Privacy in Digital Environments: Empowering Users

danah boyd[1], Carlos Jensen[2], Scott Lederer[3], David H. Nguyen[2]

[1] Sociable Media Group
MIT Media Lab
Cambridge, MA 02139
zephoria@media.mit.edu

[2]College of Computing
Georgia Institute of Technology
Atlanta, GA 30332
[carlosj | dnguyen]@cc.gatech.edu

[3] Group for User Interface Research
University of California, Berkeley
Berkeley, CA 94720
lederer@cs.berkeley.edu

ABSTRACT

This workshop seeks to address the privacy needs and concerns of users in the design of digital environments, whether they be websites, collaborative calendar systems, collaborative work environments, online communities, communications systems or ubiquitous computing environments. Each of these settings faces real and pressing challenges when it comes to protecting user privacy. (For an overview of the problems in different areas see: [15], [12], [4], [2]).

We shall seek answers to the following questions: What can we, as designers, do to increase user awareness of what our environments are doing [6], and how user information is collected and used [7]? How can we empower users to manage the ways in which they are represented in the environments, or to limit their exposure when needed? This challenge spans the breadth of CSCW systems, and beyond.

INTRODUCTION

People are deeply concerned about their privacy, and are quite adept at defining limits and maintaining barriers in the physical world. Yet, in the digital world we are no longer good privacy managers. Our motivation and interest does not disappear in the transition from the physical to the digital; the systems we use strip us of the power to become effective privacy managers.

Often these failures are attributable to us as designers and developers; our systems do not always provide access to information users need to make informed decisions about their privacy [3]. At other times we overload users with too much information, making managing their privacy too much of a burden. Even when the correct information is presented, users have no leverage; either they accept our terms or stop using our systems. Tools enabling users to control their privacy are often added as an afterthought. When users are allowed to manage their information or level of exposure, we often fail to provide the level of detail they need. At times, our interfaces are technology driven, and map poorly to users' mental models. As system designers, we have paid little attention to protecting user privacy, and even less to empowering them to take charge of their own privacy.

Some websites displaying privacy policies pay little more than lip service to the Fair Information Practice Principles laid out by the FTC in its 1998 online privacy report to congress [8] [9]. Where disclosure of privacy practices does occur, these disclosures are often incomplete [2]. Policies typically address the technologies and concerns that companies want to express, not the set of facts users need to make informed decisions. Users face a limited choice: either accept the current policy or leave the web-site. Even if users never consult the policy, consent is assumed. In fact, the simple act of loading a sites web-page implies consent to the site's policy. There is no transparency or enforcement. Not only is it difficult for users to determine how data is being used; they have virtually no recourse if a privacy violation is detected. These issues must be addressed.

THEME

There are three common definitions of privacy:

1) The right to be left alone [18]

2) Control of personal information [19]

3) Encrypted data and communications [11]

In this workshop, participants will focus on privacy from the following perspectives: the control of personal information and the right to be left alone. Although encryption is an important mechanism employed to secure private information, it will not be the focus of this workshop as there are entire conferences devoted to it.

While our focus on privacy does span the breadth of CSCW systems, we will provide examples in ubiquitous computing, online environments, and collaborative calendar systems.

Ubiquitous Computing

Emerging technologies will allow cell phone service providers to make your location information available to third parties [17]. How can we design devices and services that inform the user, in a natural and intuitive way, about

the recipients of this information and the ways it will be used?

As a development of e-commerce, companies have been tracking individuals for the purpose of marketing. Yet in light of September 11, this data is being reused in the search for terrorists without the knowledge of the observed [15]. Should users have the ability to say when and how their data can be used?

Widespread inconspicuous sensing and computation may put people under near-constant observation [12]. The accumulation and correlation of such data can contribute to richly detailed profiles of people's lives. People have demonstrated concern over the distribution and use of observation records generated in closed environments [1]. What will people's concerns be about enhanced surveillance and tracking on a grand scale? How much control will people have over such observation, and what technical means can we give them to exercise it?

Online environments

Last year, Google made over 20 years of Usenet archives available and searchable. On one hand, they should be applauded for making public records accessible; on the other, the context of the digital 'public' in 1981 was very different than it is today. The advantage of a searchable database of answers to questions is obvious; but Usenet archives contain much more than that. With Google's searchable archives and tools like Microsoft's Netscan [14], it is quite easy to aggregate data about an individual over both time and contexts.

What is the effect of searchable aggregated data on an individual's perceived identity? Are persistent cross-contextual communication archives beneficial or harmful to individual participation and community development? What happens when this archived data is used to construct reputation scores [10]? What responsibility do designers have when creating representations of individuals through their data?

Without ample cues, understanding who a stranger is online is quite challenging. Yet, profiles are not a sufficient answer; they fail to convey enough information and what they do convey is often more problematic than no information [5]. How should an individual's identity be presented? How should individuals be able to articulate who they are in these digital environments with the level of depth that their physical presentation would allow? Given the persistence of data and the lack of location-based context, how should users be able to manage the different facets of their identity? What types of control should a user have over personal data and presence information?

Collaborative Calendar Systems

In current group calendaring systems, users do not have an understanding of the context in which they and their personal information participate in the calendaring system [15]. How can they assess their privacy needs and practices without adequate feedback from the environment? Who has access to their calendar? When did they access it? From where? What did they look at? How often do they view this information? What are the social norms for this environment? Even if they are allowed to assess the digital environment, how are they going to shape the environment to meet their privacy needs and practices?

Our Approach

We seek to take a wider view to the challenge ahead, inviting participants from different areas, including social scientists, technologists, designers, legal and policy experts. Lessig, a legal scholar, offers a framework for thinking about how privacy and behavior can be regulated: through market forces, through law, through architecture (including code), and through social norms [13]. This model affords a convenient and flexible means of framing current and future challenges in digital privacy regulation. It is also important to realize that in this model of regulation, factors do not operate independently; they are interdependent and affect each other. Thus, conversations between individuals working in all these different domains are fundamentally essential. Our workshop seeks to engage people across disciplines in conversation and collaboration, although we will most likely emphasize the architectural approach.

GOALS AND ACTIVITIES

Our goals are as follows:

- Develop a common vocabulary for addressing privacy in digital environments.
- Develop a common understanding of the expectations of users within the context of various usage scenarios.
- Establish a set of ethical guidelines for researchers and developers of digital environments.
- Identify promising approaches to supporting notice and consent in digital environments.

Proposed workshop structure:

8:30 –9:00	Orientation & Introductions
9:00–9:45	Keynote Address or General Discussion of Privacy (reflecting the perspectives of all participant positions)
10:00–11:00	Privacy Scenario Exercise (small groups)
11:00-12:00	Privacy Scenario Presentations & Discussion
12:00–1:30	Lunch
1:30-2:15	Collective Discussion on Ethics
2:15-3:30	Small group evaluation of scenarios from perspective of a specific challenge (e.g. FIPs, privacy management, etc.)
3:30-4:00	Break
4:00-5:00	Presentations & Discussion

The workshop will begin with a general orientation, because we expect participants from various disciplines. This will enable us to set a tone and structure for the workshop. Prior to the workshop, participants will have submitted position papers. These position papers will be provided to all participants before the workshop. Introductions will be brief, as participants are expected to have read all position statements prior to the workshop

The next hour will be a keynote address to the workshop. From there, we will break off into groups based on common problem areas/approaches. Groups will be asked to discuss common problems and issues within the context of their approach or area. Each group will then present their findings, and time will be allotted to general discussion.

After lunch, the workshop organizers will facilitate a collective discussion on ethics. Our objective is to reach a consensus on what the ethical guidelines should be for both researchers and developers working in this area. In particular we wish to focus on identifying rights, guarantees, and expectations of users.

We will then break into groups, each discussing a specific challenge relevant to the topic. Potential topics include:

- Compliance with Fair Information Practices
- Challenges of ubiquitous computing
- Promoting self-awareness, how to visualize or convey exposure, risk and history
- Privacy management techniques, helping users manage their digital privacy
- Assessing risk and exposure when faced with missing or untrusted information

The groups will then present and discuss their findings. Following those presentations, one topic will be selected as the basis for a more in-depth discussion.

The day will end with a collective effort to identify and explicate key findings of the workshop. We look forward to presenting these findings in a poster at the conference. We hope these findings will serve to inform other software developers, researchers, designers and policy makers.

PARTICIPANTS

We seek a balanced group, composed of social scientists, technologists, designers, legal and policy experts, and others with demonstrable interest or experience in privacy-aware or identity-management technologies in existing or emerging digital environments.

Participants will be selected based on position papers submitted prior to the workshop. Proposals should be no more than three pages in length, and should address the following:

1) Frame your area of work (problem area, target population, context of work), and list some of the constraints that you and your target population have to deal with.
2) What are the main privacy concerns of your target population?
3) What are the privacy issues that you are concerned with in your work?
4) Describe your approach to addressing the problems you have identified?
5) What are your measures of privacy, and/or exposures and risk?
6) Within your field, what do you consider to be the seminal works related to this issue?

We expect around fifteen participants, but are willing to accommodate up to twenty people should the quality of papers warrant expansion. Our main objective is to ensure both a good breadth as well as depth in terms of the represented disciplines and approaches. The workshop seeks to broaden people's horizons and provide an opportunity to discuss finer points of their work. We want the workshop to create connections across fields so more interdisciplinary work can take place.

ORGANIZERS

Currently, danah boyd is a graduate student with Dr. Judith Donath in the Sociable Media Group at MIT's Media Lab. Her work focuses on developing identity management tools and interactive personal visualizations to encourage users to reflect on their digital presence. Her previous work at Brown combined computer graphics, gender theory, and visual perception; she has also worked as a software engineer, an educator and an ethnographer. Ultimately, danah is interested in using technology to empower individuals. http://www.danah.org/

Carlos Jensen is a PhD student in Computer Science at the Georgia Institute of Technology. Working with Dr. Colin Potts, his work focuses on developing end-user privacy awareness and management tools for the web. He seeks to provide solutions that both make privacy management accessible to users, and work within the current technical framework. He has previously done work on online communities, media effects on communication, and online trust. http://www.cc.gatech.edu/~carlosj/

Scott Lederer is a PhD student in Computer Science at UC Berkeley, working with Drs. Jennifer Mankoff and Anind Dey. His current efforts are focused on illuminating a user conceptual model of privacy in ubiquitous computing, though his interests also extend to novel interaction techniques and devices. He aims to empower and elevate human experience in the ubiquitous computing age. http://www.cs.berkeley.edu/~lederer/

David Nguyen is a PhD student in Computer Science/HCI at the Georgia Institute of Technology. Working with Dr. Elizabeth Mynatt, David's research focuses on ubiquitous computing environments and privacy. He is working on ways to allow users to understand how they participate in these environments, so they can shape the environments to fit the their practices, needs, values, and sensibilities. Prior to Georgia Tech, David did his undergraduate work at UC San Diego in Cognitive Science and his Master's work at the University of Michigan in Computer Science. http://www.cc.gatech.edu/~dnguyen/

RESOURCES

Logistical requirements of the workshop include: one or two data projectors with screens, two or three large whiteboards, and wired or wireless Internet access.

REFERENCES

1. Adams, A. Multimedia information changes the whole privacy ballgame. *Proceedings of the Tenth Conference on Computers, Freedom and Privacy*, April 2000.
2. Antón, A.I., Earp, J.B., and Reese, A. Analyzing Web Site Privacy Requirements Using a Privacy Goal Taxonomy, To appear: *10th Anniversary IEEE Joint Requirements Engineering Conference (RE'02)* Essen, Germany, September 9-13, 2002.
3. Bellotti, V. Design for Privacy in Multimedia Computing and Communications Environments, In Agre, P., & Rotenberg, M. Eds. *Technology and Privacy: The New Landscape*. MIT Press, Cambridge MA, 1997.
4. Bellotti, V., and Sellen, A. Design for Privacy in Ubiquitous Computing Environments. *Proceedings of the 3rd European Conference on Computer Supported Cooperative Work, (ECSCW 93)*, G. de Michelis, C. Simone and K. Schmidt (Eds.), Kluwer, 1993, 77-92.
5. boyd, d. Sexing the Internet: Reflections on the role of identification in online communities, *Sexualities, medias and technologies: theorizing old and new practices*. University of Surrey, June 21-22, 2001. 7.
6. Dourish, P. Accounting for System Behaviour: Representation, Reflection and Resourceful Action, In Kyng and Mathiassen (Eds.), *Computers and Design in Context*. MIT Press, Cambridge MA, 1997. 145-170.
7. Dourish, P. Culture and Control in a Media Space. *Proceedings of the European Conference on Computer-Supported Cooperative Work*. ECSCW'93, Milano, Italy, September 1993, 125-137.
8. Federal Trade Commission. *Privacy Online: A Report to Congress*. June 1998. Available at http://www.ftc.gov/reports/privacy3/
9. Federal Trade Commission. *Privacy Online: Fair Information Practices in the Electronic Marketplace*. A Report to Congress, 2000.
10. Fiore, A., Teirnan, S.L, and Smith, M. Observed Behavior and Perceived Value of Authors in Usenet Newsgroups: Bridging the Gap. *Proceedings of SIGCHI 2002* (Minneapolis MN, April 2002).
11. Goldberg, Ian, et al. Privacy-enhancing Technologies for the Internet. *Proceedings of IEEE Spring COMPCON*, 1997.
12. Langheinrich, M. Privacy by Design - Principles of Privacy-Aware Ubiquitous Systems. *ACM Ubicomp*, Atlanta GA, 2001.
13. Lessig, Lawrence. *Code and Other Laws of Cyberspace*. Basic Books, New York, 1999.
14. Microsoft Research. Netscan. Available at http://netscan.research.microsoft.com/
15. Palen, L. Social, Individual & Technological Issues for Groupware Calendar Systems. *Proceedings of the ACM 1999 Conference on Human Factors in Computing Systems (CHI '99).*
16. Rosen, Jeffrey. April 14, 2002. New York Times. *Silicon Valley's Spy Game*. http://www.nytimes.com/2002/04/14/magazine/14TECHNO.html
17. The Internet Engineering Task Force. Geographic Location/Privacy (geopriv). Available at http://www.ietf.org/html.charters/geopriv-charter.html
18. Warren, S., and Brandeis, L. The Right to Privacy. *Harvard Law Review*, 1890.
19. Westin, Alan F., *1967. PRIVACY AND FREEDOM. New York:* Atheneum.