NSF SBE/CISE Workshop on Cyberinfrastructure and the Social Sciences

SESSION: **3: Malevolence**

LEADS: Stephen E. Fienberg <fienberg@stat.cmu.edu>,
Shankar Sastry <sastry@eecs.berkeley.edu>

PARTICIPANT NAME: **Joan Feigenbaum**
PARTICIPANT DEPARTMENT: **Computer Science**
PARTICIPANT INSTITUTION: **Yale University**
PARTICIPANT EMAIL: **Joan.Feigenbaum@Yale.EDU**

SESSION QUESTIONS:

The Internet has led to a true revolution in communication. It provides online access to databases that only a short while ago were available to a privileged few. It supports rapid and inexpensive text-based communication in the form of electronic mail and instant messaging. In addition, it provides for the sharing of traditional information and databases as well as visual and auditory information and, to a degree, even kinesthetic information, allowing for more meaningful and realistic communicative interactions. Other advances that rely on information technology are having profound effects on the communicative experience. These include mobile phones, PDAs, distributed and embedded sensors, ubiquitous and affective computing, digital imaging and music, wearable computers, GPS devices, innovative display technologies, etc. These changes have affected not only how we conduct science, but they also have had a profound effect on many aspects of our lives, including commerce, education, health care, and other behavioral, social and cultural activities. The impact of these advances can also be very broad, for example helping to alter national boundaries and, hence, contributing to globalization.

But the very nature of the interconnected cyberworld offers a host of new opportunities for inimical behavior on the part of individuals and groups who are intent on abusing the information to which we now have access. Such malevolent behavior conflicts with the privacy and other rights of individuals and organizations whose information is shared, either in an open or a restricted fashion. Below are some questions that are intended to get you to reflect not only on the current issues surrounding malevolent behavior but also what we should be thinking about in the ever evolving cyber world.

We ask each of you to prepare a two-page statement that will address these questions and raise other issues. These statements will provide the basis for our discussion and eventually for the final product of our session during the workshop. We look forward to your statements. Please do not hesitate to contact either of us should you need any additional information. Thank you!

Questions

1. Malevolence means difference things to different intellectual communities. Explain how you interpret the word as it relates to cyberinfrastructure and describe one or more critical issues in dealing with malevolence where we can benefit from collaborative research at the interface of computer science and social science/statistics.

I and my collaborators in the PORTIA project (see http://crypto.stanford.edu/portia) use the term "sensitive data" to refer to data that can harm their owners, users, or subjects if they are misused. Of course "misuse" may not always be easy to define or detect, but, roughly speaking, it is use of the data for purposes clearly inconsistent with those for which they were created and/or collected. It may be productive to think of "malevolence" as "intentional misuse." Of course, malevolence may affect more than data; intentional misuse of any network resource (that is, use that is inconsistent with the purpose for which the resource was deployed by its owner or by whoever is responsible for it) can be considered malevolence. Research at the interface of computer science and the social sciences can help clarify the rights and responsibilities of people and organizations that use the cyberinfrastructure and can help define and implement notions such as "ownership" and "intention" in this context.

2. What features of the current cyberinfrastructure create the biggest opportunities for malevolent behavior and what type of research would allow us to begin to deal with the resulting problems.

Two of the longest lived, most robust trends in cyberinfrastructure are the ever-decreasing cost of data storage and the ever-increasing ubiquity of computers and networks in business, government, recreation, and many other aspects of daily life. Thus, more and more sensitive data about people and organizations are created, captured, and stored. It is a fundamental (and often beneficial) fact that, once data are stored on general-purpose computers, they can be used for anything; unfortunately, that includes malevolence. Interdisciplinary research that enables the formulation of appropriate-use policies, their encoding in machine-readable and machine-checkable form, and their consistent use throughout the lifetime of the relevant data would allow us to begin to deal with this problem.

3. What cyber developments are likely to raise new privacy-protection issues and how can we as a society prepare for them?

We can expect the aforementioned trends to continue. We can also expect sensor-nets and surveillance systems of all sorts to become commonplace. Once again, we should strive to accompany sensitive data by easy-to-understand and easy-to-enforce policy metadata. Note that there has already been some good work on "privacy policies," (*e.g.*, the P3P work of the World Wide Web consortium – see http://w3c.org/P3P), but there has been little or no work on pushing these policies through all of the information technology that sensitive data encounter throughout their lifetimes and making sure that these policies are enforced. It is admirable for an Internet retailer to post a comprehensive privacy policy on its website, but this policy will do no good if it is confined to the website while the data are used in the company's back-end data-processing systems.

4. Is the culture and structure of the cyber world at odds with the protection of privacy and if so what can we do about this?

Yes. In addition to the aforementioned policy work, new laws are needed that give people and organizations well defined rights to control their sensitive data and that impose harsh penalties on those who use such data malevolently. I heard Mark Hill of the University of Wisconsin make the following pertinent remark at a PI meeting in June 2004: There is less and less important information that is truly inaccessible by would-be malevolent actors; the key metric is the cost of access, and that cost is decreasing steadily.

YOUR OWN THOUGHTS:

Terminology is important, and the term "privacy" may be misleading at this point.: It is important to note that much of the sensitive information that is proliferating at great rates is not "private" in a traditional or intuitive sense of that word. No one could object to the use of "private" to describe information that should, by its very nature, be known only to one or a few people – one's private thoughts, private family life, private sex life, private communication with friends, *etc.* Sensitive medical or financial information is of a different nature: There are many (sometimes thousands of) people and machines that have legitimate reasons to access it. For good reasons, the term "private" is often equated with the terms "confidential" or "secret." Thus, use of this term naturally leads security researchers to try to "solve the privacy problem" with encryption or other techniques designed to *hide* information altogether from those without authorization to access it. However, hiding information will not help to solve the problem that most sensitive data objects (and their owners, subjects, and users) encounter, namely the fact that data created and/or collected for legitimate purposes can later be used for illegitimate purposes. We should be cautious about promulgating the use of the term "malevolent," at least until we are sure that it is not similarly misleading.

Protection of sensitive data is not a "problem" that can be definitively "solved.": As technologists and researchers, we expect to make progress and, in the best cases, to "solve" today's technical problems and then move on to tomorrow's. This is an unrealistic expectation when it comes to rights and responsibilities in cyberspace and, in particular, to the protection of sensitive data. Technological, social, and political change will pose new threats to privacy even as (or before!) the old threats are dealt with, and people will not stop complaining when they are intruded upon or defrauded by malevolent actors in cyberspace. Marc Rotenberg of the Electronic Privacy Information Organization (see http://www.epic.org) made a useful analogy at the 2005 Annual Meeting of the American Association for the Advancement of Science: Invasion of privacy is like environmental pollution. It will not be completely eliminated, but nor can it be ignored. Just as environmental-impact analyses must be part of any real-world development project, privacy-impact analyses must be part of any cyber-world development project. Research at the interface of computer science and the social sciences can develop the tools needed for these analyses.