

Privacy & Confidentiality in Internet Research

Jeffrey M. Cohen, Ph.D.
Associate Dean,
Responsible Conduct of Research
Weill Medical College of Cornell University

IRB Issues

- Research on the Internet presents new concerns to the traditional IRB issues of privacy & confidentiality
- Privacy concerns relate to whether Internet activity
 - Is identifiable
 - Constitutes public or private behavior
- Confidentiality concerns relate to inappropriate disclosure of information obtained over the Internet

Privacy

- Identifiable vs. Anonymous
 - Online participants usually use pseudonyms (screen names, handles, etc.)
 - Although not publicly linked to actual names, identities can often be “readily ascertained” (e.g., using search engine)
 - People’s online identity may be as important to them as their actual identity

Privacy

- Public vs. Private Behavior
 - Most online activity is open to the public
 - Federal regulations base the definition of “private information” on the subjects’ “reasonable expectation” of privacy
 - In many situations (e.g., chat rooms), participants expect privacy and don’t expect their activity to be studied
 - Determination of privacy more complicated than it seems

Confidentiality

- Two potential sources of breach of confidentiality
 - inadvertent disclosure
 - Investigator who sent out research database to entire Listserv
 - Investigator who’s computer was stolen
 - deliberate attempts to gain access
 - No recorded incidents of hacking research data
- Technology can provide reasonable security but cannot guarantee absolute security

Confidentiality

- Data transmitted via e-mail cannot be anonymous without the use of additional steps. Almost all forms of e-mail contain the sender’s e-mail address.
 - use an “anonymizer” - a third party site that strips off the sender’s e-mail address
- Web servers automatically store a great deal of personal information about visitors to a web site and that information can be accessed by others.

Confidentiality

- Web sites can leave "Cookies", a small file left on the user's hard drive that is sent back to the web site each time the browser requests a page from that site. Cookies can record which computer the user is coming from, what software and hardware is being used, details of the links clicked on, and possibly even email addresses, if provided by the user.

CONFIDENTIAL

CONFIDENTIAL

Confidentiality

- Degree of concern over confidentiality depends on sensitivity of the information
- Since it is impossible to guarantee absolute data security over the Internet, some extremely sensitive research may not be appropriate for the Internet

CONFIDENTIAL

CONFIDENTIAL

IRB Requirements

- Investigators are going to have to provide technical information on how they will deal these issues.
- IRBs need to have sufficient expertise on the technical aspects of the Internet in order to ask the right questions and evaluate the information provided.
- IRBs that review Internet research without sufficient expertise are not in compliance with the regulations!

CONFIDENTIAL

CONFIDENTIAL

Resources

- American Psychological Association – Report of the Advisory Group on the Conduct of Research on the Internet
http://www.apa.org/journals/amp/featured_article/february_2004/amp592105.pdf
- AAAS Report on Internet Research
<http://www.aaas.org/spp/dspp/sfml/projects/intres/main.htm>

CONFIDENTIAL

CONFIDENTIAL

Contact Info

Jeffrey M. Cohen
Associate Dean,
Research Compliance
Weill Medical College of Cornell University
425 E. 61st. St. DV-301
New York, NY 10021
Phone: (212) 821-0612
Fax: (212) 821-0580
jmc2011@med.cornell.edu

CONFIDENTIAL

CONFIDENTIAL