

## Privacy and Property on the Net: Research Questions

William Sims Bainbridge

The Net, which consists of the Internet plus the parallel networks of commercial and military data communication, raises serious issues of personal information vulnerability and ownership of intellectual property. Peer-to-peer sharing of music files over Internet, for example, challenges the existing system of copyright. Counterterrorism “data mining” systems hunt for signs of suspicious activity across vast databases of information about people who do not wish to be spied upon. Research by social and information scientists is needed to clarify many such issues.

To a sociologist, perhaps the most interesting quality of Internet and other data transmission networks is their potential to alter power relationships with respect to personal privacy and intellectual property. Both are restrictions on the free flow of information. Government is implicated in both, on the one hand by providing legal support for them and, on the other hand, by potentially violating them in the pursuit of its own goals, such as national security. Social scientists have only just begun framing research projects to learn how the Internet is liberating information from traditional restraints or to understand the likely human consequences.

Although both privacy and property are rooted deep in humanity’s evolutionary past, they are variables, and societal norms change. For example, until a century and a half ago, U.S. census takers would post their completed enumeration forms in the town square for anyone to read, but from 1850 until 1954 when Title 13 of the U.S. Code forbade publication of an individual’s records, a complex series of steps gradually increased confidentiality protections (1). Today, the Census Bureau keeps the data confidential for 72 years. This change has largely been driven by the increasing government collection of data about its citizens, to facilitate social services, taxation,

and management of the economy (2). By offering confidentiality, government hopes the public will relinquish some of its traditional privacy. Whether citizens benefit from government collection of data about them is another matter. In Maryland, every prospective juror is asked what his or her religion is, even though this information is not used in the jury selection process. The religion data are kept confidential, but all information about prospective jurors that is used in the selection is made public. Thus, we have the bizarre situation of costly information being kept

confidential precisely because it is useless.

The idea that government should regulate intellectual property through copyrights and patents is relatively recent in human history, and the precise details of what intellectual property is protected for how long vary across nations and occasionally change. There are two standard sociological justifications for patents or copyrights: They reward creators for their labor, and they encourage greater creativity. Both of these are empirical claims that can be tested scientifically and could be false in some realms (3, 4).

Consider music (5). Star performers existed before the 20th century, such as Franz Liszt and Niccolò Paganini, but mass media produced a celebrity system promoting a few stars whose music was not necessarily the best or most diverse. Copyright provides protection for distribution companies and for a few celebrities, thereby helping to

Scientists and the general public are being confronted by the changes in concepts of confidentiality and ownership caused by the rise of the Internet.

support the industry as currently defined, but it may actually harm the majority of performers. This is comparable to Anatole France’s famous irony, “The law, in its majestic equality, forbids the rich as well as the poor to sleep under bridges.” In theory, copyright covers the creations of celebrities and obscurities equally, but only major distribution companies have the resources to defend their property rights in court. In a sense, this is quite fair, because nobody wants to steal unpopular music, but by supporting the property rights of celebrities, copyright strengthens them as a class in contrast to anonymous musicians.

Image not available for online use.

The author is in the Division of Information and Intelligent Systems, the Directorate for Computer and Information Science and Engineering, National Science Foundation, 4201 Wilson Boulevard, Arlington, VA 22230, USA. E-mail: william.bainbridge@verizon.net

gists would hypothesize they thereby learn contempt for laws in general. A poll by the Pew Internet and American Life Project found that two-thirds of an estimated 35 million Americans who download music files do not care whether they are copyrighted (7). Thus, on the level of families, ending copyright could be morally as well as economically advantageous. On a much higher level, however, the culture-exporting nations (notably the United States) could stand to lose, although we cannot really predict the net balance of costs and benefits in the absence of proper research. We do not presently have good cross-national data on file sharing or a well-developed theoretical framework to guide research on whether copyright protection supports cultural imperialism versus enhancing the positions of diverse cultures in the global marketplace.

**Image not available for online use.**

It will not be easy to test such hypotheses, and extensive economic research has not conclusively answered the question of whether the patent system really promotes innovation. We will need many careful, sharp-focus studies of well-formed hypotheses in specific industries and sectors of life. For example, observational and interview research can uncover the factors that really promote cultural innovation among artists of various kinds and determine the actual consequences for children of Internet peer-to-peer file sharing.

The classic sociological theory of privacy asserts two main principles that are relevant to current information debates (8, 9). First, private relations between individuals are not merely a personal luxury but the very basis of all societies. Thus, surveillance of citizens in the name of national security undermines the very society it was intended to protect. Second, privacy and its violation are inescapably issues of power, with power defined as the ability to defend one's own privacy while being able to invade the privacy of others. Thus, citizens' privacy is violated by government surveillance, even if the information is kept confidential afterward by the agencies.

Computer scientists have begun to develop systems that would defend people's information privacy (10). For example, the World Wide Web consortium, which is the chief forum for development of Web standards, has launched the Platform for Private

cy Preferences, which automatically manages personal information when interacting with Web sites, following the explicit wishes of the user (11).

Ethically informed research can be valuable for design and implementation of information systems. Data mining, for example, can be used in noncontroversial situations like intrusion detection, when you need to defend your own data against attack from outside. But it can also be used aggressively to sift through vast troves of data, pulled together via the Net in a process called data fusion, in a way that violates the privacy of law-abiding citizens while hunting for a few criminals or terrorists. In its recent report, *Information Technology for Counterterrorism*, the National Research Council blithely suggests that all airport security baggage x-ray machines could send their pictures to a unified computer network that would monitor the collective movements of terrorists boarding different airplanes in different cities, incidentally spying on everybody else who flies (12). Proper sociotechnical design could limit the harm. For example, an automatic data-mining system could seek patterns of suspicious behavior, without allowing any human being to see the data. Then, specially sworn court officials could carefully examine the suspicious cases, before reporting to law enforcement only those few that met statutory definitions of probable cause.

What is the optimal design for such a multitiered confidentiality system? Would there be unintended consequences? Would it be acceptable to the public? Without good answers to such questions, information technology could fall under the same cloud of public suspicion that hangs over nuclear power and genetic engineering. Already, aggressive telemarketing has made people very reluctant to answer scientific surveys administered by telephone (13). Internet-based surveys have promise, but they typically lack the reliability of traditional random-sample polling. Another challenge is that impartial research on the real consequences of government security-motivated data fusion and mining may not be possible within the government security regime itself, because impartiality in science requires the research to be made public.

Like other buzzwords, "data mining" is

difficult to define precisely, and it is practically indistinguishable from a wide range of statistical and pattern recognition techniques used throughout the sciences. Thus, if it is brought into disrepute by privacy violations, science could be harmed. Similarly, many scientists use data that belong to other people, such as the social survey data archived at the Inter-University Consortium for Political and Social Research (14), so they would be affected by any change in the intellectual property regime. Internet arose as a medium of scientific communication and is an indispensable tool for scientists today. Thus, research on Internet privacy and property issues could benefit science itself, as well as society.

Only an anarchist would argue for complete government deregulation of information, but any social scientist would find it interesting to contemplate the implications of major changes in government enforcement of intellectual property rights and government collection of data about its citizens. This essay has briefly sketched theoretical issues that would underlie research on information privacy and property in the age of Internet, and policy-makers should be aware that current practices may reflect cultural lag, the conflict-ridden situation when technological development has rendered traditional norms obsolete (15).

#### References and Notes

1. R. McCaa, S. Ruggles, *Scand. Pop. Stud.* **13**, (2002).
2. D. Bell, *The Coming of Post-Industrial Society* (Basic Books, New York, 1973).
3. C. A. Ganz-Brown, *J. World Intellect. Property* **1**, (1998).
4. National Research Council, Computer Science and Telecommunications Board, *The Digital Dilemma: Intellectual Property in the Information Age* (National Academy Press, Washington, DC, 2000).
5. W. S. Bainbridge, in *Social Sciences for a Digital World*, M. Renaud, Ed. (Organization for Economic Co-Operation and Development, Paris, 2000).
6. Ipsos-Insight, "Legal issues don't hinder American downloaders" (14 March 2003); [cited from www.ipsos-reid.com 5 May 2003]; available at www.ipsos-insight.com.
7. M. Madden, S. Lenhart, "Music downloading, file-sharing and copyright" (July 2003) at www.pewtrusts.com.
8. B. Schwartz, *Am. J. Sociol.* **73** (1968).
9. W. S. Bainbridge, "Privacy," forthcoming in *Encyclopedia of Community*, K. Christensen and D. Levinson, Eds. (Sage, Thousand Oaks, CA, 2003).
10. National Research Council, Computer Science and Telecommunications Board, *Who Goes There?: Authentication Through the Lens of Privacy* (National Academy Press, Washington, DC, 2003).
11. World Wide Web Consortium at www.w3c.org.
12. National Research Council, Computer Science and Telecommunications Board, *Information Technology for Counterterrorism* (National Academies Press, Washington, DC, 2003).
13. D. El Boghdady, *Washington Post*, 8 September 2002, p. H1.
14. Inter-University Consortium for Political and Social Research at www.icpsr.umich.edu.
15. W. F. Ogburn, *Social Change* (Huebsch, New York, 1922).
16. The views expressed in this essay do not necessarily represent the views of the National Science Foundation or the United States.