

PAIT Bibliography

Kenneth D. Pimple, Ph.D.¹

*Ethical Guidance for Research and Application of
Pervasive and Autonomous Information Technology (PAIT)
March 3-4, 2010, Cincinnati, Ohio*

<http://poynter.indiana.edu/pait/>

This bibliography was compiled with the assistance of Daniel Noah Robbin. It's a work in progress; suggestions for additions or corrections would be greatly appreciated.

In addition to a citation, each item includes, when available, a URL for the item, the abstract or other description, and keywords. Some URLs may be ineffective without a subscription.

6, Perri. 2001. **“Ethics, regulation, and the new artificial intelligence, part I: Accountability and power.”** *Information, Communication and Society* 4(2): 199-229.

<http://www.informaworld.com/smpp/content~db=all~content=a713768525?words=ethics|regulation|new|artificial|intelligence|perri|information|communication|society&hash=321157399>

Abstract: A generation ago, there was a major debate about the social and ethical implications of artificial intelligence (AI). Interest in that debate waned from the late 1980s. However, both patterns of public risk perception and new technological developments suggest that it is time to re-open that debate. The important issues about AI arise in connection with the prospect of robotic and digital agent systems taking socially significant decisions autonomously. Now that this is possible, the key concerns are now about which decisions should be and which should not be delegated to machines, issues of regulation in the broad sense covering everything from consumer information through codes of professional ethics for designers to statutory controls, issues of design responsibility and problems of liability.

Keywords: artificial intelligence, robotics, digital agents, technological risk, ethics, regulation, accountability

6, Perri. 2001. **“Ethics, regulation, and the new artificial intelligence, part II: Autonomy and liability.”** *Information, Communication and Society* 4(3): 406-434.

<http://bert.lib.indiana.edu:2048/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=ufh&AN=5494209&site=ehost-live>

Abstract: This is the second article in a two-part series on the social, ethical and public policy

¹ Director of Teaching Research Ethics Programs, Poynter Center for the Study of Ethics and American Institutions, Indiana University, 618 East Third Street, Bloomington IN 47405-3602; (812) 856-4986; FAX 855-3315; <http://poynter.indiana.edu/>; pimple@indiana.edu.

The PAIT project is made possible by the National Science Foundation (grant number SES-0848097), Indiana University's Poynter Center for the Study of Ethics and American Institutions, and the Association for Practical and Professional Ethics.

This document may be reproduced and used without permission for non-profit educational purposes. Permission must be requested of the author in writing for other uses. It can be found in PDF format at <http://poynter.indiana.edu/pait/links.shtml>; look for the link to “PAIT Bibliography.”

implications of the new artificial intelligence (AI). The first article briefly presented a neo-Durkheimian understanding of the social fears projected onto AI, before arguing that the common and enduring myth of an AI takeover arising from the autonomous decision-making capability of AI systems, most recently resurrected by Professor Kevin Warwick, is misplaced. That article went on to argue that, nevertheless, some genuine and practical issues in the accountability of AI systems that must be addressed. This second article, drawing further on the neo-Durkheimian theory, sets out a more detailed understanding of what it is for a system to be autonomous enough in its decision making to blur the boundary between tool and agent. The importance of this is that this blurring of categories is often the basis, the first article argued, of social fears.

Keywords: artificial intelligence, robotics, digital agents, technological risk, regulation, accountability, autonomy, ethics, institutions, judgement, decision making, Emile Durkheim, Mary Douglas, Kevin Warwick

Andrejevic, Mark. 2007. **iSpy: Surveillance and power in the interactive era**. Lawrence, Kan., University Press of Kansas.

<http://www.loc.gov/catdir/toc/ecip0716/2007016410.html> (link to bibliographic record retrieved from the Library of Congress)

Keywords: Information society. Interactive multimedia Social aspects.

Brey, Philip. 2005. “**Freedom and privacy in ambient intelligence.**” Ethics and Information Technology 7(3): 157-166.

<http://bert.lib.indiana.edu:2048/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=phl&AN=PHL2094258&site=ehost-live>

Abstract: This paper analyzes ethical aspects of the new paradigm of Ambient Intelligence, which is a combination of Ubiquitous Computing and Intelligent User Interfaces (IUI?s). After an introduction to the approach, two key ethical dimensions will be analyzed: freedom and privacy. It is argued that Ambient Intelligence, though often designed to enhance freedom and control, has the potential to limit freedom and autonomy as well. Ambient Intelligence also harbors great privacy risks, and these are explored as well.

Keywords: Ambient Intelligence, autonomy, freedom, Intelligent User Interfaces, invisible computers, privacy, smart objects, Ubiquitous Computing

Caire, Patrice. 2007. “**Conviviality for ambient intelligence.**” Proceedings of Artificial Societies for Ambient Intelligence, Artificial Intelligence and Simulation of Behaviour (AISB’07). Olivier, Patrick and Christian Kray. Newcastle upon Tyne: 14-19.

<http://icr.uni.lu/pubs/cai07c.pdf>

Abstract: Conviviality is usually considered a positive concept related to sociability, however, further analysis reveals a negative side related to regulations. In this survey paper, we examine the multifaceted concept of conviviality and raise the question: Which definition of conviviality

can be used and made operational for ambient intelligence? We propose a two-fold definition of conviviality as a condition for social interactions and an instrument for the internal regulation of social systems. We, then, propose to use conviviality for ambient intelligence as a mechanism to reduce mis-coordinations between individuals, groups and institutions, and as a tool to reinforce social cohesion. Intelligent interfaces, for example, allow instant interactions and thereby create strong needs for coordination and regulation mechanisms that have to be addressed to ensure the safeguard of individuals against abuses, such as privacy intrusions and identity manipulations. It is therefore crucial to take into account social and cognitive factors and to address the ethical issues raised by the large scale development of ambient intelligent systems.

Carew, Peter J., Larry Stapleton, et al. 2008. **“Implications of an ethic of privacy for human-centred systems engineering.”** *AI and Society* 22(3): 385-403.

<http://bert.lib.indiana.edu:2048/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=31892182&site=ehost-live>

Abstract: Privacy remains an intractable ethical issue for the information society, and one that is exacerbated by modern applications of artificial intelligence. Given its complicity, there is a moral obligation to redress privacy issues in systems engineering practice itself. This paper investigates the role the concept of privacy plays in contemporary systems engineering practice. Ontologically a nominalist human concept, privacy is considered from an appropriate engineering perspective: human-centred design. Two human-centred design standards are selected as exemplars of best practice, and are analysed using an existing multi-dimensional privacy model. The findings indicate that the human-centred standards are currently inadequate in dealing with privacy issues. Some implications for future practice are subsequently highlighted.

Carsten Stahl, Bernd. 2004. **“Information, ethics, and computers: The problem of autonomous moral agents.”** *Minds and Machines* 14(1): 67-83.

<http://bert.lib.indiana.edu:2048/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=17020492&site=ehost-live>

Abstract: In modern technical societies computers interact with human beings in ways that can affect moral rights and obligations. This has given rise to the question whether computers can act as autonomous moral agents. The answer to this question depends on many explicit and implicit definitions that touch on different philosophical areas such as anthropology and metaphysics. The approach chosen in this paper centres on the concept of information. Information is a multi-faceted notion which is hard to define comprehensively. However, the frequently used definition of information as data endowed with meaning can promote our understanding. It is argued that information in this sense is a necessary condition of cognitivist ethics. This is the basis for analysing computers and information processors regarding their status as possible moral agents. Computers have several characteristics that are desirable for moral agents. However, computers in their current form are unable to capture the meaning of information and therefore fail to reflect morality in anything but a most basic sense of the term. This shortcoming is discussed using the example of the Moral Turing Test. The paper ends with a consideration of which conditions computers would have to fulfil in order to be able to use information in such a way as to render

them capable of acting morally and reflecting ethically

Keywords: autonomous moral agent, ethics, meaning, information, morality, responsibility

Coeckelbergh, Mark. 2009. “**Virtual moral agency, virtual moral responsibility: On the moral significance of the appearance, perception, and performance of artificial agents.**” *AI and Society* 24(2): 181-189.

<http://www.springerlink.com/content/m0606p116771nv50/>

Abstract: “In this paper, I explore how we can include artificial agents in our moral discourse, without giving up the ‘folk’ intuition that humans are somehow special with regard to morality, that there is a special relation between humanity and morality - whatever that means.” (p. 181)

Emiliani, P. L. and C. Stephanidis. 2005. “**Universal access to ambient intelligence environments: Opportunities and challenges for people with disabilities.**” *IBM Systems Journal* 44(3): 605-619.

<http://proquest.umi.com/pqdlink?Ver=1&Exp=12-02-2014&FMT=7&DID=886191801&RQT=309>

Abstract: In the years ahead, as a result of the increasing demand for ubiquitous and continuous access to information and services, information technologies are expected to evolve toward a new computing paradigm known as ambient intelligence. Ambient intelligence will be characterized by invisible (i.e., embedded) computational power in everyday appliances and other common physical objects, including intelligent mobile and wearable devices. Ambient intelligence will have profound consequences on the type, content, and functionality of emerging products and services, as well as on the way people will interact with them, bringing about multiple new requirements for the development of information technologies. In addressing this challenge, the concept of universal access is critical. This paper discusses the anticipated opportunities and challenges that ambient intelligence will bring about for elderly people and people with disabilities, envisages new scenarios in the use of ambient-intelligence technologies by users with diverse needs and requirements, and identifies some of the critical issues that will have to be addressed.

Floridi, Luciano. 2004. “**Open problems in the philosophy of information.**” *Metaphilosophy* 35(4): 554-582.

<http://web.ebscohost.com/ehost/detail?vid=1&hid=5&sid=ada08c82-3563-4f64-b28d-a80a0ca0572d%40sessionmgr10&bdata=JnNpdGU9ZWhvc3QtG12ZQ%3d%3d#db=aph&AN=13678552>

Abstract: The philosophy of information (PI) is a new area of research with its own field of investigation and methodology. This article, based on the Herbert A. Simon Lecture of Computing and Philosophy I gave at Carnegie Mellon University in 2001, analyses the eighteen principal open problems in PI. Section 1 introduces the analysis by outlining Herbert Simon’s approach to PI. Section 2 discusses some methodological considerations about what counts as a good philosophical problem. The discussion centers on Hilbert’s famous analysis of the central

problems in mathematics. The rest of the article is devoted to the eighteen problems. These are organized into five sections: problems in the analysis of the concept of information, in semantics, in the study of intelligence, in the relation between information and nature, and in the investigation of values.

Keywords: artificial intelligence, computer ethics, David Hilbert, information, knowledge, philosophy of information, semantics, Herbert Simon, information theory

Floridi, Luciano and J. W. Sanders. 2004. “On the morality of artificial agents.” Minds and Machines 14(3): 349-379.

<http://bert.lib.indiana.edu:2048/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=15104066&site=ehost-live>

Abstract: Artificial agents (Aas), particularly but not only those in Cyberspace, extend the class of entities that can be involved in moral situations. For they can be conceived of as moral patients (as entities that can be acted upon for good or evil) and also as moral agents (as entities that can perform actions, again for good or evil). In this paper, we clarify the concept of agent and go on to separate the concerns of morality and responsibility of agents (most interestingly for us, of Aas). We conclude that there is substantial and important scope, particularly in Computer Ethics, for the concept of moral agent not necessarily exhibiting free will, mental states or responsibility. This complements the more traditional approach, common at least since Montaigne and Descartes, which considers whether or not (artificial) agents have mental states, feelings, emotions and so on. By focussing directly on ‘mind-less morality’ we are able to avoid that question and also many of the concerns of Artificial Intelligence. A vital component in our approach is the ‘Method of Abstraction’ for analysing the level of abstraction (LoA) at which an agent is considered to act. The LoA is determined by the way in which one chooses to describe, analyse and discuss a system and its context. The ‘Method of Abstraction’ is explained in terms of an ‘interface’ or set of features or observables at a given ‘LoA’. Agenthood, and in particular moral agenthood, depends on a LoA. Our guidelines for agenthood are: interactivity (response to stimulus by change of state), autonomy (ability to change state without stimulus) and adaptability (ability to change the ‘transition rules’ by which state is changed) at a given LoA. Morality may be thought of as a ‘threshold’ defined on the observables in the interface determining the LoA under consideration. An agent is morally good if its actions all respect that threshold; and it is morally evil if some action violates it. That view is particularly informative when the agent constitutes a software or digital system, and the observables are numerical. Finally we review the consequences for Computer Ethics of our approach. In conclusion, this approach facilitates the discussion of the morality of agents not only in Cyberspace but also in the biosphere, where animals can be considered moral agents without their having to display free will, emotions or mental states, and in social contexts, where systems like organizations can play the role of moral agents. The primary ‘cost’ of this facility is the extension of the class of agents and moral agents to embrace Aas.

Keywords: artificial agents, computer ethics, levels of abstraction, moral responsibility

Gadzheva, Maya. 2008. **“Privacy in the age of transparency: The new vulnerability of the individual.”** *Social Science Computer Review* 26(1): 60-74.

<http://ssc.sagepub.com/cgi/content/abstract/26/1/60>

Abstract: In an ambient intelligent (AmI) environment with computing capabilities embedded in potentially every object it will be difficult (if not impossible) for users to maintain control over data generation, transfer, and use and to achieve unobservability and anonymity. Obtaining consent might not be feasible for the constant need for collection and exchange of incredible amount of data. In most cases, individuals are not aware that profiling is done, what profiles are being compiled, and what decisions may result from these profiles. Due to the overflow of information users cannot maintain knowledge of all data controllers that have some of their data or exercise their right to correct, rectify, block, or erase information collected about them. The present-day privacy legislation has several weaknesses if confronted with the AmI environment that could lead to the need for new principles on which to base new regulations, to take account of the changed context.

Keywords: ambient intelligence, privacy, profiling, surveillance, transparency

Miller, Arthur R. 1993. **“Copyright protection for computer programs, databases, and computer-generated works: Is anything new since CONTU?”** *Harvard Law Review* 106(5): 978-1073.

<http://bert.lib.indiana.edu:2048/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=7736454&site=ehost-live>

Abstract: In 1976, Congress created the National Commission on New Technological Uses of Copyrighted Works (CONTU) to examine the implications of computer and other information technologies and to advise Congress on whether it would be wise to assimilate these new technologies into the existing copyright regime. Ever since, a number of people have voiced criticisms of CONTU's recommendations, arguing for a modification of the current copyright regime or even a different legal structure for these issues. Although these critics have many different approaches, they share the fear that incorporating these technologies into the current copyright system will lead to overprotection. In this Article, Professor Miller examines these arguments and concludes that CONTU's recommendations were correct and that the current regime is flexible enough to address the critics' concerns. First, Professor Miller surveys the progress of the court decisions involving computer programs, and demonstrates that a set of coherent copyright principles is beginning to develop in this area. Next, Professor Miller discusses the new phenomenon of artificial intelligence, and utilizes copyright principles that have developed in the areas of computer programs and databases to demonstrate that at its current stage of development, artificial intelligence does not pose any significant obstacles to copyright analysis. Finally, Professor Miller addresses the claim that it eventually will be impossible to assimilate computer-generated works into the copyright system because they may have no obvious human author, and concludes not only that the caselaw contains no persuasive objection to extending copyright protection to these works, but also that such an extension would fulfill the constitutional imperative of promoting progress in these areas.

Miller, Keith, Frances Grodzinsky, et al. 2009. “**Why Turing shouldn’t have to guess.**” Asia-Pacific Computing and Philosophy Conference. Tokyo.

<http://bentham.k2.t.u-tokyo.ac.jp/ap-cap09/openconf/data/papers/13.pdf>

Abstract: Soon computers will routinely pass the Turing Test. We discuss a normative question separate from the descriptive question of feasibility: Should human developers program computers to masquerade as humans? We exclude from our discussion scientific experiments and entertainment applications. We contend that in many of the remaining cases, developers should not program machines to masquerade as humans.

Keywords: humanoid robots, ethics, Turing Test

Nissenbaum, Helen. 2010. **Privacy in context: Technology, policy, and the integrity of social life**. Stanford, Calif., Stanford Law Books.

Abstract: From the back cover of the paperback edition:

“Privacy is one of the most urgent issues associated with information technology and digital media. This book claims that what people really care about when they complain and protest that privacy has been violated is not the act of sharing information itself - most people understand that this is crucial to social life - but the inappropriate, improper sharing of information.

“Arguing that privacy concerns should not be limited solely to concern about control over personal information, Helen Nissenbaum counters that information ought to be distributed and protected according to norms governing distinct social contexts - whether it be the workplace, health care, schools, or among family and friends. She warns that basic distinctions between public and private, informing many current privacy policies, in fact obscure more than they clarify. In truth, contemporary information systems should alarm us only when they function without regard for social norms and values, and thereby weaken the fabric of social life.”

Keywords: Privacy, Right of United States. Information technology Social aspects United States. Information policy United States. Social norms.

Perusco, Laura and Katina Michael. 2007. “**Control, trust, privacy, and security: Evaluating location-based services.**” Technology and Society Magazine 26(1): 4-16.

http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?tp=&arnumber=4135773&isnumber=4135358

Abstract: Location-based services (LBS) are those applications that utilize the position of an end-user, animal, or thing based on a given device (handheld, wearable, or implanted), for a particular purpose. This article uses scenario planning to identify the possible risks related to location-based services in the context of security and privacy. The original contribution of this article is that the dilemma has been related specifically to LBS, under the privacy-security dichotomy. Here, each side of the dichotomy is divided into three key components that combine to greatly magnify risk. Removing one or more components for each set decreases the privacy or security risk. Where more elements are present in conjunction, the risk is increased.

Shin, Dong-Hee. 2009. “**Ubiquitous city: Urban technologies, urban infrastructure and**

urban informatics.” *Journal of Information Science* 35(5): 515-526.

<http://jis.sagepub.com/cgi/content/abstract/35/5/515>

Abstract: South Korea continues to lead the way in digital opportunity with its recent, innovative and ubiquitous city projects. The u-city initiative in South Korea is a national urban development project that focuses on strengthening the role of information and communication technologies in civic planning and management. This study tracks the changing dynamics driving the information society initiative of South Korea to evaluate the process of design and development of u-city. This study reviews qualitative data related to the u-city projects, describes the transformations and translation of this data in the public, political, and social discourse, and discusses the prospectus of a ubiquitous information society environment. The findings raise fundamental, practical questions about the role of ubiquitous computing in shaping our future cities. The findings show that there are more challenges ahead than prospects, despite the fact that the u-city has all the advanced technological components for a positive development. The South Korean u-city is typically more prone to problems related to the lack of social infrastructure, market restrictions, political quagmires and vested financial interests. The paper discusses the deficiencies of the South Korean approach, namely a lack of holistic approach by integrating technological possibilities with social application needs.

Keywords: urban informatics, ubiquitous city, South Korea

Shin, Dong-Hee and Won-Yong Kim. 2008. “**An analysis of Korean national Information strategy of IT839.**” *41st Annual Hawaii International Conference on System Sciences.*

<http://www2.computer.org/portal/web/csdl/doi/10.1109/HICSS.2008.48>

Abstract: Recently, the Korean government launched the IT839 project with the objective of converting Korea into a ubiquitous information society. This study investigates the role of the Korean government in the development of the national information infrastructure and the realization of IT839 vision. This paper reviews the historic and individual data related to the infrastructure project, draws on the social construction of technology theory as a framework for interpreting such data, traces the transformations and translation of this data in the public, political, and social discourse, and discusses the next generation of information infrastructure. Findings imply that despite optimistic prospects and proactive drive, uncertainty still remains with respect to where IT839 will evolve and how it will impact the new information milieu.

Turing, Alan M. 1950. “**Computing machinery and intelligence.**” *Mind* 59(236): 433-460.

<http://www.jstor.org/stable/2251299>

Abstract: “I propose to consider the question, ‘Can machines think?’ This should begin with definitions of the meaning of the terms ‘machine’ and ‘think’. The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words ‘machine’ and ‘think’ are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, ‘Can machines think?’ is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.”

Wright, David, Serge Gutwirth, et al. 2007. “**Shining light on the dark side of ambient intelligence.**” *foresight* 9(2): 46-59.

<http://www.emeraldinsight.com/Insight/viewPDF.jsp?contentType=Article&Filename=html/Output/Published/EmeraldFullTextArticle/Pdf/2730090205.pdf>

Abstract: Purpose - To identify safeguards against threats and vulnerabilities posed by the emerging world of ambient intelligence.

Design/methodology/approach - Dark scenarios were constructed to highlight the threats and vulnerabilities; safeguards are identified to address those threats and vulnerabilities and recommendations for specific stakeholders are made for implementing those safeguards.

Findings - A multiplicity of threats and vulnerabilities can be expected in the emerging world of ambient intelligence, and a multiplicity of safeguards are similarly necessary to address those.

Research limitations/implications - Additional technological research is necessary in order to develop some of the safeguards envisaged as necessary.

Practical implications - The existing legal and regulatory regime suffers from various lacunae and must be amended to address AmI challenges. Many stakeholders have important roles to play.

Originality/value - The paper identifies necessary safeguards to protect privacy, identity, trust, security and e-inclusion. It identifies specific recommendations for the European Commission, Member States, industry, civil society organizations, academia and individuals.

Wright, David, Serge Gutwirth, et al. 2008. **Safeguards in a world of ambient intelligence**, Springer.

<http://www.springerlink.com/content/978-1-4020-6661-0>

Abstract: This book is a warning. It aims to warn policy-makers, industry, academia, civil society organisations, the media and the public about the threats and vulnerabilities facing our privacy, identity, trust, security and inclusion in the rapidly approaching world of ambient intelligence (AmI).

In the near future, every manufactured product - our clothes, money, appliances, the paint on our walls, the carpets on our floors, our cars, everything - will be embedded with intelligence, networks of tiny sensors and actuators, which some have termed “smart dust”. The AmI world is not far off. We already have surveillance systems, biometrics, personal communicators, machine learning and more. AmI will provide personalised services - and know more about us - on a scale dwarfing anything hitherto available.

In the AmI vision, ubiquitous computing, communications and interfaces converge and adapt to the user. AmI promises greater user-friendliness in an environment capable of recognising and responding to the presence of different individuals in a seamless, unobtrusive and often invisible way. While most stakeholders paint the promise of AmI in sunny colours, there is a dark side to AmI.

This book aims to illustrate the threats and vulnerabilities by means of four “dark scenarios”. The

authors set out a structured methodology for analysing the four scenarios, and then identify safeguards to counter the foreseen threats and vulnerabilities. They make recommendations to policy-makers and other stakeholders about what they can do to maximise the benefits from ambient intelligence and minimise the negative consequences.

Keywords: ambient intelligence, future scenarios, privacy, risk assessment, ubiquitous computing