but the easier it is for malicious volunteers to modify it to produce incorrect results. However, since it is the volunteers who are making their own computers vulnerable and since the researchers have ways to defend the integrity of the results against malicious behaviour, it is best for the volunteer computing software to be open source.

Fairness

Volunteers deserve to be treated fairly – this is a very basic right, not negated by the fact that they are volunteers. This does not mean that they must be compensated materially for their contribution, since volunteer computing projects are typically nonprofit research efforts. Instead of compensation, the key fairness factor in such projects is acknowledgment.

A common approach to giving due acknowledgment to volunteers is credit. Credit is basically an immaterial form of compensation, an accumulative number that represents the amount of work a volunteer has contributed to a project. Based on credit it is possible to compile various statistics, giving volunteers the chance to earn some public recognition by appearing on a list of top contributors.

How credit accumulation is determined is a matter of some significance, because the measure should be objective to such an extent that keeping statistics makes sense. If it is not possible to make meaningful comparisons between the credit values of different contributors, then the credit system serves no purpose except to confirm to the volunteers that their computers are doing something for the project. The amount of credit earned is, by itself, an almost meaningless number; knowing whether it is a little or a lot requires a context, and this context is provided by the ability to compare the number among peers.

The other side of credit is that no-one should be able to gain undeserved credit, since that would be unfair to those who have acquired theirs with honest work. Conveniently, redundancy may again provide the solution. BOINC, for example, uses an accounting scheme where clients are not automatically given the amount of credit they claim for their results but receive the minimum or average of the claimed credit of all correct results instead (Anderson, 2004). This prevents clients from bolstering their credit accounts with dishonest claims.

Overall, crediting volunteers for the CPU time they have donated is a good way to acknowledge their work. Distributing credit for publishable results is a different matter, however. If the researchers write, for instance, an academic journal paper on results achieved with the help of volunteer computing, what is the proper way to acknowledge the significance of the volunteers' contribution? The answer is largely dictated by what is practical: it is simply not possible within reason to include the names of all volunteers in the paper. On the other hand, good academic form demands that if a nonauthor has contributed to a publication, the authors mention the contribution. A collective expression of gratitude to the volunteers is therefore both necessary and sufficient.

A special case worth treating separately is when a discovery can be traced to a single computer or a small number of computers. For example, it is conceivable

that if SETI@home one day achieves its goal and confirms an extraterrestrial radio signal as a transmission by an alien civilisation, it will be possible to identify the work units in which the signal was found and the clients that processed the work units. It might therefore be possible to associate the discovery with a relatively small number of volunteers, who might expect some kind of special recognition.

A fact that speaks against granting such recognition is that the distribution of work units among volunteers is a random process, albeit weighted by the amount of resources devoted to the project by each volunteer. Granting coauthorship in a scientific publication on such basis seems inappropriate, and even a special mention in the acknowledgments section might be unmerited. Perhaps the best option would be to give the names of these volunteers in a less formal context, e.g. a press release or a news item at the project website. Besides, some of the volunteers might not even want their identities published, in which case the researchers would have to respect their privacy.

A version of this scenario has already been seen at distributed.net, where volunteers have participated in competitions to solve secret-key ciphers of increasing strength. To motivate these efforts a cash prize of 10 000 USD for whoever finds the correct solution has been offered by RSA Laboratories (<http://www.rsa.com/rsalabs/node.asp?id=2100>). So far distributed.net has won two of these challenges and is organising the next one itself since RSA Labs has discontinued the contest. The distribution of prize money follows a formula: the volunteer who finds the winning key gets a part, another part goes to distributed.net for providing the necessary infrastructure, but most of the money is donated to a nonprofit organisation jointly selected by all volunteers.

The difference between this and the SETI@home scenario is that finding the secret key is not a new discovery, so there is no need to argue about who gets credit for it. The prospect of winning some cash makes participating in a secret-key challenge akin to entering a raffle, and whereas winning money in a raffle is generally considered acceptable, determining paper authorship by such means would most definitely be unacceptable. Money, unlike authorship, does not inherently belong to anyone, which leaves distributed.net free to share it in a way that achieves a good balance between advancing the public good and offering incentives to volunteers. The chosen balance seems to meet the approval of the participants, considering that when the RC5-64 cipher was solved, distributed.net itself was voted as the nonprofit organisation to receive 60% of the prize money.

Persuasion

The act of persuading is not inherently ethical or unethical. However, people may be persuaded to do ethical or unethical things, and the means of persuasion may be ethical or unethical. We conclude this section with a discussion of the persuasion techniques available to researchers wishing to recruit computing volunteers and the ethical issues associated with those techniques.

Persuasion through the use of computers has been extensively studied by Fogg (2003). Several of the persuasion techniques he has identified are already being used to recruit volunteers and to make them stay. It is hard to say to what extent researchers who use volunteer computing are aware of the theory of persuasion by information technology, but it would be good if they were, since it would both help them persuade effectively and help them avoid persuasion techniques that are ethically unsound.

The good news for researchers is that even if a project has proved difficult to sell to financiers, it may not be all that hard to find volunteers for it. It is a fairly small commitment to let one's computer do some work for someone else when it would otherwise be idle, and many people are happy to make it if they find the work interesting and agree with its goals. SETI@home is again a good example: few things are more exciting to a curious mind than the prospect of receiving a message from an extraterrestrial intelligent species, so the researchers are in a good position to write a project description that attracts curious minds.

It is good for researchers to have a grand vision – contact with another civilisation or a cure for a serious disease, for example – and it is not wrong to communicate it to potential volunteers in a fashion that appeals to their sense of romance and heroism. However, the researchers should clearly discern romance from realism and be cautious about what they promise to achieve. In the case of SETI, for instance, nobody knows for sure if there even is anyone out there for us to have contact with, so it would not be honest to imply that the project will certainly find what it is looking for. Scenarios presenting possible outcomes should be accompanied by the assumptions on which they are based and the uncertainties associated with the assumptions.

We have already mentioned the simplicity of joining a volunteer computing project and how it may be viewed as a sign that the researchers trust the volunteers. According to Fogg, simplicity is in fact one of the most powerful persuasion techniques available. If a potential volunteer is allowed to join after answering just a couple of simple questions, it is likely that he or she will. If there are many questions and the volunteer constantly finds him- or herself wondering whether he or she wants the researchers to have the information requested, it increases the likelihood that the volunteer will not complete the joining process.

Yet another way to view this is that by not requiring volunteers to supply such information as their phone numbers or even their real names, the researchers protect the privacy of the volunteers. Privacy issues may thus prove significant even if no such issues are raised by the data.

Rewards are another powerful way to persuade, and even rewards that are completely immaterial in composition may be far from immaterial in consequence; Anderson (2004) has observed the motivating effect of credit in the case of SETI@home, and Fogg (2003) confirms the general principle. Credit is particularly effective thanks to the competition it encourages among volunteers. The competition, in turn, so long as it remains friendly, fosters a certain sense of community, which is a healthy value in itself and also further strengthens the commitment of the volunteers to the project and increases the satisfaction they get from their participation.

The community aspect of volunteer computing is probably something that researchers could use to a considerably greater effect than they do now. An interesting idea would be to borrow elements from popular social networking sites such as Facebook; for instance, volunteers could have their own profile pages where they could display merit badges, awarded by the computing server when a specific level of accumulated credit is reached. Credit could also be redeemable in an online shop for virtual items, or even promotional t-shirts and other concrete items that are relatively cheap to make. In fact, the capabilities of existing services could be directly used: there could be, for example, a SETI@home Facebook application allowing volunteers to display their achievements, get in touch with other volunteers and invite their friends to join. This would provide the added benefit of simplicity, since volunteers could build their communities at sites they visit frequently anyway.

The client application can also be persuasive. This is obvious if the application is a game as in the case of Foldit, but it is true even if the client is not interactive. For example, there is Fogg's (2003) principle of attractiveness, which states that if a computing technology is visually attractive, it is likely to be more persuasive. SETI@home, among others, leverages this principle by providing colourful graphics, showing in real time what the application is doing. Volunteers can set the client to run as their screen saver, allowing them to get up to date on the progress they are making every time they glance at the computer screen.

A question worth considering is what information the client application should display to the volunteer running it. To maximise the persuasive qualities of the application, the information should be interesting to the volunteer and presented in an attractive way that makes it easy for a layman to grasp its significance. On the other hand, it is not in the interest of the researchers to let important discoveries come to public attention before they themselves are ready to publish them.

Here is another possible trade-off involving openness: instant feedback of interesting results can act as a reward for volunteers, but it also raises the question of how much information the volunteers can be trusted with before it is made officially public by the researchers.

In practice, though, this may not pose a problem, since the amount of information the client application can impart to volunteers is limited anyway because of the distributed nature of the computations and because the volunteers generally are not trained to interpret the results. Thus it seems unlikely that a volunteer could independently draw any sensational conclusions based on the output of his or her copy of the client software alone. The trust issue is averted in this case simply by considering what is really known; having the client convey an impression of imminent breakthrough might be exciting and persuasive, but such an impression would be based on guesswork at best.

Knowledge discovery as a social effort

The lesson of the previous section summarised is that there is a social aspect to knowledge discovery through volunteer computing that must not be neglected. It is important to have good technology, but it is also important to have good relationships among the people involved or the effect of the technology will be blunted. This lesson can be extended to knowledge discovery in general.

The most central position in knowledge discovery is held by the experts coordinating the discovery process. In the case of SETI@home, for example, the group of experts consists of the researchers running the project server, but generally there can be two kinds of experts: domain experts, intimately familiar with the problem domain, and technology experts, with the skills and knowledge necessary for the successful application of knowledge discovery technology. The first social relationship that has to work smoothly for a knowledge discovery effort to succeed is the one between domain and technology experts.

A good relationship between different kinds of experts requires two things: that the domain experts can communicate the problem to be solved to the technology experts, and that the technology experts can communicate the capabilities of the technology to the domain experts. Based on this mutual understanding the two groups of experts can work together to find a way to use the technology to solve the problem. The availability of general-purpose discovery tools may have diminished the role of technology experts to some extent, but even a very good tool does not yield good results when applied blindly, so there must at least be someone in the domain expert team with a solid understanding of technology.

The relationship between technology and domain experts is approximately analogous to the relationship between professionals and clients in software engineering and is therefore bound by similar ethical principles. Software engineering ethics is a rich field of study and covering it in any significant detail falls outside the scope of this paper; let it suffice to say that there are obligations to be fulfilled both ways, since there is a contract – written, verbal or unspoken – that both parties should honour. For relevant literature see e.g. (Bayles, 1989) and (O'Boyle, 2002).

The diagram in Figure 2 shows the five stakeholders of knowledge discovery we have identified. The four roles that people may play in the discovery process are denoted by ellipses, with society as a whole representing the fifth stakeholder. The arrows denote interactions among the stakeholders in terms of expectations; each arrow is annotated with the expectations that a stakeholder should seek to fulfil in its relationship with the other stakeholder. Some stakeholders also have an arrow pointing from itself to itself, meaning that members of a stakeholder group expect something from other members of the same group.

There are multiple names for some of the stakeholder groups; the ones called researchers we have also referred to as technology experts and data miners. They are the ones with the expertise required to apply knowledge discovery technology, which puts them at the centre of the figure, interacting with all other stakeholders. The group also has an expectation arrow pointing from itself to itself, because in the scientific community in particular, sharing one's results is the norm; a researcher expects to have access to new methods and tools developed by his or her colleagues around the world.

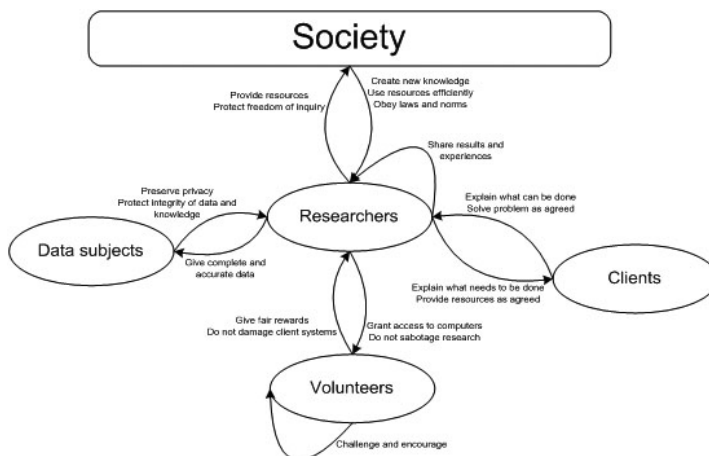


Figure 2. The stakeholders of knowledge discovery. Each stakeholder expects something from one or more other stakeholders and, in some cases, from other members of the same stakeholder group. The arrows depict these expectation-based relationships.

Clients we use as a general term for the people who know the problem domain and wish to affect it using knowledge acquired with the help of the researchers. The people we have referred to as domain experts form a subgroup of this stakeholder group. In some projects the clients do not exist as a distinct group, in which case their role is distributed among other stakeholders; in pure science, for instance, the researchers themselves are the domain experts and the acquired knowledge is received by society and other researchers.

People used in the knowledge discovery process as data sources are referred to in the figure as data subjects. Previously we have only discussed their rights, so it is worth noting that they also have at least one obligation: if they have agreed to give data to the researchers, it should be complete and accurate.

The volunteers were extensively discussed in the two preceding sections, so there is little more that can be said about them at this point. There is still a lot of untapped processor time left on the desktops of the world, so the dynamics of the researcher-volunteer relationship are well worth studying further. The arrow from this stakeholder to itself reflects the desire of volunteers to form a community; researchers would do wisely to find new ways to encourage and facilitate this.

Finally, society is the umbrella under which all of the above takes place. All the other stakeholders are subject to it and interact with it, but from the perspective of the knowledge discovery process, the researchers represent the main point of contact through which society's expectations are propagated into the process. Of course, the relationship is not one-sided – researchers also expect things from society, including both concrete necessities such as funding and abstract circumstances such as freedom of inquiry.

Conclusions

Volunteer computing is a form of distributed computing made possible by the proliferation of powerful personal computers and reasonably fast Internet connections. A project server assigns tasks to copies of a client application running on computers administered by volunteers, who allow the client to consume some of their processing resources because they find the project interesting and want to contribute to its success. Many projects have produced noteworthy results, showing volunteer computing to be a valuable tool for knowledge discovery.

Open platforms such as BOINC make it relatively easy for a research team to start a volunteer computing project, allowing it to gain access to potentially huge computing resources without making a major monetary investment. An infrastructure that solves the technical issues involved is not enough by itself, however; there

are also social and ethical issues to be solved, stemming from the fact that the researchers and the volunteers represent two groups of people trying to work out a mutually satisfactory deal. There can be no such deal if, for instance, the two groups do not trust one another. Some of the ethical issues of volunteer computing are accounted for in the design of BOINC and similar systems, but there are also issues that have not been adequately addressed and for which a technological solution may not even be possible.

This situation is illustrative of the situation in knowledge discovery in general: current process models are technically oriented and completely omit the social dimension of the discovery process. We consider this an important omission that, if it persists, will prevent knowledge discovery technology from achieving its full potential. We have therefore explored the social and ethical factors involved in volunteer computing and, by extension, knowledge discovery. We have identified the stakeholders of the discovery process and their interactions, bringing all of them together in a compact view that practitioners can use to make sure they take into account the justified expectations of all stakeholders. The mark of a truly successful knowledge discovery effort is that all participants come out satisfied, including those with no interest in the knowledge produced.

Acknowledgments

We would like to thank our colleague Dr Perttu Laurinen for his insightful comments at the early stages of composing this paper. L. Tuovinen wishes to thank the Graduate School in Electronics, Telecommunications and Automation (<http://signal.hut.fi/geta/>) and the Tauno Tönning Foundation (<http://www.tonninginsaatio.fi/>) for funding his postgraduate work.

REFERENCES

- Abdul-Rahman, A. and Hailes, S.** (2000) Supporting Trust in Virtual Communities. Proceedings of the 33rd Hawaii International Conference on System Sciences.
- Anderson, D.P.** (2004) BOINC: A System for Public-Resource Computing and Storage. Proceedings of the Fifth IEEE/ACM International Workshop on Grid Computing, 4 – 10.
- Baldassari, J., Finkel, D. and Toth, D.** (2006) SLINC: A Framework for Volunteer Computing. Proceedings of the 18th IASTED International Conference on Parallel and Distributed Computing and Systems.
- Bayles, M.D.** (1989) Obligations Between Professionals and Clients. In Professional Ethics, 2nd ed., reprinted in Johnson, D.G. (ed.), 1991, Ethical Issues in Engineering, 305–316.

Berthold, M.R., Cebon, N., Dill, F., di Fatta, G., Gabriel, T.R., Georg, F., Meinel, T., Ohl, P., Sieb, C. and Wiswedel, B. (2006) KNIME: The Konstanz Information Miner. Proceedings of the 4th Annual Industrial Simulation Conference, Workshop on Multi-Agent Systems and Simulation.

Chen, J., Fan, Q. and Xu, B. (2004) Research on the Application of Data-mining for Quality Analysis in Petroleum Refining Industry. Proceedings of the 5th World Congress on Intelligent Control and Automation, 4314–4318.

Chou, P.B., Grossman, E., Gunopulos, D. and Kamesam, P. (2000) Identifying Prospective Customers. Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 447–456.

Cole, N., Newberg, H., Magdon-Ismail, M., Desell, T., Dawsey, K., Hayashi, W., Purnell, J., Szymanski, B., Varela, C.A., Willett, B. and Wisniewski, J. (2008) Maximum Likelihood Fitting of Tidal Streams with Application to the Sagittarius Dwarf Tidal Tails. *Astrophysical Journal*, 683, 750–766.

Fogg, B. J., (2003). *Persuasive Technology: Using Computers to Change What We Think and Do*. Morgan Kaufmann, San Francisco, CA.

Gersten, W., Wirth, R. and Arndt, D. (2000) Predictive Modeling in Automotive Direct Marketing: Tools, Experiences and Open Issues. Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 398–406.

Laurinen, P., (2006). *A Top-Down Approach for Creating and Implementing Data Mining Solutions*. Dissertation, University of Oulu, Acta Universitatis Oulensis C 246.

Li, L., Tang, H., Wu, Z., Gong, J., Gruidl, M., Zou, J., Tockman, M. and Clark, R.A. (2004) Data mining techniques for cancer detection using serum proteomic profiling. *Artificial Intelligence in Medicine*, 32, 71–83.

Mierswa, I., Wurst, M., Klinkenberg, R., Scholz, M. and Euler, T. (2006) YALE: Rapid Prototyping for Complex Data Mining Tasks. Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 935–940.

Murphy, J., Sexton, D., Barnett, D., Jones, G., Webb, M., Collins, M. and Stainforth, D. (2004). Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nature*, 430, 768–772.

O' Boyle, E.J. (2002) An ethical decision-making process for computing professionals. *Ethics and Information Technology*, 4, 267–277.

Sarmenta, L.F.G. (2001) Sabotage-Tolerance Mechanisms for Volunteer Computing Systems. Proceedings of the First IEEE/ACM International Symposium on Cluster Computing and the Grid, 337–346.

Stainforth, D.A., Aina, T., Christensen, C., Collins, M., Faull, N., Frame, D.J., Kettleborough, J.A., Knight, S., Martin, A., Murphy, J.M., Piani, C., Sexton, D., Smith, L.A., Spicer, R.A., Thorpe, A.J. and Allen, M.R. (2005) Uncertainty in predictions of the climate response to rising levels of greenhouse gases. *Nature*, 433, 403–406.

Tavani, H.T. (1999) Informational privacy, data mining, and the Internet. *Ethics and Information Technology*, 1, 137–145.

Tuovinen, L. and Rönning, J. (2005) Balance of power: the social-ethical aspect of data mining. Proceedings of the Sixth International Conference of Computer Ethics: Philosophical Enquiry, 367–379.

Verykios, V.S., Bertino, E., Fovino, I.N., Provenza, L.P., Saygin, Y. and Theodoridis, Y. (2004) State-of-the-art in Privacy Preserving Data Mining. *SIGMOD Record*, 33(1), 50–57.

Zaffalon, M., Wesnes, K. and Petrini, O. (2003) Reliable diagnoses of dementia by the naïve credal classifier inferred from incomplete cognitive data. *Artificial Intelligence in Medicine*, 29, 61–79.

Zhang, C.-H., Di, L. and An, Z. (2003) Welding Quality Monitoring and Management System Based on Data Mining Technology. Proceedings of the Second International Conference on Machine Learning and Cybernetics, 13–17.

Information Systems and Future Ethics: A Reflection of Students' Current Behaviour

Aharon Yadin

Management Information Systems Department
The Academic College of Emek Yezreel

Abstract

This paper describes a study of the ethical behavior among college students, with a special emphasis on Computer Science and Information Systems students as compared to other learning disciplines. The research was trying to correlate heavy web usage, typical for CS and IS students, with the phenomenon of plagiarism. This is a follow-up study to previous research (Yadin, 2007) on plagiarism and students' abuse of information systems. The results obtained demonstrated no direct link between massive web usage and ethical misconduct. On the contrary, learning disciplines such Political Science, Education, and Economics that are not characterized by such heavy web usage, exhibited a significantly higher degree of ethical misbehavior among their students. An additional result of the current research addresses the gradual and steady increase in the number of academic papers sold via the internet. This paper concludes with a short discussion on the impact of the results.

Keywords: academic integrity, information systems ethics, plagiarism

Introduction

Rapid advancements in science and technology have brought a wealth of products and applications into people's homes, giving rise to new ethical issues, ranging from access and intellectual property rights to individual dignity, privacy, and security (Petrovic-Lazarevic and Sohal, 2004). Most academic institutes have realized the importance of these ethical challenges and have extended their curricula to include courses about the various dilemmas as they relate to their specific learning disciplines. This is especially important for IS (Information Systems) students because of the interdisciplinary role IS plays in contemporary society,

* *Aharon Yadin* is a Senior Lecturer at the Max Stern Academic College of Emek Yezreel, Management Information Systems Department. Aharon's primary teaching areas are computer architectures and business/management information systems. Prior to entering the academic world, Aharon worked in the computing industry for over 30 years. Aharon is the author of four books and consults the European Commission on software related projects and technologies.

and the vast array of new ethical issues to be considered. This paper examines trends in ethical behavior over the past two years among students with a special emphasis on CS (Computer Science) and IS students as compared to other learning disciplines.

Contextual Framework

In spite of lecturers' tendency to cover more content, in recent years successful teachers have applied a different approach based on the understanding that technology is changing at an increasingly fast pace. As Wankat and Oreovicz (1998) defined it: «The hard part is getting them (the students) to learn how to learn.» In many technological disciplines such as CS and IS, students enjoy a wealth of currently available tools, applications, and solutions, as well as an ongoing offering of new ones. To cope with these ever changing environments, CS and IS lecturers have to concentrate more on helping students learn how to learn (Chindarsi, Spafford-Jacob and Miller, 2002). In order to struggle (successfully) with the diverse range of technical issues, students are encouraged to seek alternative solutions using any and all available sources of information. Looking for on-line help and using the web and its many discussion groups and forums, is a common practice among many students. The web is extremely popular and has become the ultimate mechanism for providing diverse solutions for many types of questions. Unfortunately, while the abundance of information and the user friendly yet sophisticated search engines available on the web provide an easy and accessible platform for legitimate technical help and other information, various types of unethical abuses are made more accessible as well, such as plagiarism of texts and ideas. Since CS and IS students are using the web more significantly, this research relates to their heavy web usage and the question to be asked is whether a correlation can be made between this extensive web usage and an increase in the tendency to plagiarize among students?

Not only do CS and IS students use the web more frequently, they also learn about a special type of ethics: the information ethics. According to Moore (2005), information ethics is the field that studies the ethical issues related to the development and implementation of various information systems.

The research described herein was conducted as a follow-up study to previous research (Yadin, 2007) on plagiarism and students' abuse of information systems, mainly for economic reasons. The current research repeats the analysis on new data accumulated over the last two years with an emphasis on trends in ethical misconduct especially among CS and IS students, as compared to students from other learning disciplines.

Theoretical Background

In what follows I provide a brief theoretical background of computer related ethical misconduct, specifically concerning students during their studies.

Over the years, many researchers have addressed computer ethics' issues and have suggested many explanations for the phenomenon of misconduct. More than twenty years ago, James Moor (1985) defined the «policy vacuum» as the main reason for ethical misconduct. This policy vacuum exists when there is no standard set of guidelines to govern a computer related situation. He referred to the uniqueness of computer technology that allows for changing algorithms and so creates new solutions for which no policy exists or has even been considered. Unfortunately over the years, instead of being narrowed, this policy vacuum has widened. Other researchers claim that the policy vacuum can not be completely filled and suggest additional approaches. Bynum and Rogerson (2004) propose a hierarchical model of applying individual judgment at the top of a defined set of core ethical policies. Gotterban and Rogerson (2005) define a process for assessing software development impacts and Martinson, Anderson and de Vries (2005) relate to protecting «the integrity of science... and look[ing] beyond falsification, fabrication and plagiarism, to a wider range of questionable research practices.» In a different paper, de Vries, Anderson and Martinson (2005) explain that «a certain amount of normal misbehavior is common in the dynamic field of science.» All these researchers and many more relate to the current state of ethical misconduct. Their attention represents steps in the right direction; however, they continue to rely mainly on individual ethical understandings and judgment, without establishing the means for checking and/or validation. This may arise from the fact that most research is aimed at more mature IT workers and not at students. Generally speaking, when addressing the student population, assuming a solid individual ethical position may prove to be extremely optimistic. Students lack experience in the field and their views regarding ethical situations have not yet been tested or honed. On the other hand, students are constantly exposed to and expected to follow their institute's academic code of ethics, which fills the policy vacuum at least during their academic studies. This suggests that students should be aware of the ethical issues that govern their learning activities.

This paper analyzes trends in students' ethical behavior by addressing plagiarism in higher education. Many ethical initiatives (courses and activities) aimed at student behavior are trying to fill the policy vacuum (Barroso and Siles Fernandez 2007; Gotterbarn 1999; Boylan and Donahue 2003). This paper focuses on an adjacent phenomenon: students who sell and buy written work via the web. Selling academic papers clearly demonstrates unethical behavior, supporting the claim that principled decisions based on an individual's previous judgment pro-

vide only a partial solution to the myriad ethical dilemmas prompted by internet accessibility.

The Study

The study described herein is a follow-up to a prior research project (Yadin, 2007) that was performed on data obtained from Smarter, an Israeli academic paper bank site (www.smarter.co.il). This site resells academic papers in Hebrew (essay, research and seminar papers, book reports, and bibliographic lists) that students have already submitted in their courses. It should be noted that this is not the only site in Israel engaged in this type of activity, but it was chosen because it is the newest, and without the installation of a «garbage collection» mechanism, potentially contains less outdated and irrelevant material. The site provides a technological plagiarism infrastructure by utilizing Information Technologies (a large database where students can upload their papers, a search engine for locating papers and a «royalties» bookkeeping system). For the purpose of this study and in order to collect data about the system's usage patterns and trends, the site was sampled twice over a two year period. The first sample was taken in August 2006 and the second in August 2008. In each sample the site was scanned and all the papers' attributes (metadata used to describe the contents) were downloaded from the site and entered into a database for better and simpler analysis. The attributes provided by the site that were downloaded included:

- (1) essay (or research paper) name
- (2) discipline (e.g., biology, art, information systems, e.tc.)
- (3) academic institute where paper was submitted
- (4) submission year
- (5) total number of words
- (6) number of secondary sources
- (7) price

Because the study focused on trends in higher education, only academic papers were targeted. All other types of documents available on the site were ignored. Obtaining the information (sampling the systems information and collecting the data) was straightforward since the system provided all the documents' attributes (metadata). The study phases included: downloading all the documents' attributes (data provided by the authors and freely available), filtering out irrelevant materials, deleting anomalies and documents with illogical attributes, and analyzing the remaining documents.

The main study objective was to analyze and highlight trends in unethical behavior among CS and IS students. Due to these students' needs to cope with complex and ever-changing technical issues, they are on a constant search for outside help. In addition, they are familiar with the tools available on the web, so it was estimated that they would demonstrate a higher degree of unethical behavior. An additional linked objective was to compare CS and IS students' ethical behavior trends to those of students from various other learning disciplines.

Results and Discussion

This academic paper site does not provide any information regarding the amount of papers downloaded; however, by analyzing the uploaded paper distribution and numbers, it was possible to draw conclusions. Examining, analyzing, and comparing the downloaded data, revealed some interesting observations regarding the ethical problem of recycling academic papers: how quickly it is spreading, and which learning disciplines demonstrated the highest and lowest activities.

The total number of relevant papers increased during the sampling interval by 325% (from 1740 papers to 7392). Figure 1 depicts the number of academic papers submitted each year and demonstrates that while there was a steady increase in the number of papers submitted in the past, 2007 may represent a different trend. It should be noted, however, that the year attribute defines the year the paper was submitted to the academic institute and not the year the paper was uploaded into the site. It usually takes additional time before a paper is uploaded into Smarter. The numbers for 2008 are not final since the sample was taken in August and the data is still being analyzed. Taking into consideration that there are over 200,000 students in Israel, the 2,000 papers uploaded in 2006 represents a small and manageable problem. However, we as educators must pay more attention to the overall trend, expressed in the graph for the years 2000-2006.

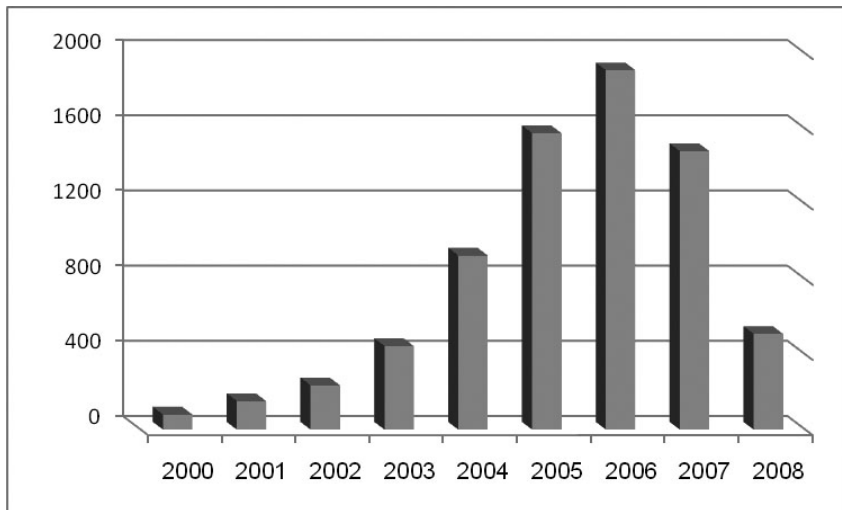


Figure 1: number of papers submitted per year

The average price requested per paper dropped slightly from 233 Israeli Shekels to 228 Israeli Shekels (roughly 60 USD or 45 Euros). This presents an additional temptation for those students who face failure in any particular course. The price of purchasing a paper has become less expensive than the tuition fee for the course and therefore an economic incentive for cheating has been fomented. The Higher Education Council in Israel controls all academic institutes and the tuition is identical in all. (The few private higher education institutes are the exception). During their studies, students are under regulations and various behavior guidelines and they know clearly that purchasing a paper and cheating by submitting it (or parts of it) are unethical. Instead of a policy vacuum, the student is facing a dilemma with real economic consequences. Various market fluctuations affected the paper price and in some learning disciplines there was an increase (Figure 2 depicts the learning disciplines with a price-per-paper increase of over 10%), while in others there was a decrease (Figure 3 depicts the learning disciplines with a price-per-paper decrease of over 10%). The data obtained provides insight about the price for papers in specific learning disciplines per each academic institute, as well as the changes in these prices, but the implications of these price fluctuations are beyond the scope of this paper.

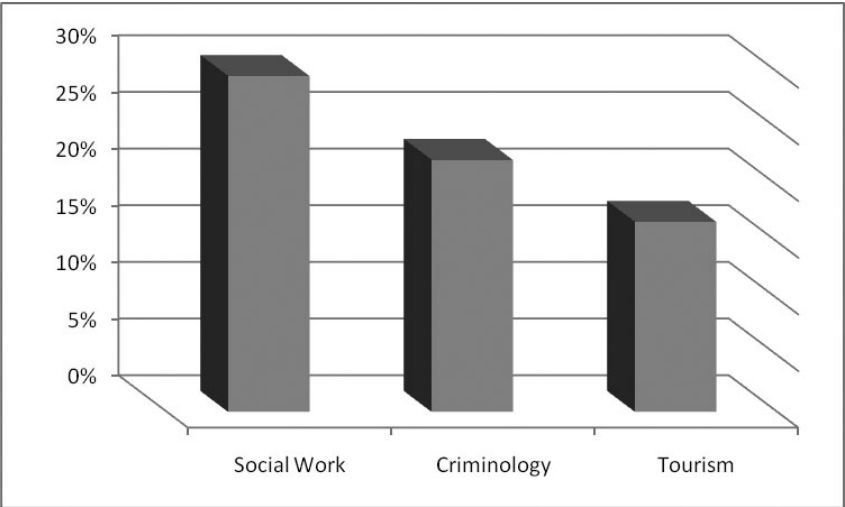


Figure 2: Price-per paper increase by learning discipline

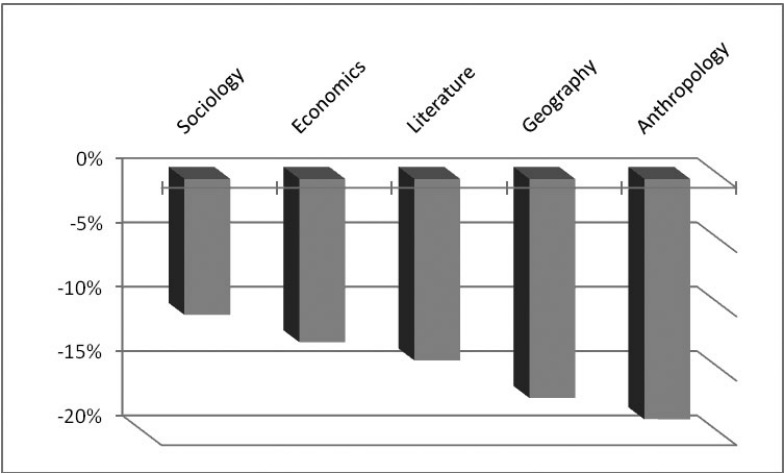


Figure 3: Price-per paper decrease by learning discipline

One of the interesting findings of the research is which learning disciplines demonstrated the highest increase in the number of papers submitted (or demonstrated the most unethical behavior). Figure 4 depicts the increase in the number of papers submitted during the sampling period for four learning disciplines. 100% means that during a two year period, between August 2006 and August 2008, the number of submitted papers doubled. The graph clearly demonstrates that CS and IS students exhibited a relatively low increase in the number of paper submitted (only 77%), while other learning disciplines, such as Political Science, which could be said to represent the country's future politicians and administrators, exhibits a 400% increase. Education, whose students will be involved in educating the next generation, exhibited a 387% increase, while Economics students exhibited a 367% rise in the number of papers uploaded. Interestingly and surprisingly, Law and Accounting were also high (around 250%). This contradicted my expectation that Law and Accounting students would display a higher degree of ethical conduct. The research question which tried to correlate heavy internet usage as part of the students' studies with a higher tendency toward plagiarism proved to be wrong. Course mandated exposure to the various web services and possibilities did not affect students' ethical behavior.

From analyzing Figure 4, one can realize that the policy vacuum is widening. In spite of the research and proposed new ethics' behavior models and ethics courses in some learning disciplines, this kind of moral misconduct is disturbing and calls for new ways of thinking and teaching.

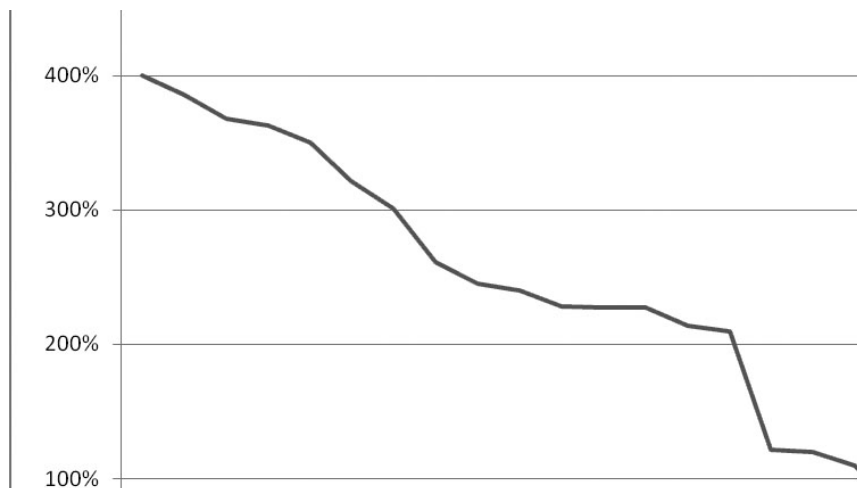


Figure 4: Increase in number of papers per learning discipline

This research strengthens the previous findings (Yadin, 2007) related to:

Course recycling in which identical term papers are submitted year after year (observed in some of the local branches of foreign universities that operate in Israel).

Duplicate papers which are papers that look identical in all attributes except one. Explanation for this phenomenon may be (a) a college student uploaded his/her paper and since it did not sell, s/he uploaded it again, but this time claimed it had been originally submitted at a more prestigious institute (a university instead of a college, for example) and (b) a student who bought the paper is trying to resell it as if s/he wrote it.

Paper writers who are not students (probably graduates) that upload their papers and are engaged in writing papers on various subjects for the purpose of selling them later. The important issue here is not the fact that there is a market for writers, but that people are willing to spend time researching and writing essays and other kinds of documents with the hope of selling them. This investment of time and effort implies that there are (many) buyers.

Concluding remarks

This paper examined the links between heavy web usage and the possible effects on ethical misconduct. The findings, based on a web site engaged in selling approximately 7,400 academic papers in its repository, proved that there is no direct link between the two. Unethical behavior is not related to heavy web usage and the wealth of application and capabilities the web provides. This misconduct should be correlated to other factors. Additional side effects of the research address the plagiarism phenomenon and the fact that websites which sell academic papers provide a temptation mechanism for misconduct in general and especially in cases of student failure of a course, when buying a paper seems the cheaper and sometimes the only alternative to salvaging the tuition fee paid for same course. This correlation has a serious impact on e-learning, the use of technology for teaching, and on the rise of cheating as well.

Although this research proved that there is no correlation between heavy web usage and ethical misconduct, the gradual increase of papers purchased from a website, as shown in Figure 1, is troubling. Students' current behavior is a good prediction for future behavior and the research does show that ethical behavior is not on the rise. In-class exams can provide the assurance that a student's work is his or hers alone. In all other cases the lecturers have to exercise extra caution due to the widespread availability of papers and their relatively cheap price. Ironically many research papers exist which deal with this troubling situation; as lecturers we have to ensure that we make proper use of them.

REFERENCES

- Barroso, P. and Siles Fernandez, A.** (2007). Treatment of Computer Ethics in ACM and ACM SIGCSE, CEPE 2007: Seventh International Computer Ethics Conference. University of San Diego, USA
- Boylan, M. and Donahue, J.** (2003). Ethics Across the Curriculum, Lanham, MD: Lexington Books.
- Bynum, T. W. and Rogerson, S., Eds.** (2004). Computer ethics and professional responsibility. Malden, MA: Blackwell.
- Chindarsi, K., Spafford, H. and Miller, J.** (2002). How can we teach students how to learn?, Teaching and Learning Forum 2002: Focusing on the Student, Mount Lawley, WA, Professional Development and Development Services, 11th Annual Forum: pp. 183-188.
- De Vries, R., Anderson, M. S. and Martinson, B.** (2006). Normal misbehavior: scientists talk about the ethics research, Journal of Empirical Research on Human Research Ethics, pp. 43-50.
- Gotterbarn, D.** (1999). Integration of Computer Ethics into the CS Curriculum: Attachment or Synthesis, SIGCSE Bulletin 31, 4, pp. 13-14.
- Gotterbarn, D. and Rogerson, S.** (2005). Responsible Risk Analysis for Software Development: Creating the software development impact statement, Communications of the Association for Information Systems, 15 Article 40.
- Martinson, B., Anderson, M., & de Vries, R.** (2005). "Scientists behaving badly". Journal of Nature 435, pp. 737-738.
- Moor, J. H.** (1985). What is Computer Ethics? Computers & Ethics, Ed. T. W. Bynum, Oxford: Blackwell, pp. 266-275.
- Moore, A., Ed.** (2005). Information Ethics Privacy, Property, and Power. Seattle, WA: University of Washington Press.
- Petrovic-Lazarevic, S., & Sohal, A. S.** (2004). Nature of e-business ethical dilemmas. Information Management & Computer Security, 12, 2, pp. 167-177.
- Wankat, P. and Oreovicz, F** (1998). Content Tyranny in October 1998 PRISM. Available: http://www.prism-magazine.org/october/html/october_teaching.htm accessed September 2008 <http://www.smarter.co.il> accessed August 2006 and August 2008
- Yadin, A.** (2007). Out of the strong, something to eat (Book of Judges 14:14), in the proceedings of I-Society 2007, International Conference on Information Society, Merrillville, Indiana, USA

Curia Novit Google: The Quest for Law in the Era of Internet Search Engines

Georgios N. Yannopoulos*

Law Faculty, National and Kapodistrian University
of Athens, gyannop@law.uoa.gr

Abstract:

Finding legal information is much easier in the era of Internet search engines and electronic legal databases. However, it can be argued that those seeking law rely heavily on “Googling”, overlooking the traditional, and sometimes still official, sources of law. The known maxim *iura novit Curia* (the Court knows the laws) could be paraphrased to *Curia novit Google*¹: those in search of legal information recognise as “law” only the results of the search engines or of the search function of electronic databases. Is that right in terms of the rule of law doctrine? Is it ethically correct? The present paper tries to address those questions.

Keywords: Law - Rule of Law - Legal Databases - Information Retrieval - Search Engines - Interpretation

From the Law of the Text to the Law of Google

Legislatures, courts and administrations, academic and professional lawyers produce gigantic amounts of legal information in jurisdictions around the globe. Since the 60's, we have evidenced the operation of a large number of legal information retrieval systems (Bing and Harvold 1977), mainly in the form of dedicated legal data bases (e.g. Lexis, Westlaw e.tc.). Following the enthusiasm of the 70's for the benefits of information technology, the Internet appeared in the 90's as the *deus ex machina* solution for the information crisis in the area of law. A vast

* *Georgios Yannopoulos* is a Lecturer at the Law School of the University of Athens, teaching Law and Information Technology. He is a graduate of the same School (1988) and has professional qualifications in Computer Programming (1984-86) and Systems Analysis (1987). He has accomplished his PhD (1996) and post-doctoral research (1998) at the University of London. In 1996-97 he was a Senior Research Fellow at the University of London. He is a Lawyer, admitted to the Athens Bar Association and has worked as in-house counsel in telecommunications and currently in banking.

1. Trademark of Google Inc.

number of legal documents from the older systems (e.g. Gopher and FTP) have been transferred to the World Wide Web, leading to dedicated legal indexes that would later be connected to the known search engines (Zizzo 1995). Apart from dedicated databases, such as Lexis and Westlaw, available for a fee, people seeking legal information, all over the world, may now access the same piece of information at the same cost through the Web. Various private initiatives around the globe, such as Legal Information Institutes (LIIs, see Mowbray et. al. 2007; for counter arguments see Bruce 2000a) have fostered the idea of dispensing legal information for free, mainly by collecting and updating links to such information in order to provide a comprehensive research facility. This usually takes the form of aggregators of legal information at national or international level.

The main legal argument is that ignorance of law or difficulty in searching for legal information is not a simple matter of professional deficiency of a lawyer, but a direct hit to the rule of law doctrine: in a democracy, citizens should have access to the law. The term “legal information”, should be considered in an expanded version in order to cover all aspects of legal documentation, i.e. laws, case-law, legal literature and in some instances further government information (e.g. registries and records), while the search for legal documents such as contract templates and blueprints of agreements cannot be ruled out. As a result, a number of countries and international organisations, following state initiatives, have abolished the Lex Gambetta doctrine, i.e. the publication of laws in an Official Gazette (Yannopoulos 2006, Kirchberger 2007) and they disseminate the text of the laws through the Internet using an “official” website, for free or for a fee. In the same vein, legislative bodies have introduced special regulations for the electronic dissemination of case-law, i.e. decisions of courts and other official bodies, while Directive 2003/98/EC on the re-use of public sector information has officially verified in the EU the idea of freely distributing the original material without exclusive rights and restrictions.

The development of that immense amount of legal information, found on a vast number of official or unofficial legal Websites around the world, has led to the reality that lawyers, in the quest of law, normally start their search by using the well known internet search engines like Google, Yahoo etc (reference to “Google” designates the totality of search engines). In technical terms these search engines use other programs, called “robots”, “crawlers” or “spiders”, which crawl the web in order to find other web pages, by assessing links from indexed webpages, or by links submitted manually. The information is then indexed in order to rank what counts as relevant and what not. The ranking of the results and hence the presentation to the user depends on the particular ranking algorithm that each search engine uses (for the solutions followed by Google see Brin and Page 1998 and

[www.google.com / corporate / tech.html](http://www.google.com/corporate/tech.html); for alleged manipulation of such data see Van Eijk 2006). The use of dedicated legal search engines has been proposed as a solution to the problems of general purpose search engines (Greenleaf et. Al, 2000, with a detailed description of the defects of search engines for retrieval of legal information) but has attracted limited attention especially in non-English speaking jurisdictions.

The new medium is accessible to everyone, but also anyone, despite of background or qualifications, may post a variety of legal information on the web. To the contrast of other traditional media, that may as well disseminate inaccurate information, it can be observed, however, that researchers rely heavily - if not exclusively - on the results produced by the above search engines. A new expression is used: "...to Google..." (Van Eijk 2006, Kirchberger 2007), just to prove the immense trust of researchers in using such engines and consequently in the relevance ranking and in the search results. The painless use of a search engine, acting as a hypothetical comprehensive "one-stop shop", has led to the abandonment of the traditional, and sometimes still official, sources of law. It is often the case that "law" is considered only what Google (or any other search engine) produces as such. In that sense, the known maxim *iura novit Curia* (the Court knows the laws or precisely: it is upon the Court to search for the Law) could be paraphrased to *Curia novit Google*: The "Court", i.e. broadly those in search of legal information, recognise as law only the results of the search engines or the outcome of the search function of electronic databases. To paraphrase another known motto: Law that does not appear in the first results page of Google does not exist: "to exist is to be indexed by a search engine" (Introna and Nissenbaum 2000, Schulz et. al. 2005). In the next sections, I am trying to address the antagonism of such practice to the rule of law doctrine, without, entering into the technical details of how to build a proper web search engine dedicated to the search of legal information. While, the legal position of search engines and issues such as data protection and intellectual property cannot be overlooked (see Van Eijk 2006, Schulz et. al 2005, Bercic 2004 and relevant trade-mark and meta-tag cases of the 90's), the crucial legal and ethical question here does not relate to the politics of the search engine, but rather to the attitude of the law-researcher: should she/he rely only on data indexed by search engines?

A Question of Authenticity

In view of the above vast amount of information in the hands of lawyers, there is limited official guidance (or regulation) as of which sources should be used. Sometimes, courts are obliged to "connect to electronic legal data banks", without specific indication as of which database is the official one. Greater legislative

bodies, such as the European Union, have introduced specific rules (see Council resolution of 20 June 1994 on the electronic dissemination of Community Law and national implementing laws on improved access conditions, OJ C 179/1-7-1994) for the electronic dissemination of legal information especially by improving access to and expanding the use of public sector information (Leith and McCullagh, 2003). Nonetheless, users, in a typical search utilise the database or the search engine that suits best their own subjective criteria. In all cases, no one challenges the issue of authenticity: Firstly, in terms of verifying that the information is undoubtedly originating from an official source: legislative body for the text of the laws, judge for decisions, author for legal literature, and secondly, that the information has not been altered or distorted during the electronic transmission. It is obvious that such assessment is critical if the information is used by a law-giving authority in its function of applying justice. It seems that the use of electronic signatures and cryptography may offer significant help in tackling this problem: In the first issue by confirming the true source of the material and in the second issue by proving that what has been received from the authentic creator has not been altered by the intermediaries, in any way. Regardless of the use of electronic signatures, the name of the institution and / or organisation, affiliation should clearly appear on the site including the names of the authors and other relevant details.

A Question of Reliability

Among the cognoscenti, there seems to be a tacit understanding that the creators of such databases and websites are self-conscious enough to provide trustworthy data (Spyridakis 2005); until the contrary is proven such entities hold the benefit of the doubt, i.e. that legal data are reliable, current and trustworthy. Although methods for assessing the substantive quality of legal information in websites have been proposed (Robinson, 2000), little examination of the reliability of private databases and collections in the following three aspects has been made:

Firstly, in the aspect of formal legal quality, i.e. by ensuring that the content is indeed the “official” text and that is updated to reflect recent changes. The “currency” function is crucial to the everyday work of a practitioner applying law and, therefore, under scrutiny when the consolidation function is undertaken by a private entity (for a detailed analysis of the public v. private benefits see Bruce 2000a). It has been argued (Bruce 2000) that the virtues attributed to private entities in relation to publishing and consolidating reliable legal texts are only accidental and reputational, while there is no objective assessment of quality factors among the online commercial publishers. A different degree of reliability could be achieved when the consolidation/currency function is undertaken by an official body (e.g. the Eur-Lex service of EU).

Secondly, in the aspect of technical quality, by ensuring that the search function of the legal material is convenient and that the user interface is helpful and user-friendly.

Thirdly, in the aspect of quantity: to be precise, in examining and verifying the extent of coverage of legal information. The issue has been resolved since the 90's when full-text technical methods have been adopted for both laws and case-law. In that sense, the user should validate if strict objectivity has been observed, as the technology is the only means for transparency: the database must, as a matter of principle, try to cover the totality of laws and decisions of a certain legal area and the user should be, at any time, aware of the full capabilities of the system. A rating system marking legal websites according to the degree of coverage (Robinson, 2000) would be appreciated, but it would not be able to cure inherent deficits of the users described in the next section.

It could be anticipated that official sources should produce and identify their own keywords, to be included in electronic indexing systems and such keywords should be in turn followed by the search engine ranking algorithm. It has been argued (Bruce, 2000) that while in traditional legal publishing, Librarians indexing cases may categorize decisions in ways that may hinder or distort their meaning or usefulness, these problems have been solved by the use of full-text retrieval techniques. However, unofficial sources still burden the task of an "unofficial" interpretation through indexing and users should be aware of that. Furthermore, the appearance and the punctuation of a legal text in electronic form have been "blamed" (Bruce, 2000) for misinterpretation of its true meaning. Equally, for web presentations, an effort to follow the text of the printed version should be made, notwithstanding the tendency in many jurisdictions to adopt only the electronic version of a text.

A Question of Search Techniques

Studies in the 80's have concluded that users in search for legal information only retrieve a small proportion of the material they really want (Bing 1987, Blair & Maron 1985). Modern search engines have adhered to traditional retrieval systems and in most cases continue to use the Boolean techniques, i.e. the combination of words or expressions to be retrieved by means of the conjunctive 'AND', the disjunctive 'OR' and the negative 'NOT' of the algebra developed by George Boole in 1847. This combination is then used to match the entries of an existing index of words or thesaurus, as it is called. The search technique used by the web search engines is not very different from the one used in older retrieval systems: "Robots" and "spiders", mentioned above, create an index using as "keys" the words found in the Uniform Resource Locators (URLs) of the webpages. The

visible improvement is the addition of the statistical counting of the words' frequency and the ranking of the documents (Schweighofer, 1999).

Notwithstanding these shortcomings, in the case of web searching two more problems arise: First, the user may find difficult to modify her / his search for something, such as Google, that everybody uses (Kirchberger 2007). Second, the ranking algorithm, mentioned above, could be "blamed" for bringing to the top "irrelevant" information and thus leading the non-meticulous and sometimes lazy and less-critical researcher to ignore data appearing in a lower position or in the next page. From the legal point of view, the end user is responsible for finding the legal material to support her / his argument. Such obligation arises not only as a matter of regulation, but also as a basic principle of professional ethics. Typical disclaimers placed in the websites indemnify the providers of the service for any liability; however, such disclaimers would neither rouse the lazy researcher from lethargy nor provide the search engine with a serious legal research facility (Mowbray et. al. 2007).

The "curse of Boolean Algebra" (Bing 1987, 1989) notwithstanding the acquaintance of users with the functions of the operators, translates a proper legal argument to only one word or two or more words combined with AND. Clearly, this has been insufficient for former electronic databases and can still be inadequate in the case of web engines where the user has to submit the query more than once to achieve proper results. It has been proposed (Yannopoulos 1997) that closed thesauri of predetermined terms must be abandoned in favour of conceptual techniques that should replace the "blank line of Google or Yahoo" with intelligent "front-ends" (see several projects in Kirchberger 2007, Quaresma & Rodrigues 2000 and the May 2009 announcement of Google to use semantic web technology). Such techniques would be able to represent the full legal argument and transform static legal information into functional legal knowledge and enhance the benefits for users.

In addition, the system should be transparent in letting the user understand (and view) the search and ranking technique used, in order to evaluate the results, to adjust the query and finally to become more critical (Kirchberger 2007). Such transparency obligation may become the object of legal regulation (see, for example, Schulz et. al. 2005 in connection to art. 6 of the e-commerce Directive 2000/31 for commercial advertising), but we should keep in mind that legal information retrieval is merely a skill that has to be taught to lawyers. Amongst other proposals, the paradigm brought up by the retrieval system itself could become an appropriate method.

Be informed: Google non novit ius

The use of the Internet to find law conforms to the ideal vision of the Web as a democratic space and a platform for social justice (Introna and Nissenbaum 2000). Those seeking legal information may have essential access to a large collection of legal materials that would otherwise be inaccessible and, finally, they are able to make better and quicker decisions. However, a valid legal decision should be based on the doctrine of valid legal sources (Yannopoulos 1997), which is a system of clearly defined rules following an established hierarchy. Legal researchers working under that principle should abide by the rules of the legal system and professional ethics and, in that sense, they should place under scrutiny the results produced by search engines.

It is beyond doubt that search engines and electronic legal databases present advantages and benefits for law-researchers and have dramatically changed the shape of legal research. Nevertheless, law-seekers fail to take into consideration that applying law is not simply a matter of making a query, which may be narrowed or broadened by the subjective ranking of the results by a search engine. To the eyes of a lawyer the idea of dissemination of justice is not merely law-finding. Seeking justice, notwithstanding the perfection of the technical means, does not discharge jurisprudence and lawyers from their obligation to make the most of their legal background and, each time, to interpret the law and legal sources. The seeker-interpreter should always be aware that *Curia novit Google sed Google non novit ius*.

REFERENCES

Berčič B., (2004). Search engines: A useful tool but are they a lawful practice? BILETA 2004, 19th Annual Conference: University of Durham, 25th & 26th March 2004.

Bing J. and Harvold T., (1977). *Legal Decisions and Information Systems*, Universitetsforlaget, Oslo.

Bing J., (1989). The Law of the Books and the Law of the Files, Possibilities and Problems of Legal Information Systems, in Vandenberghe (Ed.), *Kluwer, Computer/Law Series*, 180. Also appeared in *Computers and Law*, Vol. 54, 31.

Bing J., (1987). Designing Text Retrieval systems for 'Conceptual Searching, in *Proceedings of The First International Conference on Artificial Intelligence And Law*, ACM Press, Boston.

Blair D. and Maron M.E., (1985). An Evaluation of Retrieval Effectiveness for a Full-Text Document Retrieval System, in *Communications of the ACM* Vol. 28, No. 3, 289.

Boole G., (1847). *The Mathematical Analysis of Logic*, Cambridge.

Brin S. and Page L. (1998). The Anatomy of a large Scale Hypertextual Web Search Engine, Proceedings, The Seventh International World-Wide Web Conference, April 14-18, 1998, Brisbane, Australia, in *Computer Networks and ISDN Systems*, Vol. 30, article FP-11.

Bruce T., (2000a). Public Legal Information: Focus and Future, *Journal of Information Law and Technology*, 2000 (1).

Bruce T., (2000). Tears Shed Over Peer Gynt's Onion: Some Thoughts on the Constitution of Public Legal Information Providers, *Journal of Information Law and Technology*, 2000 (2).

Gelbart D. and Smith J.C., (1994), The Application of Automated Text Processing Techniques to Legal Text Management, in *Yearbook of Law Computers and Technology*, Vol. 8, 208.

Greenleaf G., Austin D., Chung P., Mowbray A., Davis M. and Matthews J. (2000). Solving the Problems of Finding Law on the Web: World Law and DIAL, *Journal of Information Law and Technology*, 1.

Introna L. and Nissenbaum H., (2000). Shaping the Web: Why the Politics of Search Engines Matter, in *The Information Society* Vol. 16, No. 3, p. 169-186.

Kirchberger C., (2007). Paper and stone: How technology has not changed the retrieval of legal information, yet. BILETA, 2007 Annual Conference, Herfordshire 16-17 April.

Leith P. and McCullagh K. (2003), Developing European Legal Information Markets Based on Government Information: First Findings from the Add-Wijzer Project, 18th BILETA Conference, QMW, London.

Mowbray A., Greenleaf G., Chung P. and Austin D., (2007). Improving stability and performance of an international network of free access legal information systems, *Journal of Information, Law and Technology*, (2).

Morison J. and Leith P., (1992). *The Barrister's World and the Nature of Law*, Open University Press, Milton Keynes - Philadelphia.

Quaresma P. and Rodrigues I., (2000). Automatic Classification and Intelligent Clustering for WWW Information Retrieval Systems, *Journal of Information, Law and Technology*, 2.

Robinson A., (2000), Creating a safety Net - A proposed Rating Form for Assessing the Quality of Legal Information in Websites, *Journal of Information, Law and Technology*, 1.

Schweighofer E., (1999). The Revolution in Legal Information Retrieval or: The Empire Strikes Back, *Journal of Information, Law and Technology*, *Journal of Information, Law and Technology*, 1999, 1.

Schulz W., Held T. and Laudien A., (2005). Search Engines as Gatekeepers of Public Communication: Analysis of the German framework applicable to internet search engines including media law and anti trust law, *German Law Journal*, Vol. 06, No. 10, p. 1419-1433.

Spyridakis I., (2005). Lawyers against Computers (in Greek), in *Nomiko Vima*, Vol. 53, 816.

Van Eijk N., (2006). Search Engines: Seek and Ye Shall Find? The Position of Search Engines in Law. *European Audiovisual Observatory*, Strasbourg.

Zizzo, F., (1995), Legal Resources via World Wide Web, 10th BILETA Conference, Business School University of Strathclyde, Glasgow.

Yannopoulos G. N., (1997). *Modelling the Legal Decision Process for Information Technology Applications in Law*, Kluwer, Computer / Law Series.

Yannopoulos G.N., (2006). The Right to Access Electronic Legal Databases in Greece, in *Flogaitis - Karpen - Masucci (Eds), e-Government and e-Democracy*, EPL Series, vol. LXXXI, Esperia Publications, London, p. 295.

Constraints to the Application of ICT Implants: The Concept of Self-Ownership

Karsten Weber*

Professor of Philosophy, University Opole, Poland

Preface

Those who refuse to accept the idea of ICT implanted in the human body argue that the main purpose of such technology will be surveillance and control and therefore will bring harm to people who are implanted. One can summarize this line of argumentation by claiming that ICT implants would vaporize privacy and diminish autonomy of the implant carrying person. Of course, at least in the western tradition of thinking, one can make a case that privacy and, particularly, individual autonomy are most important for the concept of a person. Therefore, so the argument goes on, autonomy and privacy must be protected and, conclusively, ICT implants must be forbidden. Yet, those you argue in favor of implants often say that in a liberal society people must be granted to deliberately decide on their own whether they would like to accept to be implanted.

Of course, this a very rough summary and it does not represent the complexity of argumentation on both sides. However, in what follows it shall be presumed that both sides are wrong in conceptualizing surveillance and control. While the one side stress that surveillance and control are nothing but threats to autonomy and privacy, the other side even does not take into consideration that there is something like a society which can be negatively affected by individual decisions.

In contrast to that, here it shall be presupposed that there are conditions and circumstances in which surveillance and control can be acceptable both from an individual and from a societal point of view. However, there will be no definition or list of situations in which surveillance and control are acceptable. Instead, the aim of this paper is to find some constraints to the application of ICT implants. Actually, one single constraint shall be discussed: The concept of self-ownership.

* Dr. phil. Karsten Weber currently is Visiting Professor for Information Ethics and Data Protection at Technical University Berlin, Germany. He also is Professor for Philosophy at the University of Opole, Poland, and Adjunct Professor for Culture and Technology at Brandenburg Technical University Cottbus, Germany. He works on Information Ethics, Political Philosophy, and Philosophy of Science.

Thus, the argumentation in this paper strongly emphasizes individual rights and opposes, for instance, utilitarian ideas.

If one takes a look on the existing literature one will learn that self-ownership almost exclusively is discussed with respect to distribution of wealth and primary goods. That means that self-ownership is used as argument in debates on distributive justice. Predominantly, Robert Nozick employed this line of argumentation in his book "Anarchy, State and Utopia" published in 1974. He wrote that book in reaction to John Rawls' hallmark book titled "A Theory of Justice" which was published three years earlier in 1971. Since these times, a quite controversial scholarly debate continuously is going on with regard to the question how a society must be shaped to be a just society. It therefore seems that self-ownership is an inadequate starting point, for there are no strong ties connecting the concept of self-ownership on the one side and the application of ICT implants on the other side. But the debate concerning the just distribution of primary goods (see for this issue, for instance, Nock 1992) will not even be touched in the paper at hand.

However, it is most obvious that the concept of self-ownership was deployed to strengthen individual autonomy and individual rights as constraints to state's and society's interference with a person's course of life. In addition, it is most apparent that the use of ICT implants and the modification of one's own body are already quite common with regard to medical applications. And since ICT implants will be perfect means for control and surveillance the question whether it is possible to find or define constraints is most important for individual autonomy, rights and the opportunity to choose one's own way of life. Thus, the concept of self-ownership may be worthwhile to look at.

Why not another Position in Political Philosophy?

Although it is not the aim of this paper to present the whole range of positions in political philosophy and the respective debates, something must be said about the choice of libertarianism as starting point of the argumentation. One could ask: Why not liberal egalitarianism as in John Rawls or Ronald Dworkin? Why not communitarianism like in Michael Walzer? The answer concerning the latter is easy to present: Most if not all communitarian scholars argue in favor of the common good as being superior to individual rights. Therefore, from a communitarian point of view, one can argue for the compulsory implantation of ICT devices if this would be a means to provide for the common good (cf. Weber 2006).

The case of liberal egalitarianism is a bit fuzzy. Particularly John Rawls concept of justice as fairness seems appropriate to deal with questions posed by ICT implants because he rigorously, as Nozick, argues against utilitarian ideas:

“Each person possesses an inviolability founded on justice that even the welfare of society as a whole cannot override. For this reason justice denies that the loss of freedom for some is made right by a greater good shared by others. It does not allow that the sacrifices imposed on a few are outweighed by the larger sum of advantages enjoyed by many.” (Rawls 1999: 3/4)

However, liberal egalitarian scholars like Dworkin, and Rawls is among them, aim towards equality of resources (cf. Moulin, Roemer 1989: 349; Christman 1991); simultaneously, ICT implants might create the possibility that mental resources of human beings, particularly cognitive capabilities, directly could be used to equalize the distribution of such resources among people. Although today that sounds like Science Fiction, liberal egalitarianism does not seem to provide for an answer concerning this possible challenge.

Therefore, libertarianism emphasizing self-ownership as an absolute constraint to infringements by state and society seems to be appropriate to discuss ICT implants and their challenges to justice.

Some Definitions

At first, it is necessary to clarify the meaning of some words which are often used in relation to the concept of self-ownership. If one talks about ownership, one has to say what is owned. In general, in this case the word “property” is used, but one has to learn that “property” has two different meanings. “Property” can be used as synonym instead of words like “attribute”, “capacity”, “feature”, “quality”, “trait”, “characteristic”, and the like. It is used to say that a particular entity has certain traits as in sentences like “Grass is green”, “Snow is white”, “A rectangle has four corners”, and so on. These sentences can be reformulated to “Green colour is a property of grass”, “White colour is a property of snow”, or “To have four corners is a property of rectangles”. For the purposes of this text one can say that in this sense property can be understood as a relation of an entity and its traits.

At the same time property has something to do with ownership. To make things as easy as can be one can say that in this sense property can be understood as a relation of a person and an entity. One also can say that it is a relation of a subject and an object. Of course, one need to clarify that kind of relation and the following quotation might help:

“Property is that which a man has the right to use and enjoy without interference; it is what it makes him as a person and guarantees his independence and security.” (Tay 1978: 10)

Property in the sense of ownership can be understood as means to support individual autonomy, particularly as means to protect a person against other persons’

and the state's infringements. Conclusively, property as the right to exclusively use certain physical spaces might be a necessary – but not sufficient – prerequisite for privacy and the distinction of private and public sphere. But here, this issue shall not be discussed further.

To understand property in the sense of ownership it is evident what self-ownership shall mean: A person owns herself. To adopt the above mentioned quotation:

Self-ownership is a person's right to use and enjoy herself and her body without interference; it is what makes her as a person and guarantees her independence and security.

Self-Ownership, Subject and Object

Self-ownership implies that a person will be subject and object at the same time because the relation of ownership or property just needs a subject and an object: Owner and owned, proprietor and property. But can a person be subject as well as object? For instance, Immanuel Kant would deny that rigorously for certain reasons. His whole line of argumentation with regard to moral imperatives would crash if persons could be objects. For if persons could be objects they would have a price and therefore could be sold and bought. But for Kant persons do not have a price but dignity which is priceless. Roughly said, this dignity makes them moral beings. Additionally, conceiving persons as objects would imply to look at them as mere means and not as ends; again, this would conflict with Kant's moral imperatives.

Nevertheless, libertarians like Robert Nozick use the concept of self-ownership for their line of argumentation and at the same time, often refer to Kant and his moral imperatives. Therefore, they need a solution for the subject-object contradiction. And this solution is really obvious and simple to recognize if one takes into account the history of philosophy and its core ideas. The solution is dualism: A person is conceived as a combination of distinct entities with totally different characteristics. These entities are body and mind, *res extensa* and *res cogitans*. Now, employing the concept of self-ownership does not create a contradiction. The subject in the relation of self-ownership is identified as the mind, the object is the body – a person has a body or owns a body, the mind is the proprietor of the body as property. This creates a hierarchy of mind and body: Mind is superior to the body.

The development of neuroscience and theories stating that mental processes could be explained merely with reference to physical and chemical processes in the human brain and central nervous system of course put this dualism into question. However, such problems shall not be discussed here. On the level of social

interactions sometimes we might talk in dualistic terms but it shall be assumed that regularly this does not really affect our behavior. Equally, the issue of free will shall not be discussed here. Liberals and libertarians always stress individual autonomy and freedom; obviously one could ask whether they must presuppose free will or merely freedom of action (cf. Glannon 2005). But again, although these are important and interesting questions, at least for philosophers, they shall not be discussed further.

For the purposes of the text in hand it is just important to note that the hierarchy of mind and body can be used as argument in favor of ICT implants and body modifications. If the real self is the mind than any kind of modification of the body does not affect the person itself but only its physical carrier. But at the same time, the dualism of mind and body might as well provide an argument at least against some technical interventions which could alter the structure of mind and therefore might inflict the real self. However, this will not be the line of argumentation used in what follows.

Constraints Imposed by Self-Ownership

To keep in mind what the concept of self-ownership shall provide for our discussion, let us quote from the first part of Eric Mack's text "Self-ownership, Marxism, and egalitarianism", in which he says that self-ownership is

"[...] the thesis that each individual possesses original moral rights over her own body, faculties, talents, and energies. Adherents of this thesis believe that it best captures our common perception of the moral inviolability of persons — an inviolability that is manifested in the wrongfulness of unprovoked acts of killing, maiming, imprisoning, enslaving, and extracting labor from other individuals. They believe that the rights of self-ownership provide individuals with the moral immunities appropriate to beings whose lives and well-being are of separate and irreplaceable moral importance." (Mack 2002: 76)

While reading Mack's essay, again, one has to learn that scholars arguing with reference to self-ownership particularly are focussed on property rights; they are concerned with the question of just distribution of wealth and other primary goods. But in Mack we also can find some hints that self-ownership could be used to argue for protection of persons against interference with their bodily integrity. In the above cited part of his text, Mack himself mentions maiming as violation of self-ownership and he refers to the well-being of humans, too. Therefore, one can argue that self-ownership sets absolute constraints to interference with a person's bodily integrity. It says that nobody, neither other persons nor the state or society or community, is entitled to violate a person's bodily integrity as long as this person refuses to consent to this infringement.

Autonomy, Paternalism, and Harming Oneself

Conclusively, something has to be said about autonomy and decision-making. An important aspect of self-ownership is that every person is entitled to decide autonomously about the course of her own life. Clearly this includes the right that this person decide that others are allowed to alter or modify her body. Since the body is nothing than property that is owned by a person, that person is entitled to do everything she wants to do with her property. Again, one can see the far-reaching impact of the above mentioned hierarchy of mind and body.

Today, tattoos or piercings are quite common examples of body modifications; with regard to self-ownership, the implantation of ICT devices into a person's body seems to be just a new instance of the same kind of decision. Of course, we all know that in the near past things were different: Just a few years ago, tattoos, piercings and even earrings for men often caused aesthetic or even moral concern. It is quite likely that during adolescence at least some of us found themselves in trouble when they came back home with an earring or something alike. As minors or teenagers our autonomy was strongly limited – actually, the concept of self-ownership is silent with regard to this problem. But today, things changed at least for adults – piercings and tattoos are commonly accepted; sometimes and in some subcultural contexts, they even seem to be a must-have.

Self-ownership implies that it is the person's decision to shape her body as she likes to – other people, communities, society or the state are not entitled to interfere with this decision, neither because of moral nor aesthetic or other reasons. But this freedom has a far-reaching implication: The autonomous person is free to take any decision she wants to – of course only as long as no other persons will be harmed. As Richard Fallon (1994: 878) puts it: “[...] autonomy represents [...] their right to make and act on their own decisions, even if those decisions are ill-considered or substantively unwise.” But at the same time a person has to bear all consequences of her own decisions; she is fully responsible for the outcomes of her decisions and she must bear them all alone. One may even harm oneself – nobody is entitled to interfere with this kind of action. As Robert Nozick stresses,

“[i]ndividuals have rights, and there are things no person or group may do to them (without violating their rights). So strong and far-reaching are these rights that they raise the question of what, if anything, the state and its officials may do. How much room do individual rights leave for the state?” (Nozick 1974: IX)

After asking this question, he answers it only a few lines below:

“Two noteworthy implications are that the state may not use its coercive apparatus for the purpose of getting some citizens to aid others, or in order to prohibit activities to people for their own good or protection.” (ibid.)

However, the concept of self-ownership allows for that we may feel a strong moral obligation to help those who are in trouble, even if we know that those persons in need caused this mess by themselves. But if those in trouble may think that they do not need or want our help, we are not entitled to force them to accept our help – for those convinced in self-ownership, any kind of paternalistic interference is strictly forbidden (cf. Scoccia 1990); self-ownership entitles a person to self-mutilation or even to commit suicide (Coleman 2005: 135).

Simultaneously, nobody is entitled to enforce us to help others – if we do not feel obliged to help others it is not allowed to force us to help. Particularly the state with its coercive potential must stay morally neutral; libertarians like Nozick stress that only

“[...] a minimal state, limited to the narrow functions of protection against force, theft, fraud, enforcement of contracts, and so on, is justified; that any more extensive state will violate persons’ rights not to be forced to do certain things, and is unjustified; [...]” (Nozick 1974: IX)

Yet, this sounds like every person is queen in her own kingdom of body. Nobody is entitled to force a person to accept ICT implants; only the person herself is entitled to decide whether she would like to accept an implantation. But it is most important to stress that this is not the whole story.

To Abandon Freedom and other Problems

Actually, from Nozick’s point of view freedom could be used to abandon freedom. In “Anarchy, State and Utopia” (Nozick 1974: 331; cf. Coleman 2005: 127), he argues that one is even allowed to sell oneself into slavery. Since a person is proprietor of her body she is entitled to do with it what she wants to do – selling her body into slavery is only one option among others. More generally speaking freedom implies that a person is free to choose to live in communities or societies which are not free. If one consents to incorporate ICT implants for surveillance purposes or which even will control his life by direct interference with his actions, from a libertarian point of view, others are not entitled to stop that person. For this would be paternalism.

Nozick himself stresses that his point of view might be quite radical and that others would not agree to it. For instance, one could argue that a person must be allowed to abandon freedom only to that extent that there must always be an exit option. This would mean that a person is entitled to subordinate herself to a restrictive regime as long as at any time that person can exit this community without being restrained to leave, without being punished or harmed. But here, we would face some severe problems. The first one is that the respective com-

munity could argue that to let go everybody who wants to go would destroy the community. For this would cause harm to other members of the community, conclusively exit must not be allowed – in fact, some communitarian scholars would argue that way.

The second problem is the question of how to treat minors and teenagers. If children are raised up in a repressive community it is very likely that they are not aware of those rights self-ownership is supposed to protect. Therefore, even if such a repressive community would grant the right to leave the community when people became adults, they might not be able to autonomously decide whether to exit or to stay.

Net Effects

Although those two problems just mentioned above are very difficult to solve, these are not the most pressing ones: There is a characteristic of concepts like self-ownership which heavily collides with current and coming real-life applications of ICT technology. We are not talking about simple computers or other ICT devices which can be examined as isolated items. Current ICT devices and therefore ICT implants are highly integrated into networks of other artifacts and into social networks of humans – actually, talking or worrying about one single ICT implant does not make any sense. The integration of ICT implants in a large infrastructure means that individuals, institutions, technology, and the like are highly meshed.

Even if we take it for granted that due to self-ownership it shall be up to a person to decide whether she would like to use ICT implants to join a certain group or to comply with the demands of a particular community, this may have the consequence that it is no more the autonomous decision of that person to leave a group or community. For the exit may cause harm to others: For instance, let us suppose that members of a certain community carry an implant that makes it possible to experience what other persons feel. It is very likely that some of the community members, perhaps even all, will develop a kind of dependence or addiction to these transmitted sensations. Obviously, as all addicted people, they will be harmed if they are not able to get what they are addicted to.

Or, to be more realistic: According to libertarians like Nozick the state has the function, and this is its only function, to protect its citizens against force, theft, fraud, and so on as well as to enforce contracts and the like. That means that a minimal state is not a defenseless state – for instance, it might heavily use technology to provide protection and security. Now let us presume that a certain kind of ICT implants could be used to measure the level of aggressiveness of the implant carrying person. If and only if all citizens would carry an implant of that

kind, such a network of implants, other devices, and law enforcing institutions adequately could serve as a viable means of protection.

Thus, it would be questionable whether the concept of self-ownership could protect persons against the compulsory implantation of such a device. Although Nozick and civil rights libertarians probably would argue that the utility of citizens' majority do not provide reasons or justifications for the state and its institutions to be allowed to interfere with the bodily integrity of a single person, for this would be interference with this person's rights, Nozick and other libertarian scholars then would have to admit that the state and its law enforcing institutions could not employ all means to realize its core functions of protection and provision of security. But if the state could not guarantee security, from a libertarian point of view the state simply would not be justified. Conclusively, we have to recognize that emphasizing individual rights in general and self-ownership in particular simultaneously may force us to accept the violation of these rights. Therefore, we face a kind of libertarian paradox or antinomy.

Final remarks

At first sight it seems unusual to discuss the issue of ICT implants, body modification, and surveillance referring to self-ownership. It is far more common to quote Foucault and other scholars following Foucault's line of argumentation. For there are a lot of really good reasons to do it that way since Foucault's main issue is power – power to control people. Surely, one can challenge whether Foucault is describing real or idealized kinds of societies – of course, “idealized” in a strange sense of the word – but one can also reasonably argue that Foucault actually is describing societies while here it was aimed towards basic moral concepts that might help to protect individuals against power and control. One of these means might be the concept of self-ownership.

To challenge the above mentioned ideas, one could argue that “self,” “body” or “person” are old-fashioned concepts which cannot describe social reality. One could stress that there is no such thing as a stable, given body, because nowadays it is possible to alter its shape and inner constitution; our bodies became fluid, alterable, an object of engineering (cf. Williams 1997) – as one can learn from the debates concerning cyborgs. Furthermore, one can argue that the body always was an entity which traits, characteristics, boundaries and the like has to be defined in a social discourse; that the body is not objectively given but socially constructed. But if even the body, which is matter, is nothing immutable, how could be the self or the person, which are not even matter, be conceptualized as something stable?

Indeed, we know that the notion of identity, self, or person appear to change caused by social change and technological progress. But, rather intuitively than knowingly, it seems that the core idea of self and person can be saved. At least, one could argue that the idea of a fluid, unstable, alterable body and person bears some obvious threat: If body and person are fluid then why not change them? Why to refuse its modification or the application of ICT implants? If we either consent or refuse to use ICT implants, eventually we have to refer to ideas implicitly included in sentences like "It's my body that only I have the right to modify". At least it is to say that we cannot talk about ourselves without reference to these old-fashioned concepts – and I would claim: even across cultural boundaries. But here it is not the time and space to discuss this further; this would be another issue for another paper.

REFERENCES

Christman, John (1991) Self-Ownership, Equality, and the Structure of Property Rights. In: *Political Theory*, 19 (1), pp. 28-46.

Coleman, Janet (2005) Pre-Modern Property and Self-Ownership Before and After Locke. Or, When did Common Decency Become a Private rather than a Public Virtue? In: *European Journal of Political Theory*, 4(2), pp. 125-145.

Fallon, Richard H. Jr. (1994) Two Senses of Autonomy. In: *Stanford Law Review*, 46 (4), pp. 875-905.

Glannon, Walter (2005) Neurobiology, Neuroimaging, and Free Will. In: *Midwest Studies in Philosophy*, 29, pp. 68-82.

Mack, Eric (2002) Self-ownership, Marxism, and Egalitarianism: Part I: Challenges to Historical Entitlement. In: *Politics, Philosophy & Economics*, 1 (1), pp. 75-108.

Moulin, Herve; Roemer, John (1989) Public Ownership of the External World and Private Ownership of Self. In: *The Journal of Political Economy*, 97 (2), pp. 347-367.

Nock, Christopher John (1992) Equal Freedom and Unequal Property: A Critique of Nozick's Libertarian Case. In: *Canadian Journal of Political Science*, 25 (4), pp. 677-695.

Nozick, Robert (1974) *Anarchy, State and Utopia*. New York: Basic Books.

Rawls, John (1999) *A Theory of Justice*. Cambridge/Massachusetts: Harvard University Press, revised edition (first published 1971).

Scoccia, Danny (1990) Paternalism and the Respect for Autonomy. In: *Ethics*, 100 (2), pp. 318-334.

Tay, Alice (1978) Law, the Citizen and the State. In: Kamenka, Eugene; Brown, Robert; Tay, Alice (eds.): Law and Society. The Crisis in Legal Ideas. London: Arnold, pp. 1-17.

Weber, Karsten (2006) The Next Step: Privacy Invasions by Biometrics and ICT Implants. In: Ubiquity – An ACM IT Magazine and Forum, 7 (45).

Williams, Simon J. (1997) Modern Medicine and the “Uncertain Body”: From Corporeality to Hyperreality? In: Social Science and Medicine, 45 (7), pp. 1041-1049.

What Should We Share?

Understanding the Aim of Intercultural Information Ethics

Pak Hang Wong*

Department of Philosophy, University of Twente

Abstract

The aim of Intercultural Information Ethics (IIE), as Ess aptly puts, is to “(a) address both local and global issues evoked by ICTs / CMC, e.tc., (b) in a ways that both sustain local traditions/values/preference, e.tc. and (c) provide shared, (quasi-) universal responses to central ethical problems” (Ess 2007a, 102). This formulation of the aim of IIE, however, is not unambiguous. In this paper, I will discuss two different understandings of the aim of IIE, one of which advocates “shared norms, different interpretations” and another proposes “shared norms, different justifications”. I shall argue that the first understanding is untenable, and the second understanding is acceptable only with qualification. Finally, I shall briefly suggest an alternative way to understand the aim of IIE.

Keywords: Intercultural Information Ethics, pragmatic arguments, pluralism, relativism, objectivism, value-based ethics

Introduction

In an increasingly globalizing world, Information and Communication Technologies (ICTs) are becoming an important part of the daily life for not only Western but also non-Western peoples. Ethical issues in relation to ICTs are therefore increasingly becoming global ethical issues. Ideally, one would like to have a set of (quasi-) universal responses to handle ethical problems related to ICTs in global and cross-cultural contexts. But is this possible? As some scholars have rightly pointed out (Brey 2007a; Capurro 2008; Ess 2005, 2006, 2007a, 2007b, 2008; Ess & Hongladarom 2007), current debates and discussions in information ethics are dominated by the ethical frameworks that are distinct from those in non-Western cultures, whose ethical frameworks may not be immediately compat-

* *Pak-Hang Wong* is a PhD Research Fellow at the Department of Philosophy, University of Twente. He is currently working on the project “Cultural, Political and Religious Ideologies and the Appraisal of New Media”, which is a sub-project of the VICI project on “Evaluation of the Cultural Quality of New Media” led by Prof. Philip Brey.

ible. For example, the arguments for the protection of privacy is often based on the individual's autonomy; these arguments may sound peculiar for Confucian cultures, which generally weigh the collective, common good over and above the benefit of individuals.

Being aware of the importance of cultural diversity and the vices of ethical imperialism, these scholars have called for careful investigations of moral systems in different cultures, when dealing with intercultural or cross-cultural ethical issues related to ICTs. As such, the aim of Intercultural Information Ethics (IIE), as Ess aptly puts, is to “(a) address both local and global issues evoked by ICTs / CMC, e.tc., (b) in a ways that both sustain local traditions / values / preference, e.tc. and (c) provide shared, (quasi-) universal responses to central ethical problems” (Ess 2007a, 102). This formulation of the aim of IIE, however, is not unambiguous. Particularly, it is unclear as to what exactly does “sustain local traditions / values / preference, e.tc.” refers to; as well as, what “shared, (quasi-) universal responses” means. In this paper, I will offer two possible understandings of the aim of IIE as characterized by Ess; namely, one that advocates (i) ‘shared norms, different interpretations’ and one that proposes (ii) ‘shared norms, different justifications’. I shall argue that (i) is untenable, and (ii) is acceptable only with qualifications. In doing so, I hope to illustrate the inadequacy of the prevailing understandings of the aim of IIE. To overcome the inadequacy, I shall briefly suggest an alternative way to understand the aim of IIE in the final section.

Intercultural Information Ethics as a Normative Project

Before explaining the two understandings of the aim of IIE, it is worth identifying the basic requirements for an adequate framework for IIE. Already pointed out by Brey (2007a) and Himma (2008), two distinct stages of IIE can, and should, be distinguished. These two stages are respectively, the descriptive analysis of different moral systems in various cultures and the normative analysis of these moral systems and the related task of formulating (quasi-) universal moral principles in response to ICTs-related ethical issues. As it is clear in Ess's formulation of the aim of IIE, both descriptive analysis and normative analysis are essential components of an adequate framework for IIE.

As IIE aims to provide responses to ICTs-related ethical issues which are (quasi-) acceptable from various cultural perspectives, what is considered to be acceptable in the moral systems of these cultures must first be identified; thus, the first stage of IIE will involve tasks such as explicating the actual moral norms and/or moral values embedded in these cultures, studying the impacts of ICTs to these cultures and their reactions towards ICTs. The empirical findings, then will provide the basis for formulating the (quasi-) universal moral principles; thus, de-

scriptive analysis of moral systems is indispensable in any adequate framework of IIE.

However, equally important for an adequate framework for IIE is the possibility to derive normative and evaluative judgments from the framework; as IIE does not only aim to describe the actual responses to the ICT-related ethical issues from a specific cultural perspective, it also attempts to provide “shared, (quasi-) universal responses to central ethical problems”. In other words, IIE is prescriptive insofar as it tells us how one ought to respond to the intercultural or cross-cultural ICTs-related ethical problems; and, for such responses to be meaningful, they have to be normative, minimally, in the sense that if party A fails to respond as specified (or, if A fails to comply to the normative, moral standard underlying the responses), the other parties can legitimately condemn A’s failure to do so.

Since IIE aims to investigate the ICTs-related ethical problems from various cultural perspectives and attempts to arrive at some agreements on how these ethical problems are to be settled interculturally or cross-culturally; therefore, it must employ both empirical findings of different cultural perspectives as well as normative analysis to determine what can, and should, be agreed upon. Once the agreements are reached, an adequate framework should also enable us to criticize and condemn those who fail to comply with the standard as specified. For any proper framework of IIE, therefore, it must have rooms for both descriptive analysis and normative analysis.

I have explained that any adequate framework for IIE must be normative, in the sense that it should allow us to criticize or condemn others morally when a party fails to follow the shared, (quasi-) universal responses; however, criticisms or condemnations are only possible, if we can reject (or, at least, restrict) metaethical moral relativism – the view that the truth or falsity of moral judgments, or their justification, is not absolute or universal, but is relative to the traditions, convictions, or practices of a group of persons. For, if metaethical moral relativism is true, then it does not make any sense to criticize or condemn people from different cultures because their judgements, actions, etc. are based on different moral systems which is equally legitimate; indeed, if metaethical moral relativism is true, there seems to be no point of making any agreements at all, they are just unnecessary. While it is clear that any adequate framework in IIE has to resist metaethical moral relativism, it is also important to remind ourselves of the importance to avoid ethical imperialism that is – to impose one’s moral system onto another culture, or simply puts, to require peoples from other cultures to judge or act according to their own norms and values.

“Shared norms, different interpretations” as the aim of IIE

It should be clear by now that Ess's formulation of the aim of IIE can be understood as both an attempt to avoid cultural-ethical imperialism (e.g. see (b)), and metaethical moral relativism, (e.g. see (c)). On a more concrete level, then, the aim of IIE as formulated is to provide shared norms for different societies with different cultures and distinct moral systems; and, at the same time, maintain the cultural diversity and respect the distinctiveness of various moral systems. The formulation by Ess, however, is not unambiguous. Particularly, it is unclear exactly what counts as maintaining cultural diversity and respecting different moral systems, e.g. (b), and when the norms are considered to be shared. If one looks at the theoretical foundations of recent debates and discussions of IIE (for summary, see Capurro 2008; Ess 2006, 2007a, 2007b, 2008; Ess & Hongladarom 2007), one may discern two prevailing understandings of the aim of IIE, namely “shared norms, different interpretations” and “shared norms, different justifications”. As I will elaborate, the first understanding holds that there can be shared norms between different cultures, but different cultures can interpret the meaning of these norms differently with respect to their own moral systems. Here, maintaining cultural diversity and respecting other moral systems amount to pluralism in the interpretation of norms; on the second understanding, cultural diversity is maintained and other moral systems are respected through a plurality of justifications of norms, which means that different cultures should share a set of norms, but justifications of these norms may be different.

In a series of paper, Ess (Ess 2006, 2007a, 2007b, 2008) has developed and defended a theoretical foundation for IIE which he calls *pros hen* pluralism; while different interpretations of Ess's theory can place it either as the “shared norms, different interpretations” approach or the “shared norms, different justifications” approach; nevertheless, the first interpretation of Ess's theory provides the best example of how a “shared norms, different interpretations” approach would look like. In remaining of this section, I shall elaborate such interpretation of Ess's theory and to show that it does not satisfy the agendas of IIE, and hence is untenable as the aim of IIE.

Ess's *pros hen* pluralism is based on Aristotle's account of *pros hen* equivocals as well as the idea of *phronesis*, i.e. Aristotle's notion of practical judgment. According to the *pros hen* pluralism, it is possible for the shared norms to take different but related meanings, as he puts (Ess 2007a, 13)

[Pros hen] equivocals stand as linguistic middle grounds between a homogenous univocation (which requires that a term have one and only one meaning) and a pure equivocation (as a single term may have multiple but entirely unrelated meanings...). Pros hen or focal equivocals, by contrast, are terms

with clearly different meanings that simultaneously relate or cohere with one another as both point towards a shared or focal notion that anchors the meaning of each.

And, it is up to different cultures to exercise their pragmatic judgment to determine their responses to the ICTs-related ethical problems; as Ess emphasizes the role of *phronesis* in his theory (Ess 2007a, 14)

[P]hronesis allows us to take a general principle (as the ethical analogue to the focal term ground two *pros hen* equivocals) and discern how it may be interpreted or applied in different ways in different contexts (as the ethical analogues to the two *pros hen* equivocals – i.e., that are irreducibly different and yet inextricably connected).

Taking the *pros hen* pluralism at its face value, what it appears to suggest is that the shared norms should be construed in the *pros hen* manner, i.e. the meaning of the shared norms is flexible and it remains so until it is being used to respond to ICTs-related ethical problems; then, the norms takes on a more concrete meaning supplied by the culture(s) involved. In fact, Ess thinks that different cultures are crucial in determining the meaning of the norms, as different culture can be complementary to the others. In this way, various cultures are irreducibly different, but they can complement the other in determining the meaning of the norms.

In defending the *pros hen* pluralism, Ess draws support from both researches in information ethics and comparative ethics; and his examples range from Bernd Carsten Stahl's notion of responsibility (and its applications in management of information systems), Association of Internet Researchers (AoIR)'s ethical guidelines for online research, comparative studies of Virtue Ethics/Ethics of Care and Confucian ethics to the notions of privacy in China and Hong Kong and many more. In these example, Ess tries to demonstrate how different parties, while sharing the same set of norms, they can nevertheless understand the norms differently.

Ess's defence is lucid and admirable, but as Capurro (Capurro 2008) quickly points out, it is unclear how *pros hen* pluralism may resolve the tension created by the irreducibility of various cultural perspectives and their complementarity; indeed, it is unclear the different interpretations itself can help strengthen the shared norms; perhaps, more importantly, allowing different interpretations of norms is too unconstrained to avoid metaethical moral relativism. While Ess's own *pros hen* pluralism allows "the interpretation of a single idea themselves remain irreducibly different from another, but connected and coherent with one another by way of their shared point of origin and reference" (Ess 2006, 218), it seems to follow that a shared norm can be interpreted, and thus, implemented

differently, as long as there is a “shared point of origin and reference”. Yet, without clearly identifying what is a legitimate “shared point of origin of reference”; any contingent facts may be employed to justify vastly contradictory interpretations, as well as conflicting implementations of a norm; and hence, pushing IIE towards metaethical moral relativism.

“Shared norms, different justifications” as the aim of IIE

While being too unconstrained, the “shared norms, different interpretations” approach cannot avoid metaethical moral relativism; one may still insist on the necessity of shared norms, and turn to focus on different justifications of the same set of norms. In other words, one may stress on a more definite meaning of a norms, e.g. protection of privacy has a more or less fixed meaning; but, at the same time, distinct justifications for the norms can be derived from different cultural perspectives. For example, Hongladarom has argued that Buddhism also agrees with the protection of an individual’s privacy; but the justification of it differs from the Western autonomy-based reasoning; in Buddhism, protection of privacy is not considered to be a protection of the individual’s rights per se, rather it is considered to be a measure against human evil, such as greed for power, material gains, etc. (Hongladarom 2007). Such approach, therefore, aims to arrive at shared norms with different justifications (from various cultural perspective). I shall call this understanding of the aim of IIE as “shared norms, different justifications”.

Here, given a standard of reasonableness, understanding plurality in terms of diverse justifications from different cultural perspectives seems to avoid relativism. However, two forms of reasonable justification, that is – ethical justification and pragmatic justification – have to be distinguished. Ethical justifications are based on the moral values within a particular moral system; as such, a moral system can ethically justify a norm if it has the relevant (set of) moral values; on the other hand, there is pragmatic justification, which “are (as far as possible) detached from any socio-political or philosophical presuppositions” (Soraker 2006, 123). Yet, neither forms of justification, as I will show, help the “shared norms, different justifications” approach to satisfy the basic requirements of IIE.

I believe a commonsensical concern for the using ethical justifications by the “shared norms, different justifications” approach is that: a particular culture simply does not have the resources to justify the norms, as it lacks the concepts at stake. For example, Brey has demonstrated that concepts such as privacy, intellectual property rights and freedom of information are missing from the Chinese, Japanese and Thai culture (Brey 2007a); and, Burk goes beyond by arguing that ethical justifications for intellectual property rights and individual privacy rights

appear to be missing in Confucian heritage (Burk 2007). The commonsensical concern, however, conflates the moral concepts related to ICTs with the more fundamental values in which these concepts are based upon. In other words, the fact that a particular culture does not, currently, have concepts like privacy, intellectual property rights, and freedom to information does not by itself entail the culture does not have the resources to justify them; moreover, to claim that a particular culture does not have any resource to justify these concepts seems to appeal to a very simplistic picture of different cultural perspectives, and thus misses their complexity.

The true worry for ethical justification, I believe, precisely stemmed from the complexity of various cultural perspectives. Indeed, if we consider the ethical debates in Western tradition, it is rather typical that a norm can be justified by an utilitarian-based ethics, while the negation of the very same norm can be justified by a deontological ethics. The very same thing can also be said of Confucian ethics, as it is often forgotten that Confucianism is not simply fixed rules derived from the canons; but, it is itself a school of thought that contains various sub-traditions, e.g. Neo-Confucianism, New Confucianism, e.tc.; and, the problem of complexity multiplies once we consider Chinese culture as a whole, which is constituted by Confucian, Daoist and Zen, and each has their own moral systems. Given the complexity of any cultural perspective, we can expect that a norm can be justified by the moral values embedded in the culture, while, at the same time, it can be rejected by the very same culture with different moral values (or, same values, but different interpretations of those moral values). The problem for “shared norms, different (ethical) justifications” therefore is one that: when different ethical justifications are equally legitimate, and these ethical justifications can justify the norms and their falsity; it is logically possible that no norm can ever be shared. In other words, it renders this approach logically inconsistent at its worse.

Hence, without an overarching position, the “shared norms, different (ethical) justifications” is not going to work. Perhaps, then we should look elsewhere to justify the shared norms; as I have pointed out, other than ethical justifications, there also pragmatic justifications (e.g. Soraker 2006), which may help to bring ‘overlapping consensus’ by highlighting the pragmatical benefits of maintaining the shared norms; indeed, as Ess’s and Soraker’s examples, e.g. emerging notions of privacy in China and Hong Kong (Ess 2008) and China’s regulations and surveillance of the Internet (Soraker 2006), forcefully demonstrate pragmatic arguments (and thus, pragmatic justifications) appear to be most effective way to justify the norms. However, pragmatic justifications is ill-suited for IIE in two senses: firstly, it risks transforming ethical problems into non-ethical problems by substituting ethical justifications with pragmatic justifications; secondly, it

seems to downplay the role of the moral systems in different cultural perspectives in IIE, when these moral systems should be at the core of IIE.

First, consider the force of pragmatic justifications in promoting or defending the shared norms; it is their being abstracted away from a particular cultural perspective, by turning the emphasis on favourable and unfavourable, that makes agreements by different cultures more easily. Hence, arguing along Soraker's line, China's regulations and surveillance of the Internet is problematic insofar as the consequence of it is unfavourable to the Chinese government; but, if China fail to response to the pragmatic justifications, it is only possible to accuse of her being pragmatically incompetent but not ethically wrong. In other words, the shift towards pragmatic justifications also signifies the shift away from making moral evaluations. An approach like this will thereby fail to satisfy the basic requirement for an adequate framework of IIE, i.e. that the framework should be normative.

Also, the shift towards pragmatic justifications also signifies a move away from any cultural perspectives; thus, rather than promoting cultural diversity and respecting different moral systems, pragmatic justifications throw away the "inter-cultural" project of IIE, and therefore, it is ill-suited for IIE.

Perhaps, the worse problem for grounding pragmatic justifications a significant role in IIE is its essential link to the economic considerations; there are, in fact, two arguments for rejecting pragmatic justifications in IIE: theoretically, it is, as I have tried to show, contradictory to the basic tenet of IIE, that is – to maintain cultural diversity and to respect different moral systems; one may even go further to argue that the use of pragmatic justifications itself is one form of imperialism; and, practically, it is likely to be an unequal tools of negotiations.

The use of pragmatic justifications can be construed as a form of imperialism, because it presupposes economic progress (or, at least some form of progress) to be the most basic value; however, the concept of economic progress is not neutral for different cultural perspective; by putting it at the core of IIE, then, run afoul of the original intention of IIE. Secondly, because of the pretended neutrality of economic progress for diverge cultures, they may be used to mask any attempt to dominate other cultures in the name of supposed economic progress. By pushing pragmatic justifications in such terms, and by allowing our sources for evaluations in economic terms, it essentially transformed the ethical issues into political struggles, in which various agendas can be coined in the spirit of promoting progress.

In sum, the justifications in "the shared norms, different justifications" ought not to be pragmatic, as the use of pragmatic justifications runs against the nature of

IIE. Yet, I have also argued that ethical justifications are not a suitable candidate for such approach once we realize the complexity of different cultures; and, with no overarching position to judge which justifications are more legitimate, the prospect of shared norms is dim.

What Should We Share?

In the previous sections, I have tried to show both understandings of the aim of IIE are problematic. I believe the problem arise from an overemphasis on producing shared norms. While I agree that having a shared (set of) norms is of practical importance, (a set of) well defined rules can help to resolve disputed ICTs-related ethical issues, particularly, in the (quasi-) legal sense, e.g. international laws, professional code of ethics, e.tc., but the emphasis on what rules can and should be shared and how to implement these rules lead to an oversight of the philosophical-theoretical foundation of IIE.

Already demonstrated in my discussion of the two understandings of the aim of IIE as producing shared norms, the resulting norms have to be either 'open' or 'thin', i.e. either it is open to different meanings or it has no substantial normative content at all; in this sense, it is possible to see the aim of IIE, for the shared norms approaches, is to produce a minimal moral denominator, that is – to specify the basic norms in ICTs-related ethical issues that can be accepted by all cultural perspectives. As I have shown, without an overarching position to arbitrate between different interpretations and/or various justifications, such minimal moral denominator is too weak to issue normative or evaluative judgments. As a normative project, the shared norms approaches falsely based its normative foundation on minimal moral denominator. While I agree with Himma that an objective moral foundation is necessary for the normative project of IIE (Himma 2007). Yet, the question remains: what, if not shared norms, can provide the objective moral foundation for IIE. The answer, I believe, is (a set of) shared values.

My call for a shift towards values is not entirely new (for examples, Brey 2007b, Bynum 2006; Floridi 2007; Johnstone 2007), and what distinguishes the values-based approaches from the norms-based approaches is their attempts to identify (a set of) basic, common values which is valid across various cultures; it remains an open question as to what the (set of) basic, common values may look like, it may take the form of human (and non-human) flourishing (Bynum 2006), being (Floridi 2007) or capabilities (Johnstone 2007); but, what is important for the values-based approaches is that: such (set of) basic, common values are defined normatively, and that we have the moral responsibility to maintain and promote these values. As such, it provides a normative foundation for arbitrating ethical issues, based on the shared values. But, how will the values-talk translate to ICTs-related ethical issues in IIE? For one thing, ICTs-related ethical issues often claim

to arise from cultures possessing different values, e.g. privacy issues in East-West context are often construed as a stand-off between community-based values and individualistic values; the shared values approaches will urge for a close investigation of the scenario and the values involved, and to determine if it is true that no shared value is available; and, it is particularly important for the values-based approaches not to overlook complexity of different cultural perspectives, in its retrieval of moral resource.

More time and efforts have to be given to fully develop a values-based IIE; here, it is important to note that the aim of IIE is not to eliminate every moral disagreements in ICTs-related ethical issues; as moral disagreement appears to be an ineliminable feature of our practical reality; but, at least, IIE should allow us to legitimately formulate normative and evaluate judgements for these issues. To do so, IIE must have a proper moral foundation; I have tried to show that the use of “shared norms” fails to provide such foundation; and hence, “shared values” appears to be a more promising candidate for the task. Moreover, as Johnstone (Johnstone 2007) has pointed out, attending to values opens up the space for issues which are marginal in the norms-based debates, e.g. well-being, digital divide, and gender issues. In other words, an added advantage of values-based IIE would be a more encompassing project than the one in norms-based IIE.

Conclusions

In this paper, I have examined two understandings of the aim of IIE, they are respectively, “shared norms, different interpretations” and “shared norms, different justifications”; both approaches aim at establishing shared norms that would be accepted by different cultures. I have tried to show that neither of them provide proper basis for IIE as a normative project. While I have not prove, in this paper, “shared values” is the ultimate aim for IIE; my modest objective is to demonstrate the problems and weakness of the approaches which focus primarily on establishing shared norms; and hence, to open up the space of discussion for values-based approaches for IIE.

REFERENCES

- Brey, P.,** (2007a). Is Information Ethics Culturally Relative? *International Journal of Technology and Human Interaction*, 3 (3), 12-24.
- Brey, P.,** (2007b). Theorizing the Cultural Quality of New Media. *Technè: Research in Philosophy and Technology*, 11(1), 2-18.
- Burk, D.,** (2007). Privacy and Property in the Global Datasphere, in S. Hongladarom and C. Ess (eds.), *Information Technology Ethics: Cultural Perspectives*, 94-105. Hershey, PA: Idea Group.
- Bynum, T.W.,** (2006). Flourishing Ethics. *Ethics and Information Technology*, 8(4), 157-173.
- Capurro, R.,** (2008). Intercultural Information Ethics, in Kenneth E. Himma and Herman T. Tavani (eds.), *Handbook of Information and Computer Ethics*, 639-665. John Wiley & Sons, Inc., Hoboken, NJ.
- Ess, C.,** (2005). "Lost in Translation"? Intercultural Dialogues on Privacy and Information Ethics. *Ethics and Information Technology*, 7(1), 1-6
- Ess, C.,** (2006). Ethical Pluralism and Global Information Ethics. *Ethics and Information Technology*, 8(4), 215-226.
- Ess, C.** (2007a). Cybernetic Pluralism in an Emerging Global Information and Computer Ethics. *International Review of Information Ethics*, 7, 94-123.
- Ess, C.** (2007b). Universal Information Ethics? Ethical Pluralism and Social Justice, in Emma Rooksby and John Weckert (eds.), *Information Technology and Social Justice*, 69-92. Hershey, PA: Idea Group Reference.
- Ess, C.,** (2008). Culture and Global Networks: Hope for a Global Ethics?, in J. van den Hoven and J. Weckert (eds.), *Information Technology and Moral Philosophy*, 195-225. Cambridge; NY: Cambridge University Press.
- Ess, C. and Hongladarom, S.,** (2007). *Information Technology Ethics: Cultural Perspectives*. Hershey, PA: Idea Group Reference.
- Floridi, L** (2007). Global Information Ethics: The Importance of Being Environmentally Earnest. *International Journal of Technology and Human Interaction*, 3(3), 1-11.
- Himma, K.E.,** (2008). The intercultural ethics agenda from the point of view of a moral objectivist. *Journal of Information, Communication & Ethics in Society*, Vol. 6 No. 2, 101-115
- Hongladarom, S.** (2007). Analysis and justification of Privacy from a Buddhist Perspective, in S. Hongladarom and C. Ess (eds.), *Information Technology Ethics: Cultural Perspectives*, 108-122. Hershey, PA: Idea Group.

Johnstone, J., (2007). Technology as empowerment: a capability approach to computer ethics. *Ethics and Information Technology*, 9, 73-87

Rooksby, E. and Weckert, J., (2006). *Information Technology and Social Justice*, Hershey, PA: Idea Group Reference.

Soraker, J. (2006). The Role of Pragmatic Arguments in Computer Ethics. *Ethics and Information Technology*, 8(3), 121-130.

“But the Data is Already Public”: On the Ethics of Research in Facebook

Michael Zimmer*

School of Information Studies, University
of Wisconsin-Milwaukee

Abstract

Using the public release of personal data collected from the Facebook accounts of an entire cohort of college students as a case study, this paper articulates a set of ethical concerns that must be addressed before embarking on future research in social networking sites, including the nature of consent, properly identifying and respecting expectations of privacy on social network sites, strategies for data anonymization prior to public release, and the relative expertise of institutional review boards when confronted with research projects based on data gleaned from social media.

keywords: Research ethics, social networks, social media, privacy, anonymity

Introduction

In September 2008, a group of researchers associated with the Berkman Center for Internet & Society at Harvard University publicly released data collected from the Facebook accounts of an entire cohort of college students. Titled “Tastes, Ties, and Time” (T3), the announcement accompanying the release noted the uniqueness of the data:

The dataset comprises machine-readable files of virtually all the information posted on approximately 1,700 [Facebook] profiles by an entire cohort of students at an anonymous, northeastern American university. Profiles were sampled at one-year intervals, beginning in 2006. This first wave covers first-year profiles, and three additional waves of data will be added over time, one for each year of the cohort’s college career.

* *Michael Zimmer*, PhD, is an assistant professor in the School of Information Studies at the University of Wisconsin-Milwaukee, and an associate at the Center for Information Policy Research. With a background in new media and Internet studies, the philosophy of technology, and information policy, Zimmer studies how new media and Internet technologies impact information flows, access to knowledge, and informational privacy.

Though friendships outside the cohort are not part of the data, this snapshot of an entire class over its four years in college, including supplementary information about where students lived on campus, makes it possible to pose diverse questions about the relationships between social networks, online and offline. (Berkman Center for Internet & Society, 2008)

Recognizing the privacy concerns inherent with the collection and release of social networking data, the T3 research team took various steps in an attempt to protect the identify of the subjects, including the removal of student names and identification numbers from the dataset, a delay in the release of the cultural interests of the subjects, and requiring other researchers to agree to a “terms and conditions for use,” prohibiting various uses of the data that might compromise student privacy, and undergoing review by their institutional review board (Lewis, 2008, pp. 28-29).

Despite these steps, and claims by the T3 researchers that “all identifying information was deleted or encoded” (Lewis, 2008, p. 30), the identity of the source of the dataset was quickly discovered. Using only the publicly available codebook for the dataset and other public comments made about the research project, the identity of the “anonymous, northeastern American university” from which the data was drawn was quickly narrowed down to 13 possible universities (Zimmer, 2008b), and then surmised to be Harvard College (Zimmer, 2008a). Reminiscent of the ease at which AOL searchers were re-identified when the search engine thought the release of search history data was sufficiently anonymized (see Barbaro and Zeller Jr, 2006), this re-identification of the source institution of the T3 data reveals the fragility of the presumed privacy of the subjects under study.

Using the T3 data release and its aftermath as a case study, this paper will articulate a set of ethical concerns that must be addressed before embarking on future research in social networking sites. These include challenges to the traditional nature of consent, properly identifying and respecting expectations of privacy on social network sites, developing sufficient strategies for data anonymization prior to the public release of personal data, and the relative expertise of institutional review boards when confronted with research projects based on data gleaned from social media.

The “tastes, ties, and time” project

Research in social networks has spanned decades, from Georg Simmel’s foundational work in sociology (Simmel and Wolff, 1964), to Barry Wellman’s analyses of social networks in the emerging networked society of the late twentieth century (Wellman and Berkowitz, 1988), to the deep ethnographies of contemporary

online social networks by danah boyd (boyd, 2008b). Indeed, the explosive popularity of online social networking sites such as MySpace, Twitter, and Facebook has attracted attention from a variety of researchers and disciplines (see boyd and Ellison, 2008). A primary challenge to fully understanding the nature and dynamic of social networks is obtaining sufficient data. Most existing studies rely on external surveys of social networking participants, ethnographies of smaller subsets of subjects, or the analysis of limited profile information extracted from what subjects chose to make visible. As a result, the available data can often be tainted due to self-reporting biases and errors, have minimal representativeness of the entire population, or fail to reflect the true depth and complexity of the information users submit (and create) on social networking sites.

Recognizing the data limitations faced by typical sociological studies of online social network dynamics, a group of researchers from Harvard University and the University of California – Los Angeles set out to construct a more robust dataset that would fully leverage the rich data available on social networking websites. Given its popularity, the researchers chose the social network site Facebook as their data source, and located a university that allowed them to download the Facebook profiles of every member of the freshman class:

With permission from Facebook and the university in question, we first accessed Facebook on March 10 and 11, 2006 and downloaded the profile and network data provided by one cohort of college students. This population, the freshman class of 2009 at a diverse private college in the Northeast U.S., has an exceptionally high participation rate on Facebook: of the 1640 freshmen students enrolled at the college, 97.4% maintained Facebook profiles at the time of download and 59.2% of these students had last updated their profile within 5 days. (Lewis et al., 2008, p. 331)

This first wave of data collection took place in 2006, during the spring of the cohort's freshman year, and data collection is scheduled to be repeated annually until 2009, when the vast majority of the study population will graduate, providing four years of data about this collegiate social network. Each student's official housing records were also obtained from the university, allowing the researchers to "connect Internet space to real space" (Kaufman, 2008a).

The uniqueness of this dataset is of obvious value for sociologists and Internet researchers. The data was extracted directly from Facebook without direct interaction with the subjects or reliance on self-reporting instruments, either of which could taint the data collected. The dataset includes demographic, relational, and cultural information on each subject, allowing broad analyses beyond more simple profile scraping methods. The inclusion of housing data for each of the four years of the study for analysis of any connection between "physical proximity,

emerging roommate and friendship groups in the real world and the presence of these two types of relationships in their Facebook space” (Kaufman, 2008a). Most importantly, the dataset represents nearly a complete cohort of college students, allowing the unique analysis of “complete social universe” (Kaufman, 2008a), and it is longitudinal, providing the ability to study how the social network changes over time.

As a result of its uniqueness, the dataset can be employed for a number of research projects that have heretofore been difficult or impossible to pursue. As one of the “Tastes, Ties, and Time” researchers noted, “We’re on the cusp of a new way of doing social science... Our predecessors could only dream of the kind of data we now have” (Nicholas Christakis, quoted in Rosenbloom, 2007).

The dataset release

The “Tastes, Ties, and Time” project has been funded, in part, by a grant from the National Science Foundation, who mandates certain levels of data sharing as a condition of its grants. As a result, the Facebook dataset is being made available for public use in phases, roughly matching the annual frequency of data collection: wave 1 in September 2008, wave 2 in the fall of 2009, wave 3 in the fall of 2010, and wave 4 in the fall of 2011 (Lewis, 2008, p. 3).

The first wave of data, comprising of “machine-readable files of virtually all the information posted on approximately 1700 FB profiles by an entire cohort of students at an anonymous, northeastern American university,” was publicly released on September 25, 2008 (Berkman Center for Internet & Society, 2008). Prospective users of the dataset are required to submit a brief statement detailing how the data will be used, and access is granted at the discretion of the T3 research team. Researchers are also required to agree to a “Terms and Conditions of Use” statement in order to gain access to the dataset, consenting to various licensing, use, and attribution provisions.

A comprehensive codebook was downloadable without the need to submit an application, which included detailed descriptions and frequencies of the various data elements (see Lewis, 2008), including gender, race, ethnicity, home state, political views, and college major. For example, the codebook revealed that the dataset included 819 male and 821 female subjects, and that there were 1 self-identified Albanian, 2 Armenians, 3 Bulgarians, 9 Canadians, and so on.

The codebook also included an account of the steps taken by the T3 researchers in an attempt to protect subject privacy:

All data were collected with the permission of the college being studied, the college’s Committee on the Use of Human Subjects, as well as Facebook.com.

Pursuant to the authors' agreement with the Committee on the Use of Human Subjects, a number of precautionary steps were taken to ensure that the identity and privacy of students in this study remain protected. Only those data that were accessible by default by each RA were collected, and no students were contacted for additional information. All identifying information was deleted or encoded immediately after the data were downloaded. The roster of student names and identification numbers is maintained on a secure local server accessible only by the authors of this study. This roster will be destroyed immediately after the last wave of data is processed. The complete set of cultural taste labels provides a kind of "cultural fingerprint" for many students, and so these labels will be released only after a substantial delay in order to ensure that students' identities remain anonymous. Finally, in order to access any part of the dataset, prospective users must read and electronically sign [a] user agreement.... (Lewis, 2008, p. 29)

These steps taken by the T3 researchers to remove identifying information reveal an acknowledgment of – and sensitivity to – the privacy concerns that will necessarily arise given the public release of such a rich and complete set of Facebook data. Their intent, as expressed by the project's principle investigator, Jason Kaufman, was to ensure that "all the data is cleaned so you can't connect anyone to an identity" (Kaufman, 2008a). Unfortunately, Dr. Kaufman was overly optimistic.

Re-identification and withdrawal of dataset

Cognizant of the privacy concerns related to collecting and releasing detailed Facebook profile data from a cohort of college students, the T3 research team – in good faith – took a number of steps in an attempt to protect subject privacy, including review by their institutional review board, the removal of student names and identification numbers from the dataset, a delay in the release of the cultural interests of the subjects, and requiring other researchers to agree to a "terms and conditions for use," prohibiting various uses of the data that might compromise student privacy.

However, despite these efforts, the team's desire to ensure "all the data is cleaned so you can't connect anyone to an identity" fell short. On September 29, 2008, only four days after the initial data release, Fred Stutzman, a Ph.D. student at the University of North Carolina at Chapel Hill's School of Information and Library Science, questioned the T3 researchers faith in the non-identifiability of the dataset:

The "non-identifiability" of such a dataset is up for debate. A friend network can be thought of as a fingerprint; it is likely that no two networks will be exactly similar, meaning individuals may be able to be identified in the dataset post-hoc.... Further, the authors of the dataset plan to release student "Fa-

vorite” data in 2011, which will provide further information that may lead to identification. (Stutzman, 2008)

Commenting on Stutzman’s blog post on the subject, Eszter Hargittai, an Associate Professor of Communication Studies at Northwestern University, sounded similar concerns:

I think it’s hard to imagine that some of this anonymity wouldn’t be breached with some of the participants in the sample. For one thing, some nationalities are only represented by one person. Another issue is that the particular list of majors makes it quite easy to guess which specific school was used to draw the sample. Put those two pieces of information together and I can imagine all sorts of identities becoming rather obvious to at least some people. (Hargittai, 2008)

Stutzman and Hargittai share a fear of the possible re-identification of the presumed anonymous Facebook dataset that has been made available to the public. Stutzman’s concern over the ability to exploit the uniqueness of one’s social graph to identify an individual within a large dataset has proven true in numerous cases (see, for example, Narayanan and Shmatikov, 2008; Narayanan and Shmatikov, 2009). Hargittai suggests that the uniqueness of some of the data elements makes identifying the source of the data – and therefore some of the individual subjects – quite trivial. Hargittai’s fears were correct.

Re-Identification

Within days of its public release, the T3 dataset was identified to have come from Harvard College (see Zimmer, 2008b; Zimmer, 2008a). Most striking about this revelation was that the re-identification of the source of the Facebook data did not require access to the full dataset itself.

Using only the freely available codebook and referencing various public comments about the research, the source of the data was quickly narrowed down from over 2000 possible colleges and universities to a list of only seven (Zimmer, 2008b). An examination of the codebook revealed the source was a private, co-educational institution, whose class of 2009 initially had 1640 students in it. Elsewhere, the source was identified as a “New England” school. A search through an online college database revealed only seven private, co-ed colleges in New England states (CT, ME, MA, NH, RI, VT) with total undergraduate populations between 5000 and 7500 students (a likely range if there were 1640 in the 2006 freshman class): Tufts University, Suffolk University, Yale University, University of Hartford, Quinnipiac University, Brown University, and Harvard College.

Upon the public announcement of this initial discovery, and general criticism of the research team’s attempts to protect the privacy of the subjects, Jason

Kaufman, the principle investigator of the T3 research project, was quick to react, noting that, perhaps in justification for the amount of details released in the dataset, “We’re sociologists, not technologists, so a lot of this is new to us” and “Sociologists generally want to know as much as possible about research subjects” (Kaufman, 2008b). He then attempts to diffuse some of the implicit privacy concerns with the following comment:

What might hackers want to do with this information, assuming they could crack the data and ‘see’ these people’s Facebook info? Couldn’t they do this just as easily via Facebook itself?

Our dataset contains almost no information that isn’t on Facebook. (Privacy filters obviously aren’t much of an obstacle to those who want to get around them.) (Kaufman, 2008b)

And then:

We have not accessed any information not otherwise available on Facebook. We have not interviewed anyone, nor asked them for any information, nor made information about them public (unless, as you all point out, someone goes to the extreme effort of cracking our dataset, which we hope it will be hard to do). (Kaufman, 2008c)

However, little “extreme effort” was needed to further “crack” the dataset; it was easily accomplished a day later, again without ever looking at the data itself (Zimmer, 2008a). As Hargittai recognized, the unique majors listed in the codebook allowed for the ultimate identification of the source university. Only Harvard College offers the specific variety of the subjects’ majors that are listed in the codebook, such as Near Eastern Languages and Civilizations, Studies of Women, Gender and Sexuality, and Organismic and Evolutionary Biology. The identification of Harvard College was further confirmed after analysis of a June 2008 video presentation by Kaufman, where he noted that “midway through the freshman year, students have to pick between 1 and 7 best friends” that they will essentially live with for the rest of their undergraduate career (Kaufman, 2008a). This describes the unique method for determining undergraduate housing at Harvard: all freshman who complete the fall term enter in to a lottery, where they can designate a “blocking group” of between 2 and 8 students with whom they would like be housed in close proximity.

Withdrawal

The announcement of this likely identification of the source of the Facebook dataset did not prompt a public reply by Jason Kaufman or others from the T3 research team, but within one week of the discovery, the access page for the

“Tastes, Ties, and Time” dataset displayed the following message, indicating that the dataset was, at least for the moment, no longer publicly available:

Note: As of 10/8/08, prospective users may still submit requests and research statements, but the approval process will be delayed until further notice. We apologize for the inconvenience, and thank you for your patience.

Then, in March 2009, the page was updated with a new message:

UPDATE (3/19/09): Internal revisions are almost complete, and we expect to begin distributing again in the next 2-3 weeks. In the meantime, please DO NOT submit new dataset requests; but please check back frequently at this website for a final release notice. We again apologize for any inconvenience, and thank you for your patience and understanding as we work to ensure that our dataset maintains the highest standards for protecting student privacy.

These messages noting the restricted access to the Facebook dataset to “ensure that our dataset maintains the highest standards for protecting student privacy” suggest that the re-identification of the source as Harvard College was correct. As of May 1, 2009, access to the dataset has yet to be restored.

Privacy and the “tastes, ties, and time” project methodology

The changing nature – and expectations – of privacy in online social networks are being increasingly debated and explored (see, for example, Albrechtslund, 2008; Barnes, 2006; Grimmelmann, 2009; Gross and Acquisti, 2005; Lenhart and Madden, 2007; Nussbaum, 2007; Solove, 2007). The events surrounding the release of the Facebook data in the “Tastes, Ties, and Time” reveals many of the fault lines within these debates. Critically examining the methods of the T3 research project, and the public release of the dataset, reveals gaps in the understanding the nature of privacy and anonymity in the context of social networking sites.

The primary steps taken by the T3 research team to protect subject privacy (quoted above), can be summarized as follows:

1. Only those data that were accessible by default by each RA were collected, and no students were contacted for additional information.
2. All identifying information was deleted or encoded immediately after the data were downloaded.
3. The complete set of cultural taste labels provides a kind of “cultural fingerprint” for many students, and so these labels will be released only after a substantial delay in order to ensure that students’ identities remain anonymous.

4. In order to access any part of the dataset, prospective users must read and electronically sign [a] user agreement.
5. The entire research project, including the above steps, were reviewed and approved by Harvard's Committee on the Use of Human Subjects.

While each of these steps reveal good-faith efforts to protect the privacy of the subjects, each has serious limitations that expose a failures by the researchers to fully understand the nature of privacy in online social network spaces, and to design their research methodology accordingly. Each will be considered below, followed by a brief discussion of some of the public comments made by the T3 research team in defense of their methods and the public release of the dataset.

Use of In-Network RAs to Access Subject Data

In his defense of releasing subjects' Facebook profile data, Jason Kaufmann, the principle investigator of the T3 project, has stated that "our dataset contains almost no information that isn't on Facebook" and that "We have not accessed any information no otherwise available on Facebook" (Kaufman, 2008c). Access to this information was granted by Facebook, but only through a manual process. Thus, research assistants (RA) from the source institution (presumably Harvard) were employed to perform the labor-intensive task of search for each first year student's Facebook page and saving the profile information. The dataset's codebook confirms that "Only those data that were accessible by default by each RA were collected, and no students were contacted for additional information" (Lewis, 2008, p. 29).

The T3 codebook notes that of the 1640 students in the cohort, 1446 were found on Facebook with viewable profiles, 152 had a Facebook profile that was discoverable but not viewable by the RA, and 42 were undiscoverable (either not on Facebook or invisible to those not within their "friend" network) (Lewis, 2008, p. 6). Importantly, the codebook notes a peculiarity inherent with using in-network RAs to access the Facebook profile data:

It is important to note that both undergraduate and graduate student RAs were employed for downloading data, and that each type of RA may have had a different level of default access based on individual students' privacy settings. In other words, a given student's information should not be considered objectively "public" or "private" (or even "not on Facebook")—it should be considered "public" or "private" (or "not on Facebook") from the perspective of the particular RA that downloaded the given student's data. (Lewis, 2008, p. 6)

The T3 researchers concede that one RA might have different access to a student's profile than different RA, and being "public" or "private" on Facebook is merely relative to that particular RAs level of access.

What appears to be lost on the researchers is that a subject might have set her privacy settings to be viewable to only to other users within her network, but to be inaccessible to those outside that sphere. The RAs employed for the project, being from the same network as the subject, were able to view and download the subject's profile data. This data – originally meant for only those within the subject's network – is now included in a dataset released to the public. As a result, it is likely that profile information that a subject explicitly restricted to only "in network" participants in Facebook has been accessed from within that network, but then extracted and shared outside those explicit boundaries.

Thus, the justification that "we have not accessed any information no otherwise available on Facebook" is true only to a point. While the information was indeed available to the RA, it might have been accessible only due to the fact that the RA was within the same "network" as the subject, and that a privacy setting was explicitly set with the intent to keep that data within the boundaries of that network. This gap in the project's fundamental methodology reveals a lack of understanding of how users might be using the privacy settings within Facebook to control the flow of their personal information across different spheres, and puts the privacy of those subjects at risk.

Removal or Encoding of "Identifying" Information

In an effort to protect the identity of the subjects, researchers note that "All identifying information was deleted or encoded immediately after the data were downloaded" (Lewis, 2008, p. 29), and that "all the data is cleaned so you can't connect anyone to an identity" (Kaufman, 2008a). Yet, as the AOL search data release revealed, even if one feels that "all identifying information" has been removed from a dataset, it is often trivial to piece together random bits of information to deduce one's identity (Barbaro and Zeller Jr, 2006). The fact that the dataset includes each subjects' gender, race, ethnicity, hometown state, and major makes it increasingly possible that individuals could be identified, especially those with a particular combination of characteristics. For example, if the data reveals that a particular subject is a white male from Montana majoring in East Asian Studies, there probably are only a few who fit such a description, especially in this small of a population. Further, the numerous one-of-a-kind individuals in the dataset (for example, there is only one Hungarian, one subject from Louisiana, and one subject majoring in Slavic Studies) could be easily identified. Repeating Hargittai's concern: "I think it's hard to imagine that some of

this anonymity wouldn't be breached with some of the participants in the sample" (Hargittai, 2008).

This reveals that even when researchers believe they have removed or encoded "all identifying information," there often remains information that could just as easily be used to re-identify individuals. The T3 researchers' belief that stripping names along is sufficient resembles the typical definition of "personally identifiable information" (PII) within the United States legal framework. As defined in California law, PII is typically limited to an individual's name or other personally identifiable elements such as a social security number, a driver's license number, or a credit card number. So long as these identifiers are removed from a dataset, it is presumed to be sufficiently anonymous.

However, others take a much broader stance in what constitutes personally identifiable information. The European Union, for example, defines PII much more broadly to include:

any information relating to an identified or identifiable natural person...; an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity.

Thus, while the T3 researchers might have felt simply removing or coding the subjects' names or other specific identifiers from the dataset was sufficient, had they followed the European Union's guidance, they would have recognized that many of the subjects' "physical, physiological, mental, economic, cultural or social identity" could also be used for re-identification. Even after removing the names of the subjects, since the dataset still includes race, ethnicity, and geographic data, as well as cultural taste information, re-identification remains a distinct possibility.

Delay in Release of Cultural Taste Data

Despite the apparent lack of use of the EU's more stringent definition of "personally identifiable information," the T3 researchers do recognize the unique nature of the cultural taste labels they have collected, referring to them as a kind of "cultural fingerprint". To protect subject privacy, the cultural tastes identified by the researchers have been assigned a unique number, and only the numbers will be associated with students for the initial data releases. The entire set of the actual taste labels will only be released in the fall of 2011, corresponding with the release of the wave 4 data.

The T3 researchers are right to recognize how a person's unique set of cultural tastes could easily identify her. Yet, merely instituting a "substantial delay" before releasing this personal data does little to mitigate the privacy fears. Rather, it only delays them, and only by three years. Researchers routinely rely on datasets for years after their initial collection: some influential studies of search engine behavior rely on nearly ten-year-old data (see, for example, Jansen and Resnick, 2005; Jansen and Spink, 2005), and these subjects' privacy needs do not suddenly disappear when they graduate from college in 2011.

Most surprisingly, despite the T3 researchers' recognition of the sensitive nature of the cultural data, they will provide immediate access to it on a case-by case basis. As the codebook reveals:

In the meantime, if prospective users wish to access some subset of the taste labels, special arrangements may be made on a case-by-case basis at the discretion of the authors (send request and detailed justification to t3dataset@gmail.com). (Lewis, 2008, p. 20)

No further guidance is provided as to what kinds of arrangements are made and what justifications are needed to make such an exception. If the T3 research team felt strongly enough that it is necessary to encode and delay the release of the subjects' "cultural fingerprints", it does not seem appropriate to announce that exceptions can be made for its release to selected researchers prior to the three-year delay. If it is potentially privacy invading content, it simply should not be released.

Terms of Use Statement

As a final attempt to protect subject privacy, researchers wanting access to the T3 dataset must (electronically) sign a Terms and Conditions of Use statement. The statement includes various covenants related to protecting the privacy of the subjects in the dataset, including:

1. I will use the dataset solely for statistical analysis and reporting of aggregated information, and not for investigation of specific individuals or organizations, except when identification is authorized in writing by the Authors.
2. I will produce no links among the Authors datasets or among the Authors data and other datasets that could identify individuals or organizations.
3. I represent that neither I, nor anyone I know, has any prior knowledge of the possible identities of any study participants in any dataset that I am being licensed to use.

4. I will not knowingly divulge any information that could be used to identify individual participants in the study, nor will I attempt to identify or contact any study participant, and I agree to use any precautions necessary to prevent such identification.

5. I will make no use of the identity of any person or establishment discovered inadvertently. If I suspect that I might recognize or know a study participant, I will immediately inform the Authors, and I will not use or retain a copy of data regarding that study participant. If these measures to resolve an identity disclosure are not sufficient, the Authors may terminate my use of the dataset. (reproduced at Lewis, 2008, p. 30)

The language within this statement clearly acknowledges the privacy implications of the T3 dataset, and might prove effective in raising awareness among potential researchers. However, the overall effectiveness and enforceability of this kind of “click-wrap” terms of service remains debated (Gatt, 2002), and it is unclear how the T3 researchers specifically intend to monitor or enforce compliance with these terms. Without any “teeth”, such an agreement might have little success in deterring any potential privacy-invasive use of the data.

Further, the lengths at which the terms of service agreement details concerns over the possibility of re-identifying individual subjects seems at odds with the researchers’ repeated claims that all identifying information has been removed from the dataset. If the latter were true, there would be little need for such strong conditions within the usage agreement. This ambiguity raises concerns over the validity of the T3 research team’s expressed belief that there are no privacy threats inherent within the dataset.

IRB Approval

As required of any research project involving human interaction, clearance for the research project and data release was provided by Harvard’s intuitional review board (IRB), known as the Committee on the Use of Human Subjects in Research. As Kaufman commented: “Our IRB helped quite a bit as well. It is their job to insure that subjects’ rights are respected, and we think we have accomplished this” (Kaufman, 2008c). Elsewhere he has noted that “The university in question allowed us to do this and Harvard was on board because we don’t actually talk to students, we just accessed their Facebook information” (Kaufman, 2008a).

Just as we can question whether the T3 researchers full understood the privacy implications of the research, we must critically examine whether Harvard’s IRB – a panel of experts in research ethics – also sufficiently understood how the privacy of the subjects in the dataset could be compromised. For example, did the IRB recognize, as noted above, that using an in-network research assistant to pull

data could circumvent privacy settings intended to keep that data visible to only other people at Harvard? Or did the IRB understand that individuals with unique characteristics could easily be extracted from the dataset, and perhaps identified? It is unclear whether these concerns were considered and discarded, or whether the IRB did not fully comprehend the complex privacy implications of this particular research project. In either case, the potential privacy-invading consequences of the T3 data release suggest a failure at some point of the IRB review process.

Other Public Comments

Beyond the shortcomings of the documented efforts to protect the privacy of the T3 dataset subjects, the researchers have made various public comments that further bring doubt into their full understanding of what is at stake with this particular research project.

For example, when a New York Times article noted the debate about whether specific consent should be obtained before using Facebook profile information in studies (Rosenbloom, 2007), one of the T3 researchers characterized it as “trying to dust up some controversy” about whether their work was invading Facebook user’s privacy, and resolved that “we’re really not abusing these students at all, all the data is cleaned so you cannot connect anyone to an identity” (Kaufman, 2008a). As alluded to above, the inability to “connect anyone to an identity” has been brought into serious doubt, and concerns over privacy and consent are not just feeble attempts to “dust up some controversy,” and should be taken more seriously.

Elsewhere, when confronted with the potential re-identifiability of the dataset, Jason Kaufman has responded by pondering “What might hackers want to do with this information, assuming they could crack the data and ‘see’ these people’s Facebook info?” and later acknowledging “Nonetheless, seeing your thought process — how you would attack this dataset — is extremely useful to us” (Kaufman, 2008b). Kaufman’s mention of “hackers”, “attacking” the dataset, and focusing on what someone might “do” with this information exposes a harm-based theory of privacy protection. So long as we can protect the security of the data from attack by hackers, Kaufman seems to argue, the privacy of the subjects can be maintained. Such a position ignores the broader dignity-based theory of privacy (Bloustein, 1964), where one does not need to be a victim of hacking, or have a tangible harm take place, in order for there to be concerns over the privacy of one’s personal information. Rather, merely having one’s personal information stripped from the intended sphere of the social networking profile, and amassed into a database for external review becomes an affront to the subjects’ human dignity and their ability to control the flow of their personal information.

The distinction between harm- and dignity-based theories of privacy are understood – and often debated – among privacy scholars, but when asked if they conferred with privacy experts over the course of the research and data release, Kaufman admits that “we did not consult [with] privacy experts on how to do this, but we did think long and hard about what and how this should be done” (Kaufman, 2008c). Given the apparent focus on data security as a solution to privacy, it appears the T3 research team would have benefited from broader discussions on the nature of privacy in these environments.

Finally, the T3 researchers regularly claim that there should be little concern over the ethics of this research since the Facebook data gathered was already publicly available. As Kaufman argues:

On the issue of the ethics of this kind of research — Would you require that someone sitting in a public square, observing individuals and taking notes on their behavior, would have to ask those individuals’ consent in advance? We have not accessed any information not otherwise available on Facebook. We have not interviewed anyone, nor asked them for any information, nor made information about them public... (Kaufman, 2008c)

This justification presents a false comparison: using an in-network research assistant to access and download an entire cohort of college students’ Facebook profile pages, each year for four years, is not the same as sitting in a public square and taking notes on random people’s behavior as they happen to pass by. The T3 researchers targeted a specific and known group of students, obtained a list of names and e-mail addresses of the students from the source university, then proceeded to systematically access, download, and process their Facebook information. The data acquired included not only the subjects gender and presumed ethnicity (which could be guessed by those in the public square example), but also their home state, nation of origin, political views, sexual interests, college major, relational data, and cultural interests. The “public square” counter-example pales in comparison to the amount of personal data systematically collected, and longitudinally linked, in the T3 research project. Suggesting that the two projects are similar and carry similar (and minimal) ethical dilemmas reveals a worrisome gap in the T3 research team’s understanding of the privacy implications of their project.

Challenges for research on/in social networks

The events surrounding the release of the Facebook data in the “Tastes, Ties, and Time” project –including its methodology, its IRB approval, the way in which the data was released, and the viewpoints publicly expressed by the researchers – reveals considerable gaps in the understanding of the privacy implications of research

in social networking spaces. As a result, serious threats to the privacy of the subjects under study persist, despite the good faith efforts of the T3 research team.

The purpose of this critical analysis of the T3 project is not to place blame or single out these researchers for condemnation, but to use it as a case study to help expose the emerging challenges of engaging in research within online social network settings. These include challenges to the traditional nature of consent, properly identifying and respecting expectations of privacy on social network sites, developing sufficient strategies for data anonymization prior to the public release of personal data, and the relative expertise of institutional review boards when confronted with research projects based on data gleaned from social media.

As made apparent to the position of some of the T3 research team that their data collection methods were unproblematic since the “information was already on Facebook”, future researchers must gain a better understanding of the contextual nature of privacy in these spheres (Nissenbaum, 1998; Nissenbaum, 2004), recognizing that just because personal information is made available to in some fashion on a social network, does not mean it is fair game for capture and release to all (see, generally, boyd, 2008a; McGeeveran, 2007; Stutzman, 2006; Zimmer, 2006). Similarly, the notion of what constitutes “consent” within the context of divulging personal information in social networking spaces must be further explored, especially in light of this contextual understanding of norms of information flow within specific spheres. The case of the T3 data release also reveals that we still have not learned the lessons of the AOL data release and similar instances where presumed anonymous datasets have been re-identified. Perhaps most significantly, this case study has uncovered possible shortcomings in the oversight functions of institutional review boards, the very bodies bestowed with the responsibility of protecting the rights of data subjects.

Overcoming these challenges is no easy task, but three steps can be taken immediately to guide future research in social media spaces. One, scholars engaging in research similar to the T3 project must recognize their own gaps in understanding the changing nature of privacy and the challenges of anonymizing datasets, and should strive to bring together an interdisciplinary team of collaborators to help ensure the shortcomings of the T3 data release are not repeated. Two, we must evaluate and educate IRBs and related policy makers as to the complexities of engaging in research on social networks. And three, we must ensure that our research methods courses, codes of best practices, and research protocols recognize the unique challenges of engaging in research on Internet and social media spaces.

The “Tastes, Ties, and Time” research project might very well be ushering in “a new way of doing social science”, but it is our responsibility scholars to ensure

our research methods and processes remain rooted in long-standing ethical practices. Concerns over consent, privacy and anonymity do not disappear simply because subjects participate in online social networks; rather, they become even more important.

REFERENCES

- Albrechtslund, A.** (2008) Online Social Networking as Participatory Surveillance [Online] First Monday. Available from: <http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2142/1949> [Accessed 3 March 2008]
- Barbaro, M. & Zeller Jr, T.** (2006) A Face Is Exposed for AOL Searcher No. 4417749. The New York Times, A1.
- Barnes, S.** (2006) A privacy paradox: Social networking in the United States [Online] First Monday. Available from: http://www.firstmonday.org/ISSUES/issue11_9/barnes/ [Accessed 12 October 2007]
- Berkman Center for Internet & Society (2008, September 25) Tastes, Ties, and Time: Facebook data release [Online] Available from: <http://cyber.law.harvard.edu/node/4682> [Accessed 30 September 2008]
- Bloustein, E.** (1964) Privacy as an Aspect of Human Dignity: An Answer to Dean Prosser. New York University Law Review, 39, 962-1007.
- Boyd, D.** (2008a) Putting Privacy Settings in the Context of Use (in Facebook and elsewhere) [Online] apophenia. Available from: http://www.zephoria.org/thoughts/archives/2008/10/22/putting_privacy.html [Accessed 22 October 2008]
- Boyd, D.** (2008b) Taken Out of Context: American Teen Sociality in Networked Publics. Ph.D. thesis, University of California-Berkeley.
- Boyd, D. & Ellison, N.** (2008) Social network sites: Definition, history, and scholarship. Journal of Computer-Mediated Communication, 13, 210-230.
- Gatt, A.** (2002) Click-Wrap Agreements The Enforceability of Click-Wrap Agreements. Computer Law & Security Report, 18, 404-410.
- Grimmelmann, J.** (2009) Facebook and the Social Dynamics of Privacy. Iowa Law Review, 95, 4.
- Gross, R. & Acquisti, A.** (2005) Information revelation and privacy in online social networks. 2005 ACM workshop on Privacy in the electronic society,
- Hargittai, E.** (2008) I think it's hard to imagine... [Blog comment] Facebook Datasets and Private Chrome [Online] Unit Structures. Available from: <http://>

fstutzman.com/2008/09/29/facebook-datasets-and-private-chrome/ [Accessed 30 September 2008]

Jansen, B.J. & Resnick, M. (2005) Examining Searcher Perceptions of and Interactions with Sponsored Results. Workshop on Sponsored Search Auctions at ACM Conference on Electronic Commerce,

Jansen, B.J. & Spink, A. (2005) How are we searching the World Wide Web? A comparison of nine search engine transaction logs. *Information Processing & Management*, 42, 248-263.

Kaufman, J. (2008a) Considering the Sociology of Facebook: Harvard Research on Collegiate Social Networking [Video] Berkman Center for Internet and Society. Available from <http://cyber.law.harvard.edu/node/4408> [Accessed 30 September 2008]

Kaufman, J. (2008b) I am the Principal Investigator... [Blog comment] On the "Anonymity" of the Facebook Dataset [Online] michaelzimmer.org. Available from: <http://michaelzimmer.org/2008/09/30/on-the-anonymity-of-the-facebook-dataset/> [Accessed 30 September 2008]

Kaufman, J. (2008c) Michael - We did not consult... [Blog comment] On the "Anonymity" of the Facebook Dataset [Online] michaelzimmer.org. Available from: <http://michaelzimmer.org/2008/09/30/on-the-anonymity-of-the-facebook-dataset/> [Accessed 30 September 2008]

Lenhart, A. & Madden, M. (2007) Teens, privacy & online social networks [Online] Pew Internet & American Life Project. Available from: http://www.pewinternet.org/pdfs/PIP_Teens_Privacy_SNS_Report_Final.pdf [Accessed 20 April 2007]

Lewis, K. (2008) Tastes, Ties, and Time: Cumulative Codebook [Online] Available from: <http://dvn.iq.harvard.edu/dvn/dv/t3> [Accessed 30 September 2008]

Lewis, K. et al. (2008) Tastes, ties, and time: A new social network dataset using Facebook. com. *Social Networks*, 30, 330-342.

McGeveran, W. (2007) Facebook, Context, and Privacy [Online] Info/Law. Available from: <http://blogs.law.harvard.edu/infolaw/2007/09/17/facebook-context/> [Accessed 3 October 2008]

Narayanan, A. & Shmatikov, V. (2008) Robust de-anonymization of large sparse datasets. *IEEE Symposium on Security and Privacy*, 2008, 111-125.

Narayanan, A. & Shmatikov, V. (2009) De-anonymizing Social Networks. *30th IEEE Symposium on Security and Privacy*,

National Science Foundation (2009) 38. Sharing of Findings, Data, and Other Research Products. Grant General Conditions (GC-1), 27.

Nissenbaum, H. (1998) Protecting Privacy in an Information Age: The Problem of Privacy in Public. *Law and Philosophy*, 17, 559-596.

Nissenbaum, H. (2004) Privacy as Contextual Integrity. *Washington Law Review*, 79, 119-157.

Nussbaum, E. (2007) Kids, the Internet, and the end of privacy [Online] *New York Magazine*. Available from: <http://nymag.com/news/features/27341/> [Accessed 13 February 2007]

Rosenbloom, S. (2007) On Facebook, scholars link up with data [Online] *New York Times*. Available from: <http://www.nytimes.com/2007/12/17/style/17facebook.html?ref=us> [Accessed 30 September 2008]

Simmel, G. & Wolff, K.H. (1964) *The sociology of Georg Simmel*. Free Press, Glencoe, Ill.

Solove, D. (2007) *The Future of Reputation: Gossip, Rumor, and Privacy on the Internet*. Yale University Press, New Haven, CT.

Stutzman, F. (2006) How Facebook Broke its Culture [Online] *Unit Structures*. Available from: <http://chimprawk.blogspot.com/2006/09/how-facebook-broke-its-culture.html> [Accessed 3 October 2008]

Stutzman, F. (2008) Facebook Datasets and Private Chrome [Online] *Unit Structures*. Available from: <http://fstutzman.com/2008/09/29/facebook-datasets-and-private-chrome/> [Accessed 30 September 2008]

Sweeney, L. (2002) k-anonymity: A model for protecting privacy. *International Journal of Uncertainty Fuzziness and Knowledge-Based Systems*, 10, 557-570.

Wellman, B. & Berkowitz, S.D. (1988) *Social structures: A network approach*. Cambridge University Press Cambridge,

Zimmer, M. (2006) More on Facebook and the Contextual Integrity of Personal Information Flows [Online] *michaelzimmer.org*. Available from: <http://michaelzimmer.org/2006/09/08/more-on-facebook-and-the-contextual-integrity-of-personal-information-flows/> [Accessed 3 October 2008]

Zimmer, M. (2008a) More On the "Anonymity" of the Facebook Dataset - It's Harvard College [Online] *michaelzimmer.org*. Available from: <http://michaelzimmer.org/2008/10/03/more-on-the-anonymity-of-the-facebook-dataset-its-harvard-college/> [Accessed 3 October 2008]

Zimmer, M. (2008b) On the “Anonymity” of the Facebook Dataset [Online] michaelzimmer.org. Available from: <http://michaelzimmer.org/2008/09/30/on-the-anonymity-of-the-facebook-dataset/> [Accessed 30 September 2008].

