Navigating Computer Science Research Through Waves of Privacy Concerns: *Discussions among Computer Scientists at Carnegie Mellon University*

Latanya Sweeney, editor

July 2003
CMU-ISRI-03-102

Also appears as Computer Science Technical Report
CMU-CS-03-165

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213-3890

Abstract

Computer Science research and practice are raising growing privacy concerns among the public and government. Computer technology's increasing ability to capture, organize, interpret and share data about individuals raises questions about what privacy practices computer science researchers should adopt, if any. These issues are already very real in ongoing research projects in the School of Computer Science (SCS) at Carnegie Mellon University, from mining databases of individual transactions, to studying how people use the web, to mounting cameras in lounges, to building hallway robots that capture data about passers by, to building intelligent workstation assistants that learn user habits. This paper introduces the nature of privacy concerns often related to computer science research, explains potential benefits and risks (especially of abuse and misuse) and examines traditional and innovative methods for providing privacy assurances in research. Examples are provided from projects at Carnegie Mellon's School of Computer Science.

Keywords: Computer Science Research, privacy, Institute Review Board, de-identification, policy

1. Research Introduces New Problem

Recent news articles ignited public concern over issues of privacy in emerging technologies. Concern is founded, but when presented in alarmist ways, public and legislative response can be so drastic as to inhibit scientific research and progress and thwart possible benefits to society. Privacy concerns are new to almost all computer science researchers. Attitudes are varied. By understanding privacy issues, perceptions, regulations, and laws coming to bear on computer science research, computer science researchers can make informed, educated modifications in research practices. This paper is provided as a beginning examination based in part on the first of two brainstorming sessions with computer science faculty at Carnegie Mellon University's School of Computer Science.

> Carnegie Mellon may be ideally positioned to make a constructive response. Specifically, a coalition of Carnegie Mellon's data collection, data interpretation/inferencing, and data privacy faculty could take a public position that acknowledges the legitimate concerns, explains the potential benefits and associated risks (especially of abuse and misuse), shows a technical basis for a reasonable policy that permits (and encourages) progress, and proposes ground rules for both research and practice that provide safeguards against the risks. [Mary Shaw, Carnegie Mellon, 2003]

Carnegie Mellon finds itself in a unique position. On the one hand, it has traditionally pioneered new emerging technologies, so it is not surprising to find active research on many new emerging technologies related to data mining, robotics, face recognition and personal assistants. However, Carnegie Mellon is the only university to have an active research group devoted to developing privacy technology. The Data Privacy Laboratory works with real-world stakeholders to construct technology for analyzing risks and solving privacy problems.

2. Described Uses of Emerging Technologies Incite Privacy Action

As technology continuously changes, so does the nature of some computer science research. Perhaps the technology and research are changing so fast that the consequences are more than the public, reporters, lawmakers, or even computer scientists can realize or prepare for. Here are some reactions in the news.

> The Defense Advanced Research Projects Agency (DARPA) stimulates outside-the-box thinking that has given us the Internet and the stealth bomber. On occasion, however, DARPA goes off half-cocked. Its Total (now Terrorist) Information Awareness (TIA) plan to combine all commercial credit data and individual bank and academic records with F.B.I. and C.I.A. dossiers, which would have made every American's life an open book has been reined in somewhat by Congress after we privacy nuts hollered to high heaven. [William Safire, "Dear Darpa Diary," New York Times, June 5, 2003]

These kinds of characterizations of the TIA project by Safire and other reporters generated serious privacy concerns.

In January 2003, Senator Feingold, Democrat-Wisconsin, introduced legislation to place a moratorium on data mining research and deployment efforts at the U.S. Department of Defense. Senator Wyden, Democrat-Oregon, introduced a similar anti-data mining bill, but it was limited

to TIA. A broad coalition of public interest groups, ranging from the American Civil Liberties Union to the American Conservative Union urged Congress to take action against TIA. [Ryan Singel, "Bills: Down with Citizen Database," Wired, Jan 17, 2003]

Congress did respond, but in doing so, data mining research was spared what could have been a horrible blow to computer science research funding in this area. Privacy concerns are not limited to TIA, but are inherent in many emerging technologies.

Two years earlier, in January 2001, police in Tampa, Florida tested face recognition technology during Super Bowl XXXV, scanning faces of people in crowds, comparing them with images in a database of digital mug shots. In August of that year, a councilwoman in Jacksonville, Florida introduced legislation banning the use of face recognition technology by the Sheriff's Office and other city agencies. [Dibya Sarkar, "Florida City Moves to Ban Face-recognition System," USA Today, Aug 23, 2001]

Similar legislation was predicted to emerge in other city councils and legislatures, but the events of September 11, 2001, dramatically reversed the projected down-turn expected for funding in face recognition research. These recent encounters are viewed as warnings by some computer science researchers, who worry that future encounters may lead to serious loss in research support and who themselves have concerns over the perception and deployment of the technology they create.

> The media is currently hypersensitive to privacy issues stemming from various DARPA projects. Recent editorials spill-over concern over the TIA program into other DARPA programs. In the latest case of DARPA's Lifelog project, cool research is getting superficially and incorrectly slandered as Orwellian. [Robert Collins, Carnegie Mellon, 2003]

Technologies that record and/or share person-specific information are at the center of these privacy concerns. These technologies are emerging in part due to the robust nature of new algorithms, operating on ever smaller machines having faster processors with larger, less expensive storage capacities, and communicating over ubiquitous networks.

> Not too long ago simply having video surveillance was not a big deal to most people unless one was using the local 7-11 as a late night ATM. However, current computational engines and networks enable human identification and other biometric software to continuously screen imagery and other sensors and correlate that information with non-directed/opportunistic search of ill-structured data. [Dave McKeown, Carnegie Mellon, 2003]

3. Computer Science Research Promises Fruitful Benefits

Not all areas of computer science research are affected by privacy issues. The areas having to address privacy concerns typically involve technologies that record and/or require sharing person-specific information.

In the crudest classifications possible, computer science research can be divided into: (1) theoretical computer science, which has a close relationship to logic and mathematics; (2) programming languages and systems, which concerns the general development and operation of physical computer systems and networks; and, (3) artificial intelligence (AI), which has a long-term vision of producing machines that can think, reason and function comparable to humans, but

whose short-term visions often center on constructing smarter machines. Most of the research facing privacy concerns fits into the crudely classified third group, AI. Some research in the systems group, on network administration and security, also faces privacy issues. Little to no research in theoretical computer science faces privacy concerns, with the noteworthy exception of cryptography.

Historically, privacy problems faced by cryptographers have often made allies of computer science researchers, privacy advocates and libertarians, who together supported technology that provided secure, private communications shielded even from the government.

> [Immediately following] the terrorist attacks in New York City and Washington D.C. on September 11, 2001, [were]... renewed calls among some lawmakers for restrictions on the use and availability of strong encryption products. ... Civil liberties and privacy advocates strongly oppose any attempts to require key escrow, key recovery or other means of accessing encryption keys, arguing that they are an unjustified restriction of individuals' fundamental privacy rights, detrimental to security, costly, subject to massive abuse, and ultimately ineffective crime prevention methods. Technology and security experts [including researchers] also oppose any restrictions on encryption, arguing that they would damage consumer trust in e-commerce transactions. [Electronic Privacy Information Center (EPIC), www.epic.org/crypto/, 2003]

In more fine-grained classifications, computer science research involving human-computer interaction, personal robots and assistants, biomedical applications, data mining, sensor technology, ubiquitous computing, cybersecurity, and data privacy (a new emerging area aimed at providing technical solutions to privacy problems) are more likely today to face privacy controversies than are any other research areas in computer science.

It is important to appreciate that past research in AI resulted in many notable discoveries.

> Laboratories whose focus was AI first conceived and demonstrated such well-known technologies as the mouse, time-sharing, high-level symbolic programming languages (Lisp, Prolog, Scheme), computer graphics, the graphical user interface (GUI), computer games, the laser printer, object-oriented programming, the personal computer, email, hypertext, symbolic mathematics systems (Macsyma, Mathematica, Maple, Derive), and, most recently, the software agents which are now popular on the World Wide Web. [David Waltz, "Artificial Intelligence: realizing the ultimate promises of computing." NEC Research Institute and the Computing Research Association. 1996]

Numerous societal gains may be realized by the very same research currently generating privacy concerns. Reliable face recognition technology, for example, can have many worthy purposes.

> Improved face recognition would have benefits in fighting terrorism. Accurate face recognition could confirm identity at border crossings, airports, and secure government buildings and sites. In the private sector it would be useful for verifying identity in credit card transactions, ATM machines, and for access control to buildings. [Henry Schneiderman, Carnegie Mellon, 2003]

Capturing video data on daily life activities can endow research that may substantially enhance human interactions with computational environments and robots.

> Video understanding of human beings is an emerging area of computer vision. Success in this area will lead to more natural human-computer interfaces, such as "smart spaces" that are capable of understanding human actions and providing assistance relevant to a person's goals. Mobile robots must understand human form and action, in order to successfully interact with (or even to avoid running over) humans in the scene. A social robot that interacts with the same group of people every day should recognize the identity of each individual, and be able to recall the interactions it has had with that person in the past. Due to the high intrinsic value of human beings, a robotic system can never be said to be intelligent until it is capable of observing, recognizing and interpreting the actions of the people around it.
>
> Fundamental research into the realm of human activity understanding is hampered by a lack of recorded video data. Current databases for activity analysis involve staged activities carried out by subjects who are aware that they are being observed. The self-consciousness of a subject performing a scripted activity leads to unnatural body motions and behavior. To stimulate work on a deeper level, observing and recording people performing routine daily activities in public spaces is needed. The goal is to collect a series of databases of natural human behavior to spur development of video understanding algorithms that can observe a space and quickly identify the temporal patterns of human activity in that space (where people are likely to go, and what they are likely to do).
>
> Work on recognizing human activity also has obvious applications to the development of automated security systems. There are immediate needs for such systems in commercial, law enforcement and military applications. Although surveillance cameras are already prevalent in banks, stores, and parking lots, video data currently is used only after the fact as a forensic tool, thus losing its primary benefit as an active, real-time medium. Automated video understanding algorithms hold the promise of providing constant real-time monitoring of surveillance video to alert security officers to a burglary in progress, or a suspicious individual loitering in the parking lot, while there is still time to prevent the crime. Non-commercial applications include patrolling national borders, logging routine maintenance tasks in nuclear facilities, and providing secure perimeters around military bases and embassies. [Robert Collins, Carnegie Mellon, 2003]

DARPA's new RADAR project at Carnegie Mellon might put a useful, intelligent clerical assistant on every computer.

> If successful, the DARPA-funded, EPCA/RADAR project would have the same effect as hiring two full-time support staff for each computer user in the country. The goal is to provide the same kind of support that a personal human assistant would provide -- improving the user's productivity by helping to find and organize information, scheduling meetings on the user's behalf, etc. The difference is that the RADAR software would provide this support rather than hiring people (most computer users obviously cannot afford to hire full-time support staff). The impact would be a broad-based boost to worker productivity across the private sector, as well as to government workers who use computers in

their everyday work (e.g., intelligence analysts, law enforcement professionals, medical professionals). [Tom Mitchell, Carnegie Mellon, 2003]

Personal assistants may someday free humans from many tedious and complicated tasks currently required to organize and schedule personal and professional life.

> Modern computer users spend a large amount of time searching for desired information and organizing the information they currently maintain on a desktop. Executives in companies hire secretaries to manage much of this information, but every day users must wrestle with e-mails, documents, and the World-Wide-Web to acquire the information they need. Personal computational assistants can help users in these tasks. For example, the RADAR project is building the following assistants:
>
> Information Assistant: Track down key information desired by the user, relying on a number of external resources, including search engines, user profile filters, access to past email and documents, task-based summarization, and even emerging question answering technology. The information assistant will automatically organize the information discovered for the user and provide assistance in the construction of a WWW site for the user.
>
> Scheduling Assistant: Combine the calendars and preferences of multiple users to schedule meetings, briefings, seminars, and other activities whether the agents learn reasonable times, durations, prioritizations, and conflict resolution strategies.
>
> Email assistant: Focus on sorting, classifying, managing emails, including automatically responding to scheduling requests (via the activity scheduler), learning what to filter out, grouping content-related emails and summarizing them (particularly useful for rapid catching up after a trip or other down time), etc. [Dan Siewiorek, Carnegie Mellon, 2003]

Computers capable of managing personal communication may dramatically improve productivity (and possibly reduce stress) by helping with the bombardment of phone calls, faxes, and email messages that typifies the contemporary office.

> A hallmark of modern managerial and professional work is that it is communication intensive [9]. Managers and professionals have large numbers of spontaneous communication sessions with multiple partners over the course of a single work day [12]. Modern technologies, including electronic mail, instant messaging, pagers, wireless email devices, and cell phones, have made communication more convenient in more locations, providing many people the fresh and rich information they need to do their jobs [6].
>
> However, informal, spontaneous communication comes at a cost-interruption. Because of interruptions, managers often think through important issues in three minute blocks of time. Perlow's [11] fieldwork among software engineers shows that engineers grab help from their colleagues and the interruptions cascade, leading to crises of productivity. Tetard [17] demonstrates that interruptions disrupt ongoing thought, and O'Connaill and Frohlich [8] show that managers fail to return to the interrupted activity almost 50% of the time.

> Developing intelligent assistants to help overburdened managers and professionals handle their communication may immediately enhance the productivity of valuable managers and professionals in government, industry and academia. [Robert Kraut, Carnegie Mellon, 2003]

Notions of ubiquitous computing, whether realized as networks of sensors and cameras or as coordinated secondary sharing of collected data, can yield personal benefits to aging populations and those living with debilitating diseases.

> The goal of [the Personal Robotic Assistants for the Elderly ("nursebot")] is to develop mobile, personal service robots that assist elderly people suffering from chronic disorders in their everyday life. We are currently developing an autonomous mobile robot that "lives" in a private home of a chronically ill elderly person. The robot provides a research platform to test out a range of ideas for assisting elderly people, such as:
>
> Intelligent Reminding: Many elderly patients have to give up independent living because they forget. They forget to visit the restroom, to take medicine, to drink, or to see the doctor. Our project explores the effectiveness of a robotic reminder, which follows people around (hence cannot get lost).
>
> Tele-presence: Professional care-givers can use the robot to establish a "tele-presence" and interact directly with remote patients. This makes many doctor visits superfluous. Our robot is a platform for tele-presence technology that connects patients with care-givers through the Next Generation Internet (NGI).
>
> Data collection and surveillance: A range of emergency conditions can be avoided with systematic data collection (e.g., certain types of heart failures). This reason alone can make service robots succeed in the home care business.
>
> Mobile manipulation: Arthritis is the main reason for elderly to give up independent living. A semi-intelligent mobile manipulator, that integrates robotic strength with a person's senses and intellects, can overcome barriers in manipulating objects (refrigerator, laundry, microwave) that currently force patients to move into assisted living facilities.
>
> Social interaction: A huge number of elderly people are forced to live alone, deprived of social contacts. The project seeks to explore whether robots can take over certain social functions.
>
> Two factors suggest that now is the time to establish mobile robots in the home-care sector: First, for the first time we actually have the technology together to develop robots that exhibit the necessary robustness, reliability, and level of capability. Second, the need for cost-effective solutions in the elderly care sector is larger than ever before. [Nursebot Project, www-2.cs.cmu.edu/~nursebot/, 2003]

These research endeavors require sharing unprecedented amounts of person-specific information with computer science researchers in order for promised technology to be realized. Improvements to face recognition algorithms may require studying video images of human

subjects gathered from real-world settings, implying a need for a lack of individual consent, and comparing captured images to driver license or identification card photos, requiring that these images be shared. Work needed to construct personal assistants capable of organizing and scheduling personal and professional events may require sharing archives of email messages and electronic schedules with researchers, initiating concerns as to whether the consent of all parties listed on the email messages or included in scheduled events is needed, and if so, the practical nature of the effort involved in acquiring them. Research to build communication assistants may require studying electronically captured phone, fax, email and in-person communications, generating concerns related to wire-tapping, video, and surveillance laws.

> To develop some parts of RADAR will require access to a moderately large corpus of realistic -- but not necessarily real -- E-mail traffic. We are pursuing several possible methods for obtaining such a corpus: using a synthetic data set, using real mail that has been anonymized, or using project-internal mail for which all parties have given consent. We envision that RADAR, when deployed, will simply help its individual user to organize information already in that user's possession. [Scott Fahlman, Carnegie Mellon, 2003]

4. Harms in Research are Real

Notwithstanding the fruitful results that may be realized from all areas of computer science research, privacy disasters in research have been documented. History points to events in medicine.

> It is important to distinguish two ways in which a research initiative might be objectionable or problematic: subjects experience a physical, social, or economic HARM, OR, subjects are WRONGED in the sense that their interests in privacy, for example, are violated (even if they do not experience a setback or damage as a result). The Jewish Chronic Disease Hospital Case provides a nice example of the latter situation. Previous studies had shown that people with cancer take longer to expel foreign cancer cells from their bodies than do otherwise healthy people. Dr. Chester Southam and his colleagues wanted to measure the rate at which foreign cancer cells would be rejected from the bodies of people who were suffering from illnesses other than cancer. To do this, they injected foreign cancer cells into the skin of twenty-two debilitated residents of the Jewish Chronic Disease Hospital in Brooklyn, New York. In order to ensure sufficient enrolment in their study, the researchers did not inform subjects of the nature of the injections-that they contained live cancer cells-because they wanted to avoid irrational fears about the word "cancer." To the researchers, the importance of the research and the low probability of harm provided sufficient reason to justify the use of chronically debilitated hospital patients without further burdening them with irrational worries about contracting cancer. To members of the Board of Regents of the University of the State of New York, however, these considerations did not obviate the fact that the researchers subjected vulnerable individuals to non-therapeutic procedures without their informed consent, thereby violating the subjects' rights to control the disposition of their person and to be free from unwanted molestation.
>
> Several points about this case are worth emphasizing. First, none of the research subjects experienced a physical harm from the injections. But they were nevertheless wronged by the breach of their privacy interests. Second, the Board

> of Regents was especially concerned about this wrong for an additional reason. That is, such wrongs damage the trust of the public in medical research and researchers, threatening the long-term interests of both researchers and the larger community. Third, this is a case where both the researchers' interests in conducting the study, and the residents' privacy interests, could both have been fully met if the researchers had been willing to take the privacy interests of participants into account. Fourth, it should be noted that it was in response to the public outcry over this and similar cases that federal oversight mechanisms for human subjects research were instituted. Finally, it is essential to retain a balanced view of this case. In particular, the outcry over this case should not be seen as an anti-research sentiment. This was not "medical research gone mad" as was the case with many of the Nazi experiments. This was sound science pursuing an important hypothesis. But it was carried out in a way that showed a distinct lack of respect for the legitimate interests of those whose lives it touched. The lesson: research is important and should be fostered, but it should also be carried out in a way that is properly respectful of the legitimate interests of those whose lives it touches. Abuses in medical research have led to today's institutional practices of having research which involves human subjects be reviewed by a panel called the Institute Review Board (IRB). [Alex London, Carnegie Mellon, 2003]

Historical events leading to today's IRB process began most notably in the 1920's with Nazi Germany, were further fueled by reports of abuses in the United States in the 1960's and 1970's, and culminated in the Belmont Report, whose ethical principles are embodied by today's IRB process. Regulation requiring IRB review of research involving human subjects was promulgated in 1981.

> In 1946, 23 Nazi physicians went on trial at Nuremberg for crimes committed against prisoners of war. These crimes included exposure of humans to extremes of temperature, performance of mutilating surgery, and deliberate infection with a variety of lethal pathogens. During the trial, fundamental ethical standards for the conduct of research involving humans were codified into the Nuremberg Code. … The two most important conditions [set forth] were the need for voluntary informed consent of subjects and a scientifically-valid research design that could produce fruitful results for the good of society…
>
> [In the 1960s America's general attention to human and civil rights increased, at a time when newspapers reported that researchers in New York injected elderly, indigent people with live cancer cells, without their consent, as described earlier.]
>
> [In the 1970s it was revealed that] since the 1930's, approximately four hundred black men in Tuskegee, Alabama, had been involved, without their knowledge, in a lengthy study (the Tuskegee Syphilis Study) on the natural history of syphilis. These men were systematically denied penicillin even after its introduction as the standard treatment for the disease…
>
> In 1979 the Belmont Report was published identifying fundamental ethical principles for the protection of human subjects. [National Institutes of Health, Guidelines for the Conduct of Research Involving Human Subjects at the NIH, 1995].

The Belmont Report contains three ethical principles that form the cornerstone of today's IRB process. (1) "Respect for Persons" acknowledges the dignity and autonomy of individuals. Human subjects give informed consent to participate in research. Vulnerable populations, such as the elderly, the young, or the disadvantaged often require additional protection. (2) "Beneficence" protects individuals by seeking to have research maximize possible benefits while minimizing possible harms. Careful examination of research design, including alternative ways of conducting the research, is undertaken to realize the best research benefit possible while minimizing risk to humans. (3) "Justice" treats subjects fairly. Subjects should be carefully and equitably chosen to insure that certain classes of individuals are not systematically selected or excluded unless there are scientific reasons for doing so. Examples include minorities and women.

In 1981, the principles founded in the Belmont Report were promulgated as Title 45, Code of Federal Regulations, Part 46, Protection of Human Subjects (45 CFR 46), requiring IRB approval for all federally-funded research involving human participants. Research institutions receiving substantial federal research dollars must provide an IRB panel to review and approve all research conducted at the institution involving human subjects. IRB approval must be granted before research begins.

Today's IRBs insist that computer science research protocols must address fundamentally the same risks (e.g., violation of privacy, legal risks, psychosocial stress) and provide the same level of protection as any other type of research involving human participants. This belief is not sparked by any incident specific to computer science research as much as by what has become a tradition of protection required in federally funded research, as learned and inherited from the past.

The appropriateness of fit with computer science research is not clear, but no official complaints or opinions have surfaced. With the rapid evolution of computer technology and fundamental changes to computer science research itself, constant advances are likely to pose ongoing challenges to the IRB process.

At present, all studies involving human subjects, including those in computer science, must (a) ensure that the procedures fulfill the principles of voluntary participation and informed consent, (b) maintain the confidentiality of information obtained from or about human participants, and (c) adequately address possible risks to participants including privacy risks.

5. Privacy Concerns are New to Computer Science Researchers

Unlike researchers in medical, psychological, and social sciences, who are accustomed to IRB review, computer scientists have historically not had to address privacy issues within research practices. Traditionally, computer science research involved abstract and hypothetical examples. Outside of work on computational tools for education, such as Papert's Logo [10], and on expert systems in medicine, such as Mycin [3], virtually no computer science research involved the kind of real-world, person-specific data and observations of individuals that gives rise to today's privacy concerns in some recent computer science research projects.

This shift in computer science research is due in great part to two trends: (1) the field's increasing ability to capture and share large volumes of person-specific information and (2) the field's increasing development of methods to use that information to develop more useful machines.

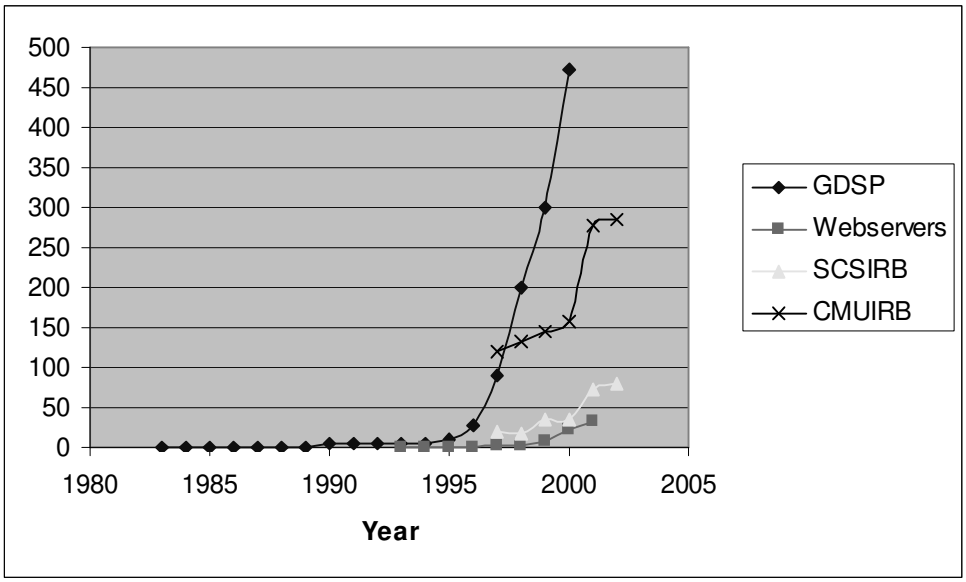

Figure 1. Comparison of growth in global disk storage per person (GDSP), active webservers, and research protocols reviewed by Carnegie Mellon's IRB. GDSP is reported in megabytes (MB) per person, and webservers is reported in millions. SCSIRB is the number of SCS research protocols reviewed, while CMUIRB is the total number of protocols reviewed.

In an attempt to characterize and compare the growth in disk storage capacity, active webservers and computer science cases reviewed by Carnegie Mellon's IRB, data were collected from several sources and plotted in Figure 1.

Global disk storage per person (GDSP) is the amount of rigid disk drive space sold in a year divided by the adult world population. It provides a measure of the amount of disk storage available to capture person-specific information. This information has grown substantially over time. Similarly, the number of active websites on the World Wide Web has grown tremendously. These are both plotted in Figure 1 [16].

The total number of research protocols reviewed by Carnegie Mellon's IRB, including those from SCS, has grown over time, as shown in Figure 1. Increases were also reported in the number of cases from social science departments from 81 in 1997 to 137 in 2002, the business school, 4 in 1997 to 18 in 2002, and the information science (IS) and policy school, from 8 in 1997 to 15 in 2002. [Susan Shingle, Institute Review Board, Carnegie Mellon, 2003]

There are very few data points to examine in Figure 1, but inspection implies some interesting trends. Significant growth in the number of computer science research protocols reviewed by the IRB was experienced.

> SCS is a factor in the increase in the number of protocols being submitted to the IRB. However, it is not the only factor. There is a general trend that suggests SCS accounts for increasingly larger percentages of the protocols submitted to the IRB. A linear fit has an R^2 of significance of ≈ 0.81 (where 1 is a perfect correlation). This fit is mainly hindered by 1998's total, which appears to be an outlier. Removal of this point provides an R^2 of ≈ 0.9. The total for 1999 appears a little high, but the general growth continues after that year.

> Plotting the number of protocols reviewed by the IRB versus GDSP with no time lag provides a linear fit with R^2 of almost 1, and of ≈0.94, when IRB totals are shifted back 3 years. However, the absence of GDSP data for 2001 and 2002 and the shape of the curve of the IRB cases make it difficult to make any sweeping claims on relationships between GDSP and the number of IRB cases. Both are growing, but it is inconclusive that GDSP is the main contributing factor for the total growth in IRB protocols reviewed. [Bradley Malin, Carnegie Mellon, 2003]

6. What's a Computer Scientist to Do?

Reactions by computer science researchers to privacy issues in their research can be characterized by three overlapping positions. In this writing, these are crudely tagged as: (1) "technology trumps privacy;" (2) "technology research is policy neutral;" and, (3) "computer scientists take responsibility."

In the "technology trumps privacy" position, computer science researchers take stock in past accomplishments and computational benefits enjoyed by society, thereby relying on a belief that if society is forced to chose, it will chose advancements in computer technology over privacy.

> We are losing our privacy not for some nefarious purpose but for the best of reasons: safety, security, and generally improved quality of life. Indeed, surveillance cameras in public spaces, parking garages, and around our neighborhoods can deter crime and speed its solution. Cameras keep tabs even on the police themselves – helping avert the next Rodney King incident. Video cameras, linked to the Web, let us check traffic, and they offer the peace of mind that comes from seeing our kids having fun at home or preschool while we're at work. And video cameras are just the most conspicuous component of the rapidly propagating surveillance infrastructure: at more and more schools, Internet usage is monitored to ensure students don't visit pornography sites; similar systems watch over employees to guard against illicit actions and possibly against harassment, sexual or otherwise. [Robert Buderi, "Our Surveillance Nation." Technology Review, 106 (3) April 2003]

Some go further and recommend avoiding any attention to privacy concerns until after the benefits of the deployed technology are realized. Warnings against this position have been voiced, cautioning that unforeseen dangers could be unleashed forever.

> We have put ourselves in a position where even though the motivations are good, the outcome could be chillingly bad. Not only governments, but also private groups and individuals, have real opportunities to amass an unprecedented amount of information about each of us – the people with whom we have associated, what we have bought, where we've been, maybe even where we are. Such data, gathered for seemingly benign reasons, could nonetheless be used for suppression of free speech, idea control, and to finger people for some future crime – purposes that rumble the foundations of free societies. [Robert Buderi, "Our Surveillance Nation." Technology Review, 106 (3) April 2003]

In the "technology research is policy neutral" position, computer science researchers do not contemplate any privacy or social implications that may be inherent in the construction or existence of the technology they seek to build. Instead, these computer scientists want to pursue their research, leaving any related privacy issues to social scientists, policy makers, lawyers, and

others. Warnings have been voiced that such positions are themselves human value decisions, and computer science researchers cannot escape making them.

> Many of us when we design and implement computer technologies focus on making the technology work -- reliably, efficiently, and correctly. Rarely do we focus on human values. Perhaps we believe in value-neutral technology. Perhaps we believe that issues of value belong only to social scientists, philosophers, or policy makers. Neither belief is correct. In their work, computer system designers necessarily impart social and moral values. Yet how? What values? Whose values? For if human values are controversial, then on what basis do some values override others in the design of, say, algorithms, network security, and databases? ... [There are] a collection of human values -- freedom from bias, autonomy, trust, informed consent, accountability, accessibility, access, and moral personhood -- that have been central to … design criteria and methodologies that help to account for human values in the system design process. [Batya Friedman, Human Values and the Design of Computer Technology, Cambridge University Press, 1997]

In the "computer scientists take responsibility" position, computer scientists want to take the initiative to educate others about risks and benefits to ensure balanced research guidelines and best practices.

> At present, there are approximately 7,000 pieces of legislation before Congress and the state legislatures dealing with privacy issues. Many of them are very poorly thought out. Emotion is running fairly high on privacy, with new revelations each week that give even rational minds something serious to ponder. Computer scientists can take a leadership role for the computer science research community by working out a compromise position.
>
> The difficulty in such situations is that neither side is willing to recognize the legitimate need nor concerns of the other and so both take extreme positions. If computer scientists are going to do work at Carnegie Mellon that can potentially be seen as privacy-invading, or providing tools that can be used to invade privacy, then computer scientists here at Carnegie Mellon should take a step back and develop a set of research guidelines stating what can and cannot be done. [Michael Shamos, Carnegie Mellon, 2003]

Some believe assuming responsibility is a necessary condition to insure viability of future technology.

> Most computer scientists can no longer afford to do their work in an ivory tower and rely on the social scientists and lawyers to make decisions about limits of its use. First, policy makers and lawyers may not fully understand the technology. Second, decisions will often be made as a reaction to biased or sensationalized public opinion. Third, policy decisions are often crude and sub-optimal, and tend to legislate over simple technical remedies. Finally, there is a horrible temporal mismatch -- policy can be a function of years but new technology is a function of months, so policy enacted on today's technology may be totally inappropriate for tomorrow's and policy supporting technology today can prohibit it tomorrow. Computer scientists can and must insulate their creations from such risk. [Latanya Sweeney, Carnegie Mellon, 2003]

7. Privacy Technology Can Help

If developments in computer technology have raised privacy issues, then many believe computer technology can be instrumental in resolving them.

Two kinds of privacy issues arise in computer science research: (1) privacy issues inherent in applications of developing technology; and, (2) privacy issues related to information or practices needed to develop technology. These are interlinked. Reactions to editorials and news accounts of uses can spawn reactions that threaten research funding. IRB approval for research practices provides no assurances or legal protections for applications. However, in both cases, privacy technology can help.

Privacy technology refers to existing and emerging technologies that provide privacy assurances in the collection, sharing and use of person-specific information. The oldest form of privacy technology stems from computer security, but today's privacy concerns are much broader and include policy enforcement and techniques for rendering data anonymous.

> Traditional areas of computer security concern authentication, authorization, and encryption. Technologies in these areas provide assurances that a parcel of information is only viewed or shared with a person (or machine) authorized to view or receive it. To avoid eavesdropping or snooping, the information may be encrypted, thereby limiting access to those having proper electronic keys. Using traditional computer security technology to store, send and receive parcels of person-specific information can provide some privacy assurances against intruders and snoopers, but the bulk of privacy concerns appearing in the press cannot be resolved by computer security approaches.
>
> Voiced privacy issues relate to information given away without illegal intrusion or theft. Sensitive information about individuals (e.g., finances, health, beliefs, and habits) can be learned from sharing seemingly innocent information often by linking shared information to other available data. Anxiety has also been expressed over ways to enforce and monitor promised uses of shared information. Technologies that address these problems are part of an emerging area called data privacy, and described problems relate to works in "data anonymity" and "privacy rights management", respectively. "Anonymity" in computer security is not data anonymity, but instead relates to the ability to send/receive a parcel of information anonymously. Data anonymity technology, in contrast, controls what inferences can be learned from the content of shared information. Data privacy is a new and ripe area for computer science research, especially for those researchers developing emerging technologies that ignite privacy concerns.
>
> A 1994 Harris-Equifax consumer privacy survey focused on how the American public feels about having their medical records used for medical research and how safeguards would affect their opinions about such systems and uses. This survey implies strong public support for sharing personal health information in which persons contained in the information cannot be identified at all. The goal of pioneering work in data anonymity is to construct technology such that person-specific information can be shared for many useful purposes with

scientific assurances that the subjects of the data cannot be re-identified. The result is termed de-identified data.

A growing number of laws and regulations, as well as case law decisions, concerning person-specific data sharing allow de-identified data to be shared more freely. A recent example is the Health Insurance Portability and Accountability Act (HIPAA), which describes in detail permissions for receiving and handling personal health information. Hospitals, physician offices, and others are bound by HIPAA. If personal health information held by them is to be generally shared with others, then the receiving party must qualify for sharing and must agree to be bound by parts of the regulation. Severe civil and criminal penalties are attached. However, HIPAA includes a provision that if the personal health information is sufficiently de-identified, using accepted scientific standards, then the de-identified data can be shared free of additional agreements and threat of penalties. Similarly, human subjects regulation that governs IRBs provides an exemption for research involving scientifically de-identified data.

Data anonymity offers tremendous promises to researchers and society by allowing needed information to be shared freely while privacy (identity) is protected. This is no easy task because it takes very little information to uniquely identify individuals. For example, 87% of the U.S. population is uniquely identified by {date of birth (month, day, year), gender, 5-digit ZIP}. Work in this area is very new and the first research group devoted to this area is Carnegie Mellon's Data Privacy Lab. We already have some solutions and noted progress.

Within months, we will put into real-world practice de-identification technology to enable sharing of de-identified data for bio-terrorism surveillance. Emergency room visits and other healthcare encounters will be reported daily to the state's public health department under the authority of public health law. Collected health information will be filtered in real-time by a self-contained machine called a CertBox, which automatically edits combinations of fields (often demographics) so that released information relates to many people ambiguously. Settings are preset for a specific population and set of data fields and then sealed to prohibit tampering. CertBox technology de-identifies health information in accordance to the scientific standard of de-identification allowed under HIPAA. The resulting de-identified data is then shared with bio-terrorism surveillance systems. CertBox technology (more generally termed a "privacy appliance" by DARPA) allows us to certify that resulting data are properly de-identified and to warranty that resulting data remain practically useful for anomaly detection algorithms in bioterrorism surveillance. Since September 11, 2001, the American public has felt it must chose between safety and privacy. CertBox technology demonstrates that the American public can now enjoy both safety and privacy.

Other data privacy technologies are maturing. We know how to automatically detect and remove (or replace) explicitly identifying information in unrestricted text, such as letters, notes and email discussions [15]. We know how to sufficiently de-identify facial images in video surveillance data such that no matter how good face recognition technology may become, the identity of the de-identified face images cannot be reliably recognized without due process or equivalent safeguard, even though the resulting video may still reveal suspicious

behavior [7]. More work is needed before de-identifying video surveillance data will be ready for real-world deployment.

Our Data Privacy Lab is not the only place where significant work in this area is being done. At Purdue University, researchers have developed coordinated data sharing methods in which sensitive information can be protected by carefully partitioning and coordinating releases across a network of data holders [5, 17]. Researchers at Stanford are examining computational ways to de-identify DNA sequences. Pursuits are underway to develop real-time protocols for distributing algorithms and queries across a network of data holders so that aggregate information can be computed across the network without actually sharing person-specific data values. Some theoretical solutions are published using multi-party computation [4]. Increasingly more work on data mining algorithms that protect from disclosing unusual inferences is the topic of annual computer science workshops. ACM is hosting its second Workshop on Privacy in Electronic Society. Research is also emerging from ATT Research, IBM Research, and Microsoft Research, independently, to extend ways to specify policy preferences and enforce them [1, 2, 3]. Teresa Lunt at Parc has a research grant from DARPA to develop a privacy appliance for general counter-terrorism surveillance.

Privacy technology can assist computer science researchers in acquiring IRB approval in two ways: (1) reducing privacy risks to human subjects during the research process; and, (2) increasing possible societal benefits of the developed technology. If a research effort can be conducted with only scientifically de-identified field structured data, textual documents, or facial images, for example, then doing so lowers privacy risks and additionally, renders the protocol eligible for an IRB exemption. Protocols that include developing privacy protections within the research of an emerging technology may increase the research's perceived benefits to society, thereby weighing it more favorably in the IRB's risk versus benefit analysis.

Some of the best privacy protections are apt to be realized when developed within the emerging technology. Research funds must be made available for researchers who are developing emerging technologies that pose privacy concerns to also develop integrated privacy solutions. Proposals should include privacy technology options.

For example, a personal assistant sending email messages to assist me in scheduling meetings should not reveal sensitive information about me inappropriately. I should have expressible controls and stated assurances of what will (or will not) be shared. If instead, the personal assistant actively reports my activities to others (a boss or the government) without my knowledge or consent, as pictured by Safire, then rather than merely helping me, the assistant also helps others observe me and may compromise the privacy of those with whom I communicate. The difference between being my helper or my keeper (relating to my ability to control the technology or be controlled by it) initially relies on decisions made by researchers who develop the technology and the funding agencies that support them. Technology aligned with society's traditional norms for freedoms results when privacy protection tools are integrated within emerging

> technology as it emerges, thereby making adoption easier and the technology more powerful. [Latanya Sweeney, Carnegie Mellon, 2003]

This paper has introduced the nature of privacy concerns related to computer science research, and examined the traditional framework of IRB review for providing privacy assurances in research, as well as introduced innovative approaches to solving privacy problems as part of computer science research. This manifold protection is in emerging privacy tools integrated into research and in emerging methods of research attentive to privacy concerns. Future work dealing with the relationship between computer science research, emerging technologies and humans will have much to address. IRBs are a fairly fixed part of any institution that receives federal funds for research and much can be learned about them. Computer science research increasingly involves more human subjects and much can be learned about that development. Technology resulting from computer science research can protect or thwart privacy and much can be learned from monitoring outcomes. The idea of protecting the privacy of a human subject is certainly not new, but research concerning privacy protections afforded the virtual identity of individuals, as realized in the data people leave behind, is very new. What happens in the intersection between computer science, data privacy, and human subjects is an urgent issue. The outcome will affect those topics and all of society.

Acknowledgements

Tremendous gratitude goes to Marshall Warfield for editorial suggestions and overall assistance. The un-cited views expressed in this document are for information purposes, discussion generation, and are those of the editor alone. They do not necessarily reflect the views of any other faculty member at Carnegie Mellon University (CMU) nor do they express any official position of CMU or the School of Computer Science (SCS). The editor thanks the following people (alphabetically) for contributed ideas, comments, suggestions and support: Chris Atkeson, Sylvia Barrett, Robert Collins, Scott Fahlman, Ralph Gross, Alexander Hauptmann, Robert Kraut, Alex London, Bradley Malin, Dave McKewon, Tom Mitchell, Jack Mostow, Henry Schneiderman, Mary Shaw, Susan Shingle, Mel Siegel, Dan Siewiorek, Michael Shamos, Anthony Tomasic, and Benjamin Vernot. No claim is made that these people agree with all statements presented in this document. Appreciation also goes to the SCS faculty for creating an environment that made this work possible. Support for this work was provided by the Data Privacy Laboratory at CMU and the editor appreciates the administrative support of Sherice Livingston for meetings and note-taking and Kishore Madhava for assistance in surveying CMU IRB cases. This work is part of a collection of writings in honor of Latanya Sweeney's receipt of the distinguished Patient Advocacy Award from the American Psychiatric Association.

References

[1] Cranor, L., Langheinrich, M., Marchiori, M., Presler-Marshall, M. and Reagle, J. "The Platform for Privacy Preferences 1.0 (P3P1.0) Specification." W3C Recommendation. 16 April 2002.

[2] Cranor, L. and Wenning, R. Why P3P is a Good Privacy Tool for Consumers and Companies. GigaLaw.com. April 2002.

[3] Davis, R., Buchanan, B., and Shortliffe, E. Production rules as a representation for a knowledge based consultation program. Artificial intelligence. 1977, v8, pp. 15-45.

[4] Lindell, Y. and Pinkas, B. “Privacy preserving data mining.” In CRYPTO. Springer-Verlag, August 20-24 2000.

[5] Kantarcioglu, M. and Clifton, C. “Privacy-preserving distributed mining of association rules on horizontally partitioned data.” In ACM SIGMOD: Workshop on Research Issues in Data Mining and Knowledge Discovery (DMKD). Madison, Wisconsin, June 2, 2002.

[6] Mintzberg, H. "The Manager's Job: Folklore and Fact" Leadership. Robert P. Vecchio ed. 1997.

[7] Newton, E., Sweeney, L. and Malin, B. Preserving Privacy by De-identifying Facial Images. Carnegie Mellon University, School of Computer Science, Technical Report, CMU-CS-03-119. Pittsburgh: 2003.

[8] O’Connaill, B., & Frohlich, D. “Timespace in the workplace. Dealing with interruptions.” Proceedings of the Conference on Human Factors in Comupting Systems (CHI’95), Denver, CO, 1995.

[9] Panko, R.R., “Managerial Communication Patterns.” Journal of Organizational Computing, 2, 1992, 95-122.

[10]Papert, S. "A Computer Laboratory for Elementary Schools." Computers and Automation. 1972

[11]Perlow, L. “The Time Famine: Toward the Sociology of Work Time.” Administrative Science Quarterly, Vol. 22, 1999, Issue 3, p. 287.

[12]Reder, S. and Schwab, R. “The communicative economy of the workgroup.” Office: Technology & People, 1998, v.4.3, pp. 77-196.

[13]Schunter, M. and Powers, C., “The Enterprise Privacy Authorization Language (EPAL 1.1).” IBM 2003 (Available at www.zurich.ibm.com/security/enterprise-privacy/epal/).

[14]Steinbock, B., Arras, J. and London, A., eds. Ethical Issues In Modern Medicine, Sixth Edition, New York: McGraw Hill, 2002.

[15]Sweeney, L. “Replacing Personally-Identifying Information in Medical Records, the Scrub System.” In: Cimino, JJ, ed. Proceedings, Journal of the American Medical Informatics Association. Washington, DC: Hanley & Belfus, Inc., 1996.

[16]Sweeney, L. “Information Explosion.” Confidentiality, Disclosure, and Data Access: Theory and Practical Applications for Statistical Agencies, L. Zayatz, P. Doyle, J. Theeuwes and J. Lane (eds), Urban Institute, Washington, DC, 2001.

[17]Tetard, F. ”On Fragmentation of Working Time: a Study of Causes and Effects on Work Interruptions.” IAMSR Research Report. 9/1999.

[18]Vaidya, J. and Clifton, C. “Privacy preserving association rule mining in vertically partitioned data.” In ACM SIGKDD. Edmonton, Alberta Canada, July 23 - 26, 2002.