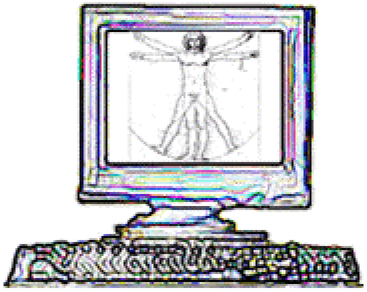
ETHICAL AND LEGAL ASPECTS OF HUMAN SUBJECTS RESEARCH ON THE INTERNET



A REPORT OF A WORKSHOP
June 10-11, 1999
Washington, DC

Mark S. Frankel, Ph.D. and Sanyin Siang
Scientific Freedom, Responsibility and Law Program
Directorate of Science and Policy Programs
American Association for the Advancement of Science
1200 New York Avenue, NW
Washington, DC 20005

November 1999

http://www.aaas.org/spp/dspp/sfrl/projects/intres/main.htm

This report was funded under a contract with the U.S. Office for Protection from Research Risks. The report's contents reflect the views of those participating in a workshop convened by the AAAS Program on Scientific Freedom, Responsibility and Law in June 1999. They do not necessarily represent the views of the American Association for the Advancement of Science or the Office for Protection from Research Risks.

INTRODUCTION

The Internet has become an important form of communication in modern society, with a forecast of 500 million online globally by the year 2003.[1] Its increased use and accessibility have led to a burgeoning of cyber communities, where people of like minds and common interests transcend geographical barriers and communicate with one another on a range of subjects, some trivial, some controversial, and some intensely private.[2]

The vast amount of social and behavioral information potentially available on the Internet has made it a prime target for researchers wishing to study the dynamics of human interactions and their consequences in this virtual medium. Researchers can potentially collect data from widely dispersed populations at relatively low cost and in less time than similar efforts in the physical world. As a result, there has been an increase in the number of Internet studies, ranging from surveys to naturalistic observation. Examples of recent research include the Carnegie Mellon Human-Computer Interaction Institute's investigation of the social and psychological effects of Internet use at home[3]and a University of Pittsburgh researcher's study on Internet addiction.[4]

New Internet research offers great potential for improving scholarship in a wide variety of fields and for assessing the very practical impacts of an increasingly critical technology. Indeed, this potential was recognized in the August 1998 report of the President's Information Technology Advisory Committee, when it recommended that the federal government expand its research portfolio on the "social and economic impacts of information technology diffusion and adoption."[5]

The ease with which the cyberspace medium facilitates these types of studies also raises issues about the ethical and legal dimensions of such research and the norms and policies that have traditionally governed its conduct. The ability of both researchers and their subjects to assume anonymous or pseudonymous identities online, the complexities of obtaining informed consent, the often exaggerated expectations, if not the illusion, of privacy in cyberspace, and the blurred distinction

[1] October 1999 International Data Corporation Survey on Internet Usage

[2] Schrum, L. "Framing the Debate: Ethical Research in the Information Age." *Qualitative Inquiry*. 1995 1(3):311-326.

[3] Kraut, R. Patterson, M., Lundmark, V., Kiesler, S, Mukophadhyay,T & Scherlis, W. "Internet Paradox: A Social Technology that Reduces Social Involvement and Psychological Well-Being?" *American Psychologist*. 1998. 53(9):1017-1031. http://homenet.andrew.cmu.edu/progress/research.html

[4] Young, K. *Caught in the Net: How to Recognize the Signs of Internet Addiction-And a Winning Strategy for Recovery*. John Wiley & Sons 1998.

[5] "Information Technology Research: Investing in Our Future." Report of the President's Information Technology Advisory Committee. 1999 http://www.ccic.gov/ac/report/

between public and private domains fuel questions about the interpretation and applicability of current policies governing the conduct of social and behavioral research involving human subjects.

The Office for Protection from Research Risks (OPRR), the agency responsible for oversight of federally funded research by the U.S. Department of Health and Human Services involving human subjects, has received inquiries from researchers and Institutional Review Boards (IRBs) members seeking guidance regarding research in this area. Many IRBs recognize their unfamiliarity with the nature of Internet research and their lack of technical expertise needed to review related research protocols. To both protect human subjects and promote innovative and scientifically sound research, it is important to consider the ethical, legal, and technical issues associated with this burgeoning area of research. Researchers, IRBs, and policy makers need to know the questions to ask as the first step in developing appropriate responses.

To contribute to that effort, AAAS and OPRR convened a workshop on “Ethical and Legal Aspects of Human Subjects Research in Cyberspace” in June 1999. The workshop was intended to explore the relevant issues and lay the groundwork for further involvement in these matters by professional and online communities, research institutions, and government agencies. Participants were drawn from OPRR and an array of fields, including social science, ethics, law, and computer science. Over the course of one-and-a-half days, they fleshed out the relevant issues in online research and considered the role of IRBs. This report and its action, research and education agenda are products of the workshop.

This is very much an exploratory study. We make no pretense of being comprehensive or definitive. There is a vast amount of research taking place on the Internet; we had neither the resources nor the time to catalogue and examine it systematically. One should not attempt to generalize from this effort to the larger body of research. That awaits additional study. Nevertheless, we believe this effort identifies a set of issues that provides a basis for fostering further analysis and dialogue among the various players. We invite others to join us in raising the consciousness of all those committed to advancing scientific research in a way that ensures appropriate protections for human subjects.

BASIC PRINCIPLES OF CONDUCTING HUMAN SUBJECTS RESEARCH

The current ethical and legal framework for protecting human subjects rests on the principles of autonomy, beneficence, and justice. The first principle, autonomy, requires that subjects be treated with respect as autonomous agents and affirms that those persons with diminished autonomy are entitled to special protection. In practice, this principle is reflected in the process of informed consent, in which the risks and benefits of the research are disclosed to the subject. The second principle, beneficence, involves maximizing possible benefits and good for the subject, while minimizing the amount of possible harm and risks resulting from the research. Since the fruits of knowledge can come at a cost to those participating in research, the last principle, justice, seeks a fair distribution of the burdens and benefits associated with research, so that certain individuals or groups do not bear disproportionate risks while others reap the benefits. This report is organized around these central principles.

BENEFITS AND RISKS

One of the fundamental principles of research ethics, beneficence, obligates researchers to maximize possible benefits from the research and minimize harms and risks to their subjects. Benefits can be defined as gain to society or science through contribution to the knowledge base, gain to the individual through improved well being, or empowerment of the individual by giving him or her a voice. Harms may include death and injury, psychological abuse, loss of privacy and public exposure and may not only affect individuals, but specific population subgroups as well. Over the years, guidelines and requirements such as informed consent and the protection of privacy and confidentiality have been developed and modified to reinforce this ethical principle in the physical world. As the Internet continues to offer researchers both a tool and a medium for research, there is a need to reexamine how the principle of beneficence and current guidelines and requirements translate into the virtual domain, and whether they provide an adequate foundation for protecting human subjects. Whether the benefits and risks in online research are less or more than what occurs in the physical world remains to be determined as research enters this new technological frontier. We raise the issues below simply to indicate the potential for risk in Internet studies that warrants assessment as this research proceeds.

No research involving human subjects should occur without some expectation of benefit, whether it be the advancement of science and new understanding, or a direct benefit to the participating subjects. Researchers' claims about the benefits of their research will rest in large part on their ability to collect useful data. But conducting research on the Internet raises questions about data sampling techniques and the validity and reliability of the data collected. For example, the Internet appeals to researchers because of its access to a potentially wide geographical and diverse population. However, this may also be one of the pitfalls in such research, since it is quite easy to mislead others about one's geographical location, gender, or race.[6] The reality may be that the research population is skewed in gender, race, and geographical distribution. Moreover, studies have revealed the existence of a racial and economic divide among Internet users,[7] further compounding the issue of non-representative sampling. As a result, the claims of benefit may suffer from a skewed data set that leads to misleading findings, and perhaps misguided policy if the data are relied upon by policy makers. Resolving these sampling issues is critical to the conduct of certain types of research on the Internet.

Much more so than in the physical world, virtual communities are very fluid, with new participants joining daily and others withdrawing and then perhaps returning at a later time. This feature of online communities complicates efforts to conduct debriefings and follow-up research in order to assess the long-term benefits or harms to subjects.

With respect to benefits, Internet research can contribute to the growing pool of knowledge on the new phenomenon of online communities and interactions. It allows the researcher to do so conveniently, and grants him or her potential access to a geographically and culturally diverse population. In some cases Internet research will provide greater convenience than research in the physical world for someone with online access to participate in the study.[8, 9] It has also been shown that interviews conducted via e-mail allow for greater clarification of concepts and involvement and empowerment of the participants than face to face interviews.[9] Furthermore, Internet research enables some individuals or populations, who might not be able to or willing to do so in the physical world, to participate in the research, hence giving some a voice that they would not otherwise have outside of online research.[10]

Subjects are less likely to experience physical injury in online research than in the physical world, but that should make us no less vigilant of research on the Internet. For example, one of the most

[6] Waskul, D and Douglass, M. "Considering the Electronic Participant: Some Polemical Observations on the Ethics of On-Line Research." *The Information Society*. 1996 12:129-139.

[7] "Falling Through the Net: Defining the Digital Divide." Report of the U.S. Department of Commerce National Telecommunications and Information Administration 1999.

[8] Hewson, CM, Laurent, D, and Vogel, CM. "Proper Methodologies for Psychological and Sociological Studies Conducted via the Internet." *Behavior Research Methods, Instruments, & Computers*. 1996 28(2): 186-191.

[9] Murray, CD and Sixsmith, J. "E-mail: a Qualitative Research Medium for Interviewing?" *International Journal of. Social Research Methodology*. 1998 1(2): 103-121.

[10] Bier, M., Gallo, M., *et al*. "Personal Empowerment in the study of Home Internet Use by Low-Income Families" http://www2.educ.ksu.edu/Projects/JRCE/v28-5/Bier/article/textonly.htm

common forms of Internet research is the survey. Traditionally, survey research has been thought to pose little risk to participants compared to other, more intrusive methods because participants possess greater control over the extent of their participation, and their identities are typically kept confidential. While survey research online shares many characteristics of traditional survey research, it may increase the subject's risk of identity exposure since subjects are transmitting their responses via the Internet and may not be aware of or sufficiently protected from the potential accessibility to their personal information by others.

This lack of understanding by participants, and sometimes researchers as well, of the technical and storage capabilities of Internet technologies may elevate the risk. The risk of exposure can surface at different stages of research, from data gathering, to data processing, to data storage, and dissemination. During data gathering, researchers conducting a sensitive study may not be aware that the participant is sharing an e-mail account or is not the owner of the computer that they are using to communicate. The researcher, unaware of the situation, may respond to a confidential e-mail and the receiver may be the owner of the computer with whom the participant shares an e-mail address. Furthermore, participants may not be sophisticated enough to know that there is a record of the exchange in a cache somewhere on their system or saved in their Internet service provider's server's log files. The possibility also exists that an e-mail may be sent to the wrong address, leading to potential embarrassment, or worse, for the participant. Furthermore, as data are accumulated and stored over the years, outdated or poorly designed security measures may create more opportunity for risky exposure.

During data dissemination and publication, the assessment of harms and benefits online becomes more complicated by pseudonymous identities. In Internet research, researchers may encounter the presence of pseudonyms in place of "real" identities. Many participants in online communities and other types of computer mediated communications (CMC) use one or more pseudonyms. Researchers are obligated by federal policies and professional ethics to provide special consideration for vulnerable members of the community, such as children and persons of diminished mental capacity. The use of pseudonyms leads to the possibility that vulnerable populations not normally recruited for a study could be included without the researcher's knowledge. Yet, researchers have traditionally disregarded pseudonyms as real identities and have quoted them directly along with the names of the newsgroups in their published research.[11] Yet, one workshop participant observed that

[11] King, SA. "Researching Internet Communities: Proposed Ethical Guidelines for the Reporting of Results." *The Information Society*. 1996 12:119-127.

online, people invest in their pseudonyms the way they invest in their real identities within a physical community. Hence, a researcher who attributes a quote or other information to an online identity and references the community studied could, within the confines of the online community, trigger reactions by community participants at specific individuals that may lead them to experience psychological distress.

Questions are also raised about how much a researcher should quote directly from online texts and whether her or she should give the name of researched community. In cyber fieldwork, researchers can have largely unprecedented access to people's conversations and stories. Studies have documented the tendency of people to become more open online than they are in person.[12, 13] Under a false or exaggerated expectation of privacy, participants may reveal more than what they might have done under conditions in the physical world. Furthermore, e-mail conversations can be archived without the participants' knowledge. Against this backdrop, direct reference to the researched community and public exposure may negatively affect and adversely impact the dynamics of an online community.[11.] In his paper on "Researching Internet Communities," King referred to the reaction of a member of an e-mail discussion support group who, after being cited without permission, felt that the "support group" is no longer a "safe environment" for discussion and for help.[11] This sense of violated privacy for the group as a whole is also illustrated in the aftermath of a 1994 study[14] of online self-help groups for sexually-abused survivors. These negative reactions are not inevitable, and they may well be rare. In fact, in the limited timeframe of this project, several studies were identified in which researchers went to great lengths to protect their subjects.[15] Nevertheless, problems can happen and vigilance in preventing or minimizing them is required.

INFORMED CONSENT

A vital component of the ethical discourse on human subjects research is the process of informed consent, which recognizes the autonomy of research subjects by sharing with them the power of

[12] Reid, E. "Informed Consent in the Study of On-line Communities: A reflection on the Effects of Computer-Mediated Social Research." *The Information Society.* 1996 12:169-174.

[13] Childress, CA and Asamen, JK. "The Emerging Relationship of Psychology and the Internet: Proposed Guidelines for Conducting Research." *Ethics and Behavio.r* 1998 8(1):19-35.

[14] Finn, J and Lavitt, M "Computer Based Self-Help Groups for Sexual Abuse Survivors." *Social Work with Groups.* 1994 17:21-46.

[15] See, for example, David Jacobson, "Impression Formation in Cyberspace: Online Expectations and Offline Experiences in Text-based Virtual Communities," in *Journal of Computer-Mediated Communication*, 1999 5(1) http://www.ascusc.org/jcmc/vol5/issue1/jacobson.html and Geoffrey Z. Liu,

decision making. Already complex in its application in the physical world, the process of informed consent can be further complicated by features of the Internet.

The issues that arise regarding the process can be captured in three questions: When is informed consent required; how can it be obtained; and how can it be validated? Three features of the Internet -- the blurred distinction between the private vs. public domain, its easy conductivity for anonymous and pseudonymous communications, and its global and easy accessibility -- pose difficulties for interpreting and implementing the requirements of informed consent.

In human subjects research, the distinction between public and private domains is important for determining when informed consent is required, since researchers may be exempt from obtaining consent for data collected from the public domain, such as data collected from television, public records, radio, printed books, conferences, or in public spaces such as parks. Data from online newsgroups and usenet support groups are readily accessible to anyone, and, if archived, accessible to the public months or years after messages were posted.

Some researchers interpret cyberspace to be part of the public domain since newsgroups, listservs, Internet Relay Chats (IRCs), and Multi-User Dungeons (MUDs) they observe are as accessible to anyone as a television or newspaper interview. These researchers believe that the responsibility falls on the disseminators of the messages to filter out what they might consider revealing or private information.[11] Hence, they adopt the position that this type of research should be exempt from the informed consent requirement.

Other scholars disagree with this interpretation, arguing that researchers have an ethical responsibility to understand how the diverse forums of the Internet work and how the users of these forums form expectations about what and where they are communicating. They see the greatest risk for cyberspace participants occurring in the situation where members remain unaware that their messages are being analyzed until the results of the research are published. Moreover, if the results are published in such a way that members of a virtual community can identify their community as the one studied without their knowledge, psychological harm may result.[11] These scholars argue that even though the information is public, communicants may perceive a degree of privacy in their communications. One workshop participant gave the

“Virtual Community Presence in Internet Relay Chatting,” in *Journal of Computer-Mediated Communication*, 1999 5(1) http://www.ascusc.org/jcmc/vol5/issue1/liu.html example of a substance abuse online support group. Although the e-mail list that hosted the group is publicly accessible, its members expected a degree of privacy comparable to that at an Alcoholic Anonymous meeting.

Workshop participants suggested that the same private vs. public distinction cannot be drawn for all computer mediated interactions, given the differences among the various types of CMC and the wide variations of groups that exist under each type.   One participant recommended that the evaluation criteria for the level of sensitivity that members of a particular online community may expect be proportional to the community's level of "accessibility."

If it is determined that informed consent is required for a particular research protocol, researchers and IRB members must next grapple with how to obtain it. The informed consent process involves three components: relating the information to subjects; ensuring that subjects comprehend the information; and obtaining the voluntary agreement from subjects to participate. Researchers are charged with the responsibility of determining what information should be conveyed to subjects in the consent process. In the physical world, this information may detail the possible risks from the research, such as the side effects of a particular drug for a clinical trial or the various avenues of potential exposure in a psychological study. The Internet is a new venue for research, and the technology is not always well understood, by scientists or their subjects. For example, should informed consent online include information on the technology component that is part of any transmission? If yes, what details should be provided?

The ease of anonymity and pseudonymity of Internet communications also poses logistical difficulties for implementing the informed consent process. As mentioned in the Benefits and Risks Section of this report, it is difficult for researchers to know with certainty relevant characteristics of their subjects, such as their age or mental competency, for determining types of risks. For example, minors could respond to a study involving inappropriate materials for their age without the researcher's knowledge. Furthermore, as in the physical world, some populations, due to gender, geographical location, or race, may be more susceptible to certain types of risks.[16] Such uncertainties compound the difficulties of determining what types of information should be imparted to subjects that would aid them in deciding whether to participate in the research.

Anonymity and pseudonymity not only complicate efforts to determine what information should be conveyed to subjects, they also affect researchers' ability to gauge the subject's understanding of the research risks. In the physical world, the researcher and subject can engage in face-to-face dialogue that can help to ascertain whether the subject adequately comprehends. Such dialogue is not a characteristic of the Internet, and the distance between researcher and subject is widened even further when the researched community consists of hundreds or more anonymous or pseudonymous members. Are there models in the physical world, perhaps in naturalistic research protocols, in which informed consent from a "community representative" could suffice for research on the entire community? And how should researchers proceed when studies are based on interactions among community members and some members refuse to give informed consent?[17]

Just as research subjects can be cloaked in anonymity and pseudonymity, so can researchers, raising the issue of deception. Deception occurs when a researcher intentionally misinforms or does not fully disclose relevant information to subjects in cases when informed consent is required. On the Internet, group discussion formats make it relatively easy for researchers to engage in covert or unobtrusive observation. An investigator can record the online conversations of a community without making her presence as a researcher known. Alternatively, she can pose as a member of the community, giving false information in order to study the reactions and behavior of community members. Guidelines in the physical world allow for deception in the study of human phenomena, providing that the research has considerable prospective scientific, educational, or applied value, that there are no alternative methods for achieving the expected results, that the risks to subjects are minimal, and that sufficient explanation or a debriefing will be given to participants as soon as possible following the conclusion of the research.

[16] Mark S. Frankel, *Viewing Science and Technology Through a Multicultural Prism*. AAAS 1993.

[17] The complexities associated with obtaining informed consent when researching communities has been widely debated in both the social and biomedical sciences literature. See, for example, Ruth Macklin, "The Problem of Adequate Disclosure in Social Science Research," in Tom L. Beauchamp, Ruth R. Faden, R. Jay Wallace, Jr., and LeRoy Walters (eds.), *Ethical Issues in Social Science Research*, The Johns Hopkins University Press, 1982, pp.209-210, and "Proposed Model Ethical Protocol for Collecting DNA Samples," *Houston Law Review*, 1997 33(5):1443-1447.

Without a clearer understanding of the benefits and risks associated with Internet research, it may be difficult to justify deceptive practices online. Even if deceptive practices by researchers were permitted, how is the debriefing to be conducted so that it reaches the entire virtual community? Online communities, like their physical counterparts, may be in constant flux. However, tracing members who leave the virtual community prior to the debriefing may be more difficult than the counterpart situation in the physical world. Users of one e-mail address may easily close that address and switch to another without leaving a trail as to their new address, or chatrooms and listservs may delete the e-mail addresses of members who leave after a certain period. Additionally, the "faceless" nature of online interactions may further complicate the debriefing process. In the physical world, there are groups whose members remain anonymous and in which members may also be in flux. However, should such a group consent to study by a nonobstrusive researcher, its members are able to observe the researcher taking field notes or tape recording a session.[6] Likewise, the researcher can "see" who is leaving the group and who is entering, and inform them of the research upon departure from or arrival into the group. The faceless nature of online interactions may not always allow for this course of action.

Finally, there is the nature of the consent form and the validity of the process. In the physical world, informed consent is secured with a written signature on a consent form (telephone surveys, however, may simply rely on a verbal consent). Online, the equivalent would be a click to a statement such as "I agree to the above consent form." But how valid is such consent when the age, competency, or comprehension of the potential subject is unknown? A key issue to resolve is how informed consent can be authenticated online.

PRIVACY AND CONFIDENTIALITY

Harm to human subjects can occur with the invasion of subjects' privacy and the violation of confidentiality. Invasions of privacy happen when research participants lose control of the types of personal information revealed about themselves. Privacy provides people with some protection against harmful or unpleasant experiences – against punishment and exploitation by others, against embarrassment or lowered self-esteem, against threats to the integrity and autonomy of the individual. Invasions of privacy can increase the likelihood of harm because they deprive the individual of that protection.[18] Violations of confidentiality occur when

[18] Kelman, HC. "Privacy and Research with Human Beings." *Journal of Social Issues* 1977 33(3):169.

information about a research participant is disseminated to audiences for whom it was not intended without the subject's authorization.

As noted earlier, some commentators contend that cyberspace postings are open to public scrutiny, and thus do not fall under the protection of confidentiality. Others maintain that there is a difference between what is publicly accessible and publicly distributed. For example, an online support group may be publicly open to anyone who wishes to participate in its discussions, but its members may perceive the exchange of information as a very private matter. Hence, in order to answer the question of when a researcher is obligated to take steps to protect the privacy and confidentiality of his/her subjects, one must first delineate the boundaries of the private domain in cyberspace.

One workshop participant approached the issue by examining it from two perspectives. The technological point of view perceives the Internet as an exchange of data files as products of online activity, such as the archives of online discussion groups. It is seen as a form of electronic book that has been written collaboratively. Categorization of these archives as "public books" or "private journals" depends on the accessibility of the data files. The other perspective involves the psychological construct of cyberspace, and calls for a distinction between the public and private domains based not on the accessibility of the data, but on the psychological perception of the subjects with regard to the information. The technological point of view may define a set of data as publicly accessible, electronically written books, but the providers of the data may perceive these as electronically captured records of private conversation occurring in a quasi-spacial domain.   This workshop participant favors an approach that first develops a technological understanding of the issue and then expands this understanding to include the psychological perspective of the participants.

In a paper on proposed ethical guidelines for the reporting of results in Internet research, Storm King adds further detail to the concept of accessibility vs. perceived privacy.[11] Similar to the previously described point of view, he believes that an online group can be evaluated on several dimensions in order to determine if the requirements of protecting privacy and confidentiality are necessary or justified. First, Internet communities can be classified according to the degree of group accessibility, a factor that represents how accessible a particular Internet forum or community is to the public. On one extreme are unmoderated Usenet bulletin board groups and on the other are private e-mail groups with unpublished subscription addresses and enforced requirements for participation. In the middle are MUDs (Multi-User Dungeons), where the address is available to the public, but participants are constrained by internally prescribed or available activities that are not considered public. Second, Internet communities can be classified according to varying degrees of perceived privacy by their members. At the extreme of low perceived privacy are scholarly e-mail discussions, where the aim of participation is to promote the widest possible dissemination of ideas. At the extreme of high perceived privacy are support groups for very socially sensitive topics, such as certain medical or mental conditions that could create a social stigma. Regardless of the level of group accessibility, many members of high perceived privacy groups post messages with the expectation that only others that understand, respect, and support their situation will read their notes. The question remains, how should the public and private domains be defined for research in cyberspace?

Another feature of online research that complicates the application of existing guidelines and policies regarding privacy and confidentiality are anonymous and pseudonymous communications. Government regulations[19] and ethical guidelines[20] obligate researchers to protect the privacy of research subjects**.** For research in the physical world, researchers can comply with this obligation by disguising the identity of their subjects. However, the anonymity of online interactions is fundamentally different from that of traditional practices. Many online users have pseudonyms that in themselves might qualify as sufficient disguise in the physical world, but which have comparable values to real identities in the online world.

Cyberspace interactions possess the timeliness of spoken conversations and the endurance of printed matter. Some scholars treat communication on the Internet like a spoken conversation and argue that as long as the real identity of the participants is protected, it would be ethically possible to cite fragments of electronic messages from virtually any source.[21] Other scholars find this interpretation problematic because communication on the Internet is typed rather than spoken, leaving a physical record that can be archived or otherwise preserved. In the cyberspace domain, it is much more difficult for a researcher to ensure subject privacy, since a determined reader of a study who knows the name of the studied group could trace the message and discover the login name of the person who sent the message. Supporters of this position cite a 1994

[19] See "Protection of Human Subjects" 45 CFR 46.

[20] For example, see American Psychological Association. "Ethical Principles of Psychologists and Code of Conduct." *American Psychologist*. 1992 47(12):1597-1611.

[21] Herring, S. "Linguistic and Critical Analysis of Computer-Mediated Communication: Some Ethical and Scholarly Considerations." *The Information Society* 1996. 12:153-168.

study[14] in which researchers collected and analyzed texts from a cyberspace support group for sexually-abused survivors. Although the published article altered the names of participants in the discussion group, many of the participants could clearly identify themselves since it quoted directly from the text of their cyber postings, and they felt violated. Moreover, one workshop participant observed from her experience in working with online communities that people invest just as much into their online identities as they do in their real ones. Hence, it is may not be enough to protect the real identities. It may be necessary to protect the online identities of the research subject as well. How, then, should a researcher cite online text without violating the privacy and confidentiality of his or her subjects? How much description of an online community should a researcher provide?

A lack of understanding among researchers and potential subjects regarding the technical components/limits of the Internet may further complicate issues of privacy and confidentiality. With respect to technology, confidentiality can be compromised during data transmission and storage. Unauthorized persons may be able to access messages transmitted through the system as multiple copies of messages are transferred from computer network to computer network. E-mail is also vulnerable in transmission as a result of computer or human error. E-mail communication may sometimes be re-routed to unanticipated locations, perhaps with minimal security systems, due to technical malfunctions within the computer network. There are also occasions when people mistakenly sends e-mail to the wrong address, or when they wish to respond privately to a message posted on a listserv, but end up sending a private response to the entire listserv group.

Furthermore, more than one person may have access to an e-mail account. Persons within the sender's family may have access to e-mail that is sent from the home computer. As an example, a research participant sends e-mail to the investigator regarding a sensitive personal issue and the participant's spouse or other family member may access and view a saved copy of the sent e-mail. Conversely, the investigator sends a response that references this sensitive issue, and a spouse or other family member may open and view the investigator's response, thereby inadvertently breaching confidentiality. And an employer may gain access to a participant's e-mail sent from an office computer. Many office computer systems make routine copies, separate from the sender's copy, of e-mail sent from office computers, and employers may have a legal right to monitor and read employee's e-mail.[22]

With respect to data storage, privacy and confidentiality may be inadvertently breached when the researcher stores the data on a computer with Internet access and unauthorized persons hack into the system. Even if material is erased from a computer, this only involves making the disk space formally occupied by the erased material available to be written over by new files. Until such time as this disk space is reissued, the erased material may still be recovered.

Researchers, subjects weighing whether to participate in a study, and IRB members reviewing research protocols need to be aware of the potential technological breaches of privacy and confidentiality in order to minimize risks to subjects in the course of Internet research.

JUSTICE

Of the basic principles governing human subjects research, justice is perhaps the most elusive in terms of application and understanding. Justice can be interpreted as "fair, equitable, and appropriate treatment in light of what is due or owed to persons." A person has a valid claim based in justice when he or she is owed something. An injustice occurs when certain persons are denied benefits that are rightfully due to them or when burdens are not distributed fairly.[23] With respect to human subjects research, application of the principle of justice is inextricably linked to fair distribution of the burdens and rewards of research.   This bears directly on the selection and recruitment of participants, where justice is invoked to ensure that subjects are selected for reasons directly related to the problem being studied instead of for their easy availability, compromised position, or tractability.

Applying the principle of justice to Internet research is based in part on identifying the benefits and risks of the research and assessing how they are distributed. Yet, as noted earlier, the Internet poses several challenges in attempting to identify and measure benefits and risks. More work is needed on defining what constitute benefits and risks in Internet research.

[22] Doyle, R. "Privacy in the Workplace." *Scientific American* 1999.

[23] Beauchamp, TL and Childress, JF. *Principles of Biomedical Ethics 4th Edition*. Oxford University Press 1994.

The feature of anonymous and pseudonymous communications further complicates the application of justice with respect to human subjects research on the Internet. Such communications make it difficult to distribute the rewards of research when subjects are anonymous or when a researched community is in constant flux with the identities and numbers of its members unknown. Furthermore, anonymity and pseudonymity in cyberspace raise the issue of how researchers can be properly inclusive in their research designs. In the physical world, researchers need to consider factors such as gender, race, and age in the selection of their subjects to account for differences between genders and racial and age groups, where they exist, to maximize the generalizability of their results, and to ensure compliance with federal guidelines. This may be more difficult to achieve in the Internet environment, where people go to great lengths and take considerable pride in protecting their anonymity. And given the present social and economic disparity, both domestically and internationally, in Internet access and usage, achieving a fair distribution of burden and rewards may prove elusive in Internet research, at least for the immediate future.

CONCLUSION

Internet research raises a number of complex issues for the scientific community, research subjects, and policy makers. The preliminary discussions begun at the AAAS-OPRR workshop need to continue, with greater focus on the extent to which the issues take on different qualities on the Internet than they do in the physical world, and what that means for policies and guidelines intended to protect human subjects. In order for these discussions to be most useful, they must be firmly grounded in an understanding of Internet technology and its impact on human subjects research. The Internet as an evolving medium for research can be viewed as a paradigm of the larger question of how we deal with value-laden questions associated with technological advances and the capabilities they confer. This powerful technology needs to be better understood if we are to ensure that effective mechanisms are in place to protect human subjects.

Workshop participants developed a set of recommendations related to research and education as well as for action that should be considered as part of a general strategy for dealing with the challenges posed by Internet research. They are presented here as a basis for galvanizing further dialogue on these important issues.

RESEARCH AND EDUCATION AGENDA

In order to assess what, if any, changes in education and policies are needed to deal specifically with online research involving human subjects, various issues need to be examined and clarified, and research undertaken to determine the adequacy of existing educational efforts, guidelines, and policies.

- Clearly delineate the types of online research that would require compliance with federal guidelines on human subjects studies. Traditionally, a human subject is defined as "a living individual about whom an investigator (whether professional or student) conducting research obtains data through intervention or interaction with the individual, or identifiable private information." Given certain features of the Internet, this definition may need to be reassessed: the blurred distinction between public and private domains raises the question of what is considered "private information" in cyberspace; and the traceability of online communications and the amount of emotional investment that some people put into their online identities challenge notions of what is considered "identifiable." Clarity on these matters is critical if promising research is not to be unnecessarily deterred.

- Assess the risks and benefits associated with different research methods used in online research, ranging from surveys, in which questions are posed to participants, to observational research, in which participants remain unaware of the researcher's presence. More work is needed to conceptualize and measure benefits and risks in Internet research.

- Improve understanding of the vulnerabilities of research subjects with respect to Internet research. The enhanced capacity for involving international participants in online research makes it imperative that researchers be sensitive to cultural or political factors that affect the vulnerability of those participants.

- Increase knowledge about the structure of Internet communities and their similarities and differences with physical communities.

- Delineate the boundaries of private vs. public space on the Internet.

- Survey existing literature and practices regarding the conduct of Internet research in order to develop a taxonomy of ethical and legal issues.

- Identify aspects of existing ethical guidelines and policies that can be applied to Internet research.

- Examine how IRBs currently handle ethical issues in Internet research.

- Assess existing methods and resources for encouraging responsible online research.

- Examine existing curricula on the ethics of human subjects research and develop strategies for including a component on Internet research.

- Develop case studies that illustrate the ethical and legal issues associated with Internet research.

ACTION AGENDA

Research and education should be joined by new initiatives that are intended to improve the quality of Internet research while promoting adherence to sound ethical research practices.

- In their proposals, researchers should be specific about the possible benefits and harms to their subjects, how they plan to minimize risk exposure, and their methods of securing informed consent from prospective subjects.

- Researchers should employ the concept of community consultation in planning research and interpreting results. This should include a dialogue with the researched communities regarding perceived benefits and harms, their expectations of privacy in different CMC environments (e.g., listservs,, MUDs, IRCs), and the information prospective subjects believe they should know to make decisions about research participation.

- Researchers should consult with their institution's technology system administrators regarding the technical aspects of their research so that they are knowledgeable about the power and limits of this research medium.

- Discussions of the ethical, legal and policy issues associated with Internet research should be broad-based and include international representation.

- Since many Internet users invest in the development of their online personas, there should be consideration of whether these pseudonyms should be treated as real identities and hence, afforded the same types of confidentiality protection.

- As we are only beginning to grasp the complexity of online research involving human subjects, IRBs should carefully evaluate requests for exemption of Internet research protocols.

- IRBs should consider having members of the virtual communities studied represented in their deliberations.

- IRB members should be familiar with the various methodologies associated with Internet research and assured that procedures used by researchers will safeguard participants. They can be aided in this task by including persons knowledgeable about online technologies in their deliberations.

- Professional societies should be encouraged to develop ethical guidelines and educate researchers on Internet ethics.

- Special certification for online research should be considered.

- Companies developing Internet technology should consider ways to design their products to help researchers conduct scientifically and ethically sound research.

- A national resource network of Internet researchers, ethicists, and technical personnel should be created to respond to inquiries from IRB members, researchers and subjects regarding the technical and ethical components of online human subjects research protocols. Inquiries and responses should be posted online as part of a publicly accessible FAQ bulletin board.

OPRR should identify ways to transfer information about Internet research to IRBs and researchers:

- Set up a Web site for researchers with links to examples of informed consent forms for Internet research.

- Develop a set of "points to consider" for researchers and IRBs that will help alert them to the ethical and legal requirements of human subjects research online when developing and reviewing proposals.

- Incorporate education about Internet research into national, regional, and local workshops.