Revisiting Research Ethics in the Facebook Era: Challenges in Emerging CSCW Research

Amy Bruckman[1], Karrie Karahalios[2], Robert E. Kraut[3], Erika Shehan Poole[1], John C. Thomas[4], Sarita Yardi[1]

School of Interactive Computing Georgia Tech Atlanta, GA	Computer Science University of Illinois at Urbana-Champaign Urbana, IL	HCI Institute Carnegie Mellon University Pittsburgh, PA	IBM T.J. Watson Research Center Yorktown Heights, NY

{asb,yardi,erika@cc.gatech.edu}, kkarahal@cs.uiuc.edu, robert.kraut@cs.cmu.edu, jcthomas@us.ibm.com

ABSTRACT
New research paradigms in CSCW come with unforeseen ethical challenges. In particular, online social research tests the boundaries of public observation, third-party disclosure, and anonymization methods. Furthermore, there are differences in norms about what is and is not ethical among various research disciplines studying the Web. This workshop will bring together CSCW researchers who are interested in ethical challenges and best practices for studying the Web. From this workshop, we will foster the development of a network of researchers who will help shape university and corporate best practices for online research.

Author Keywords
Research ethics, regulation, methods, web research, social and legal issues

ACM Classification Keywords
K.4.1 Public Policy Issues; K.7.4 The Computing Profession.

General Terms:
Legal Aspects.

INTRODUCTION
As more people have access to the Internet, and more people are using the Internet in new and interesting ways, the opportunities for researchers to understand social behavior online has expanded immensely. With these new opportunities come new responsibilities. A range of open questions emerge: do researchers need 3rd party consent to view friend pages of a Facebook user who has consented to show the researchers her own profile page? Is it ethical to phish subjects to understand how to design better security protocols? When should researchers anonymize versus attribute creative content on the Web? If someone could potentially identify a subject even with anonymization, can the study still be conducted?

The goals of this workshop are to: 1) identify and debate emergent questions and tensions in online research; 2) develop and document techniques, best practices, and pitfalls in addressing these emergent questions; and 3) inform policy and practice among the broader research community through documentation for conducting ethical research online.

BACKGROUND
Human subjects research has been closely regulated in the United States after a number of egregious violations of research ethics came to light, such as the Tuskeegee syphilis study in 1974, in which study participants suffered the effects of syphilis untreated for decades after effective treatments were known. In response to this and other events, Health and Human Services put forth Title 45 CFR 46 of Federal Regulations in 1975 [7]; the Belmont Report [6] outlining principles for ethical human subjects research was issued in 1979. Other countries, notably the United Kingdom and Australia, also have similar regulatory oversight over research. However, as the contexts and nature of research has evolved, especially with the growth of the Internet, new ethical questions emerge that early regulatory mechanisms were not designed to address. As a result, there is much that these regulations and guidelines do not answer in this context. This workshop brings together researchers interested in discussing the challenges of how to conduct ethical online research.

MOTIVATION
Ethics in online research is not a new topic (e.g. [2]). However, the rapidly expanding scale and broadening scope of types of online research in cross-disciplinary communities demands attention within the CSCW community. While researchers and ethics review boards often rely on dichotomies like "public" versus "private," "published" vs. "unpublished," and "anonymous" vs.

"identified" [4], these categories are imprecise. The workshop will continue the discussion of themes uncovered at the CHI 2009 Special Interest Group on Research Ethics in the Facebook Era [1]. Over 50 researchers attended this SIG, from a range of research interests and with a wide variety of expertise. SIG participants expressed interest in attending a follow-up workshop hosted at CSCW 2010. This workshop targets a broad cross-section of CSCW attendees.

Topic 1: Web data collection and analysis
In offline public settings, activities such as observation of a public village square is, by most accounts, considered an ethically sound research approach. But in an online setting, boundaries between public and private are less clear. In what contexts should researchers seek consent from participants? When, and by what means, should people who use online services be notified that online research is occurring? When, if at all, should anonymization or masking of data be required? Does data that is obtainable online automatically count as "public"?

While Internet research is often no more risky than traditional research methods, the risks and safeguards against them differ, and change over time [5]. Terms of service, for example, vary across sites, and are governed by both legal policy (e.g. COPPA) as well as corporate policy. Further, online social networking research cannot be reasonably conducted, in many cases, if consent must be obtained from every potential data source. How should potential study participants be notified that their data may be used for research? What are the prerequisites for a weaker form of disclosure or 'opt-out' technique? Finally, current practices for anonymizing large datasets may not be adequate for many types of online research. Even if researchers take reasonable steps to anonymize large datasets, it is possible that one's identity can be revealed.

Topic 2: Disparate norms across research communities
Many universities, government research labs, and private organizations maintain committees (e.g. Institutional Review Boards) to oversee human subjects research conducted by their organizations. Yet, rules and regulations vary across regions and universities. As online research expands into a broad cross-section of disciplines, there is little agreement on how to conduct ethically sound research. In studies of social networks, for example, researchers in computer networking and data mining study sites like Twitter and Facebook, but are guided by different methodological approaches, norms, and conventions. Kinds of research that may be ethically acceptable in one field may not be in another. The role of site policies and terms of service are weighted differently among different academic communities and cultures.

Given the diversity of CSCW membership, which includes researchers worldwide from both industry and academia, what should its position be toward conducting ethical online research? How can we engage with other communities who are also interested in online research?

Topic 3: Pedagogical approaches and issues
Online research is increasingly being taught across disciplines and departments in higher education courses. Some courses are focused on network theory and analysis, while others are focused on the design and community aspects of the Web. Some include a mandatory ethics lecture and a class-wide IRB protocol under which students can then conduct individual projects [2,3]. However, as these types of courses grow, there is a need for consistent pedagogical approaches to teaching students ethics in the context of online research.

How should educators train students (including both future researchers as well as those who may not become researchers) to deal with ethical issues in online research? What should students be taught about when and how to capture, analyze, and report Web data?

ONGOING ACTIVITIES
We hope that this workshop facilitates future collaborative endeavors among participants, and creates a network of researchers who can share advice and best practices about online research. If appropriate, the results of this workshop will result in the publication of one or more papers discussing techniques, best practices, and pitfalls of Web research in a special issue of a journal or a magazine such as Interactions or CACM. We also intend to identify issues that can be brought to the attention of the SIGCHI Public Policy Committee and the ACM US Public Policy Committee (USACM).

REFERENCES
1. Bos, N., Karahalios, K., Musgrove-Chávez, M., Poole, E., Thomas, J. and Yardi, S., "Research ethics in the facebook era: privacy, anonymity, and oversight," *CHI 2009* ACM Press (2009), 2767-2770.
2. Bruckman, A., "Studying the amateur artist: A perspective on disguising data collected inhuman subjects research on the Internet," *Ethics and Inf. Tech. 4*, 3 (2002), 217-231.
3. Bruckman, A., "Teaching Students to Study Online Communities Ethically". *Journal of Information Ethics 15*, 2 (2006), 82-98.
4. Hudson, J. and Bruckman, A. Using empirical data to reason about internet research ethics. *Proc. ECSCW*, 2005.
5. Kraut, R., Olson, J., Banaji, M., Bruckman, A., Cohen, J. and Couper, M., "Psychological research online: Report of Board of Scientific Affairs' advisory group on the conduct of research on the internet," *American Psychologist 59*, 2 (2004), 105-117.
6. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. The Belmont report, U.S. Government Printing Office, Washington, DC, 1979.
7. U.S. Department of Health and Human services. Title 45 Public Welfare, Part 46 Protection of Human Subjects.