

Ethics in Global Internet Research

Yvonne Wærn
Department of Communication Studies
Linköping University
SE 581 83 Linköping,
Sweden
e-mail: yvowa@tema.liu.se

Abstract

Two issues of research ethics are presented here: the need for research and the need for respecting human dignity. It was found that current research does not meet the need for research in global Internet use. English speaking Internet users are still over-represented in most Internet research. This leads to a biased view of Internet use. As to the need for human dignity, several documents were analyzed. Only one dealt directly with Internet research. The main requests for respecting human dignity were found to be: informed consent and anonymity. However, as regards Internet research, there is still some discussion about when data on the Internet may be considered to be public versus private. The recommendation is easy: Investigate to a greater extent how ethical issues are considered in different cultures.

Introduction

Is there anything special about Internet that requires us to take particular considerations as to ethics in global Internet research? Of course, that depends upon what we mean by "ethics" and "Internet research". I will present some arguments that we have to take particular consideration, if within ethics we include both demands on selecting a proper research topic and demands for respecting human dignity.

Basics of ethics

Ethics can of course have a lot of different meanings, ranging from philosophical considerations to practical issues. Let us consider some basics that underlie the consideration of ethics in research. We may start with utilitarianism, conceived in the 19th century by Jeremy Bentham and John Stuart Mill. According to their ideas ethical actions are those that provide the greatest balance of good over evil. For whom, we may ask? The utilitarians answer: The greatest good for the greatest number. As a critique against this stance we can say that also the minority should have the right to have some good. People have different, often conflicting, goals. Therefore,

we have to solve the issue of balancing goals among people in order to be able to act ethically.

This leads us to the rights approach. The rights approach has its roots in the philosophy of the 18th-century thinker Immanuel Kant and other, similar thinkers. Here the individual's right to choose for her/himself is stressed. For research ethics the following rights are central:

- The right to the truth: to be informed about matters that significantly affect choices to be made.
- The right of privacy: We have the right to do, believe and say whatever we choose in our personal lives as long as we do not violate the rights of others.
- The right not to be injured: We have the right not to be harmed or injured.
- The right to what is agreed: We have a right to what has been promised by those with whom we have freely entered into a contract or agreement.

Another approach is the common-good approach that assumes that community members are bound by the pursuit of common values and goals. This notion originated more than 2,000 years ago in the writings of Plato, Aristotle and Cicero. The common-good approach challenges us to view ourselves as members of the same community, reflecting on broad questions concerning the kind of society we want to become and how we are to achieve that society. We could say that in our Western society, there is a common agreement that research may lead to some common-good.

Research ethics

To a great extent, research ethics reflect the three approaches mentioned above. The utilitarian approach is reflected in the conception that research is good when it produces good results. If we combine this with the common-good approach, we may hold that only that research is good which leads to some common-good. However, taking the rights approach, we also see that there are some balances to strike with respect to research. Researchers themselves may overestimate the value of their studies to the common-good, and not adequately respect the rights of the people studied. Although the research may be seen to lead to some results that benefit society in the long term, this is not always clear to the people studied.

Most scientific boards have been concerned with research ethics. My aim with this paper is to present some of the general issues derived from various documents related to research ethics.

Method

The method used is information search on the web. Since this is a rather new research method, I will present here some of the basics in my use of it. Firstly, some reasons for using a web-based method. Although many people argue that "everything is to be found on the web", this is of course not

true. However, most of the newest information is available on the web. This is not the case in library databases that suffer from constant updating problems. Further, a search on the web has the advantage that I, as a researcher, get closer to the information sought, than when I have to work through a documentalist. However, as with all information searches, it is important to accept that all data will not be covered, and that there are limitations that have to be known by the user of the data. I list those I have considered below.

First, it is important to know the search engine used. Search engines give rather different results and coverage. Thus, in order to be able to judge the information gathered, the reader has to know which search engine(s) have been used.

Second, the search words are important. This is true because search engines work differently with search words. Also, the coding of the documents to be searched differs and thus the words used are crucial.

Third, the date for access has to be recorded and reported. The reason for this is that web-sites disappear or change servers rather often. The reader should have a chance to find the site. In case the server has been changed, the author (in this case I) should make all efforts to find the new site. In cases where readers later find that servers have been changed, I would appreciate being informed about this.

Fourth, all results from the search cannot be presented. Only those that have bearing upon the issue investigated will be included. In this particular article, I have not recorded material excluded. In other kinds of reviews of the web, it may be crucial to record all material found.

Documents used for illuminating research ethics

The search was performed with the search engine "Google", with the search words "research ethics". The search was performed several times, from January 1999 to May 2001. Most research ethics found on the web pertains to medical issues, these have been excluded. Also, much related to ethics concerns professional ethics, not ethics concerning investigations. The sites I will use below fall into the domain of social science research and have been stable over the period investigated. Among those found, I have chosen sites that are based on some kind of authoritative source. The presentations and conclusions are based on the following documents: (HSFR Etik, 1996)¹, Guidelines for Research Ethics NESH (1999), Ethical guidelines from the American Psychological Association (1992), Ethical guidelines from Calgary, National Scientific Research Council, Canada, (Accessed 2001-05-25). Most work has of course been done regarding "traditional research", that is research based on interviews, questionnaires

¹ The Swedish guidelines (HSFR) are at present not available on the net, since the organization of Swedish research boards has been changed. However, it is available (for closed use) to researchers related to the Ethics discussion at the Association for Internet Researchers.

or observations of people "in real life". I have been able to locate one single document that is concerned with ethics for Internet research, i.e. Frankel & Siang, (1999).

Results

The results from the search are given below. First, the general ethical principles found are presented then details are given concerning these principles.

Ethical principles

All guidelines acknowledge the need for research. This need includes the demand that the research is directed towards relevant issues and that it sustains a high quality. This implies that available methods and knowledge should be ameliorated and further developed. Furthermore, most space in the documents is dedicated to issues related to respecting the dignity of the human subjects included in the research. This demand implies that nobody should be hurt, psychically or physically by the research. None of these demands are absolute. They must be balanced against each other and against other demands. In each scientific study a trade-off must be considered between the value of the expected increase in knowledge and probable risks in terms of negative consequences for the participants in the research. I will approach these two demands, one at a time.

Need for research

Let me start by quoting from the workshop reported by Frankel & Siang, (1999):

"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. " Op.cit. Frankel & Siang, (1999), p. 3

I would like to pursue the issue of data sampling to some extent. Current research on the Internet is mainly concerned with Internet use among English-speaking people, or the research is mainly presented in the English language. I propose that this may give a biased idea of what Internet use is about.

First some background to the claim about the language/culture bias. In a recent investigation, the following figures as to the use of the Internet have been presented.

(Search in Google with search words: Statistics and Internet and language. Accessed Feb 7, 2001.) Source:

<http://www.gltreach.com/globstats/index.php3>

A graphical overview over the percentages is found in figure 1.

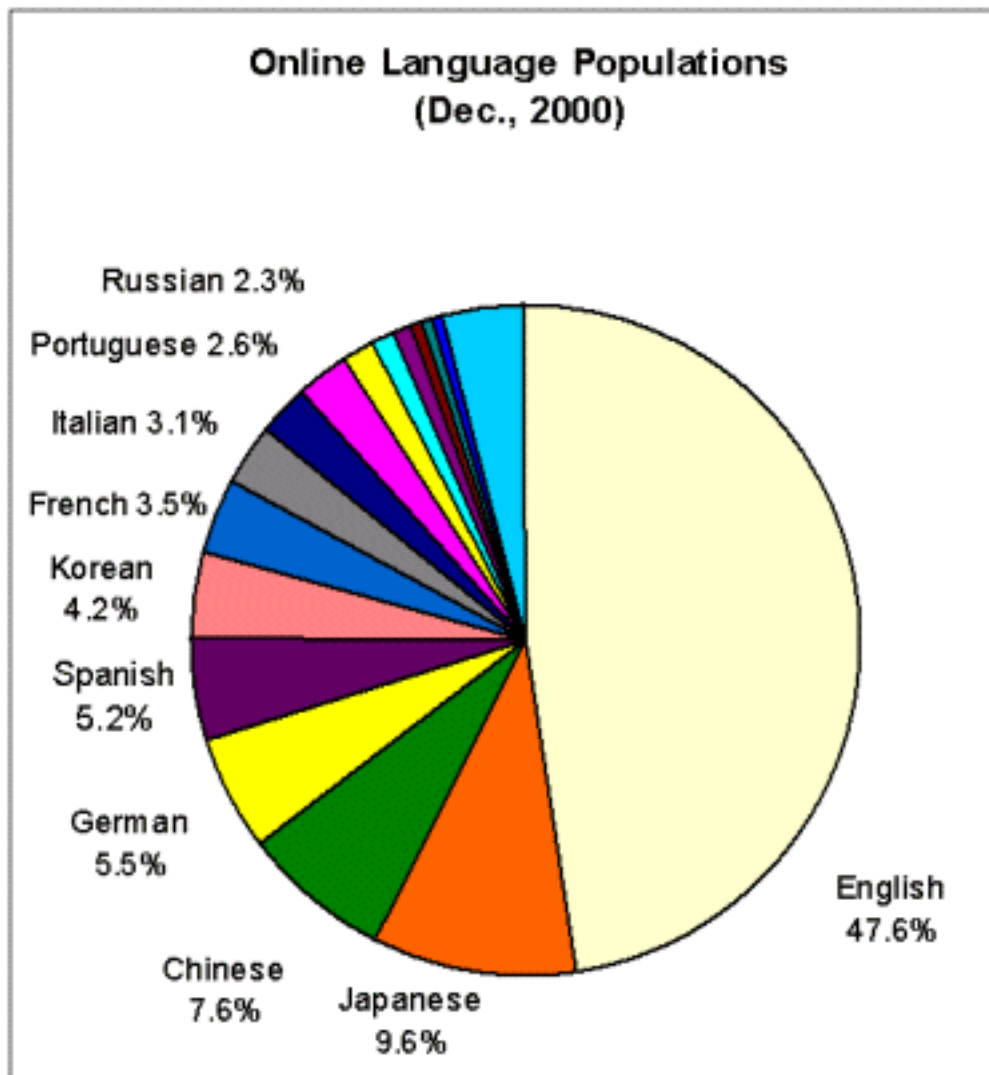


Figure. 1. Overview of language users online.

Source: <http://www.gireach.com/globstats/index.php3>

We see that the percentage of people online far from matches the percentage of the world's population. In some cases, there is a much higher percentage of online persons than offline persons (English-speaking), in other cases a much lower percentage online than offline (the Asian Languages). Now, let me compare these statistics to some recent web-sites, offering information about research on Internet use and communication. I searched with the search engine "Google" using the search words "research and communication and computers". Languages: German and Spanish. Results from February 6th, 2001. I chose this search approach because I wanted to get as much as possible (for me understandable) material that was not in English. One relevant bibliography is found at the following site:
<http://www.rzuser.uniheidelberg.de/~mbeisswe/biblio.html>
 This one presents a bibliography over research on chat-use.

Another relevant source is found at the following site: <http://www.netzwissenschaft.de/>. From this one, I chose the presentation of publications from 2000, dealing with all kinds of Internet issues. I now counted the occurrences of texts in English, German and other languages in these two sites. The statistics follow below.

Table 1. Statistics from two German bibliographies related to Internet research

Topic/Site	English	German	Other
Chat:			
http://www.rzuser.uni-heidelberg.de/~mbeisswe/biblio.html	111 63%	64 34%	11
Books, 2000			
http://www.netzwissenschaft.de/	199 65%	104 34%	1

It is obvious that the percentage of German-related publications is far greater than the use of the Internet shown above. The percentage of languages other than English or German is very low, indeed. The two bibliographies, although concerned with very different topics, show an astonishing similarity in terms of the distribution over languages, only that the "chat" bibliography includes a higher percentage of languages other than English and German. When I asked one of the originators, he told me that he listed books, the titles of which he would be able to understand, - which is of course a reasonable strategy. So, we find that reports of research related to computer-mediated communication are biased towards English or German reports. Does this matter? Will the coverage of issues be different from when other language communities are approached?

I have no firm answer, but I have started investigating this issue. I only want to give one example, from a list originating in Latin America, and called "Mistica" (address: www.funredes.org/mistica/). (Accessed several times, from Jan 29 to March 20th 2001). A continuation of this project can be found at: www.funredes.org/olistica/. One of the goals of this list is the following: "design and experiment collectively **new democratic paths** to bring original proposals of what should be **the Information Society in their region.**" (stress in the original text). (Mistica, May 2000).

We may find this goal in English-speaking lists as well, however current research does not reflect this to any great extent. Another aspect of covering other kinds of language cultures is that we will meet quite different circumstances. Mistica also works with the building of new communication

methods. One of the questions posed is the following: "How to integrate the radio with the Internet to serve communities without telephones?" This question seems rather far from the recent efforts to create a need for "broad-band" in Sweden.

Conclusion as to the relevance of research

I have tried to show that the research on Internet communication is far from relevant to the actual practice of Internet communication. I based the argument on the existence of different language cultures on the Internet, whereas the research is greatly biased towards English- and German-written communication. There is of course a danger in this kind of bias if we are going to talk about general Internet usage. Since the English and German writing participants probably to a greater extent belong to a highly developed, Western culture, the issues addressed will be quite different from issues addressed in other cultures. In particular, our issues seem to be biased towards using high-technology media and addressing "luxury" use of Internet communication. If our research is to be considered as of great value, we should also consider the use of the Internet by other, less privileged users.

Respecting Human Dignity

In the guidelines from the Natural Sciences and Engineering Research Council of Canada, the following is written about human dignity:

"Respecting human dignity implies protecting the multiple and interdependent interests of the person - from bodily to psychological to cultural integrity".

This means that we should not try to persuade people to participate in our studies, if they hesitate in doing so. Also, that the study should be performed in such a way that the cultural integrity should be protected - something that may be difficult in an Internet context, as shown above. Within this general concept of human dignity we find in all documents the following issues to be paid respect: Free and informed consent. This is an important aspect of respecting human dignity. I shall continue investigating this issue further below. Vulnerable persons should be guarded against abuse and exploitation. Such persons may be for instance children, but also people who are dependent upon us as researchers: i.e. patients, students.

Elaborated considerations for human dignity are the following:

Privacy and confidentiality. " In many cultures, privacy and confidentiality are considered fundamental to human dignity " (Canadian ethics rules). Thus, research should protect the access, control and dissemination of personal information.

Respect for justice and inclusiveness. This respect implies on the one hand that no one particular group should be burdened by participating in onerous

research. On the other hand, it implies that no groups should be excluded that might benefit from the research performed. This issue was discussed at some length above.

Balancing harms and benefits. It is self-evident that research should provide more benefits than doing any harm, in general. It may be difficult to predict to what extent research will benefit one particular group of people, and it is especially difficult to predict if the participants are the ones to benefit. However, research should always minimize harm. This may seem to be most relevant where medical research is concerned, but we may also conceive of harming individuals' self-esteem, or their reputation among others, as well as the reputation of organizations.

Recommendations as to respecting human dignity

Now to the recommendations in the documents I have analyzed. The common suggestions from the documents analyzed are derived from the above aspects of human dignity and consist in the following main requirements on research:

- -the researcher should get "informed consent" from the participants
- - the researcher should preserve the anonymity of participants.

Let us start by considering when these two principles are applicable within Internet research. There is some disagreement among Internet researchers concerning whether or not research on data taken from the Internet should be regarded as research with human subjects, and thus lie under the principle of respecting human dignity. Researchers are allowed to collect, without consent, data from the public domain, such as data collected from television, public records, radio, printed books, conferences, or in public places such as parks.

There are two differences between the Internet and other media, however. The first lies in the fact that other public media usually have an intermediary between the person writing and the reader. This intermediary is legally responsible for the material presented to the public. The Internet media seldom has this kind of intermediary. Some electronic groups may have a moderator, but the responsibility of this person is far weaker than in public media. Thus, people writing within the Internet may feel less watched, and thus express themselves more freely, maybe even less responsibly. Second, some people who use the Internet may feel this as a private medium, far from newspapers or even public parks. The distinction between public and private domains is important for determining when consent is required. Is the Internet to be considered public or private?

In the document presented by Frankel & Siang (1999), the workshop participants have gone into great depth as to when informed consent should be sought. As pointed out by them, one of the issues concerns

" the blurred distinction between private vs public domain". As a solution to the disagreements between researchers, it is proposed that the issue should be treated with consideration both of media and of people. Some Internet media may be less private than other ones - organizational web-

sites as well as professional web-sites are definitely public, whereas usenet news postings are more private and personal e-mails are of course the most private medium.

It is too simple, however, to only talk about the technicalities of media. The context of the use is also important. Participants in a distribution list or in a usenet group may feel that they belong to a certain group of people, where they share private experiences. This is often the case with psychological support groups. In this case, the postings to this group are more private than postings to professional groups, aimed at the distribution of information or knowledge. In some closed groups, participants may be inclined to reveal more private information than they would do in face-to-face situations. In those cases, an intrusion by a researcher who does not inform of his/her investigation may disturb the dynamics of the group (King, 1996).

We see that no general rules can be given as to whether consent should be sought or not. Instead, the researcher should follow his/her sense of ethics and, preferably, seek the opinion of an ethical committee. If we and/or the ethical committee find that consent is required, what has then to be done? This issue is covered in next paragraphs.

Information

All participants who take part in an investigation must be informed about the intent of the investigation. The information should cover all the parts of the investigation that may affect their willingness to participate. It is interesting to see that most documents stress that the information given should be understandable by the prospective participants. This means that a possibly esoteric research purpose has to be expressed in every-day language, without any professional jargon.

This information may of course be more or less detailed. It should always include names and addresses of the responsible researchers, in order that participants may get further information. It is important that possible harm or damage is pointed out. Of course the researcher has to describe the possible utility of the research and offer the participants to partake of the results of the study at a later date. All documents also point out that the information should include the opportunity for participants to leave the study whenever they find it disagreeable or want to leave without any particular reason. There are discussions on how far the researcher may go to persuade the participants to stay with the study, in case they would like to quit. The information about the study should, if possible, be given before the study starts. However, it may also in some cases, where information might harm the purpose of the study, be given after the study. In this case, it should be provided as soon as possible.

We find various ways of giving this information within Internet research. In some cases, researchers have constructed special research sites. All persons who want to participate in the activities here, are informed that this is a research site as well as about the general nature of the research. In

other cases, researchers ask participants in an open channel to join them (and one another) in a particular channel - be it a discussion group or a chat-channel - in order to participate in the research. In both cases, the information about research leads the participants to choose whether they want to join or not. In other cases, when it is deemed that such self-selection is not desired, it is possible to inform participants after the data has been collected. Here, too, the issue of consent is crucial, and we proceed with this issue below.

Consent

When people choose to participate in an activity where research is being performed, this does not necessarily mean that they consent to being included in that research. Many research sites have an additional requirement that participants fill in a consent form before their data may be included in the research. The consent form should of course also be written in understandable language. Preferably, the consent should be given on paper, in handwriting. Some Internet research procures consent through the Internet itself. This might be less of a good choice, since it may be difficult to gauge if the person giving the consent is the same person as the one who is going to be studied. In some cases, it could be possible that other persons give their consent. This may happen for children or other wards. It may also be the case that the owner of an electronic discussion list or a newsgroup gives a general consent, depending on the nature of the list. In this latter case, the owner should explain to the participants that consent to participate in research is given by joining the list. We thus have a combination of self-selection to participate in the discussion and self-selection to participate in research.

Anonymity

The last general aspect of integrity refers to the issue of preserving the anonymity of the participants. In "real life" research there are quite explicit rules about how to go about maintaining anonymity. First, of course, all personal references are changed or deleted. However, the researcher may want to keep a personal record of the names of the participants, for further data treatments. Such a record has to be kept safe from persons who are not involved in the research. In particular, a participant can expect that data (including identification) is not used for other purposes by other researchers, or that identifiable data does not reach the public, for instance through journalists.

When we approach data from Internet research, we get a certain problem with the traditional procedure. First, in many Internet communication situations, people are either anonymous or use pseudonyms. In terms of ethical considerations, this presents the problem alluded to above, i.e. that the Internet identity studied does not correspond to the person who signs the consent form.

Another, more difficult issue concerns the treatment of pseudonyms. A researcher who has worked a lot with the aspect of ethics in Internet

research, writes as follows: "Researchers have traditionally disregarded pseudonyms as real identifiers and have quoted them directly along with the names of the newsgroups in their published research. (p 5) " (King, 1996). King questions this practice and points out that pseudonyms in cyberspace may be loaded with similar affections and values as real names in the world outside of cyberspace are. If people want to stay anonymous also within the cyber community chosen, the pseudonyms used cannot offer them any anonymity. Thus, in the information as well as in the consent form, the researcher has to take the decision of whether pseudonyms will be revealed or not. In some cases, it might be desirable from the research point of view to use pseudonyms, i.e. when the pseudonyms per se give information about the ambiance of the community. In such a case, consent must of course be sought to use the pseudonyms.

The same applies to names of virtual places. Often a particular virtual place is identified by its name. Frankel and Siang point out that "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." (op.cit. p 7)

The questions raised by these considerations are voiced by Frankel and Siang as follows: "How should a researcher cite an online text without violating the privacy and confidentiality of his or her subjects? How much description of an online community should a researcher provide?" (op. cit. p 13)

We see that the issues of the anonymity of Internet users are far from solved. This should not excuse researchers approaching the Internet from not making every attempt currently possible to preserve the anonymity of the research subjects.

Discussion and conclusion

The general conclusion is that research should provide more benefit than harm. However, the exposition here shows that it is problematic to propose that no harm is done, and even more so to claim what benefit research gives.

One clear weakness of current Internet research, a weakness that reduces its possible benefit, is related to its cultural bias. It was shown above that language users with English as their native language by far are over-represented in most Internet research. This may lead to biased conclusions concerning issues that are at stake in Internet discussions. Another aspect of this bias is that this article itself is based on an English biased medium, i.e. web-sites. I propose, however, that a search based on a library information retrieval method would not give a very different result. Also in traditional information retrieval practices, such as coding and thesauri, English is the dominating language.

We may thus conclude that the ethical guidelines found are based on Western culture in general and Anglo/Saxon culture in particular. It may well be the case that these guidelines place less value upon establishing trust and an intimate relationship between the researcher and the subject than other cultures. On the other hand, it might place higher value on privacy than other cultures. A continued investigation of ethical issues in various cultures is therefore greatly needed for research with the aim of studying global Internet use.

References

- American Psychological Association(1992). Ethical Principles of Psychologists and Code of conduct.
<http://www.apa.org/ethics/code.html>
- Frankel, M.S., Siang, S. (1999). Ethical and Legal Aspects of Human Subjects Research on the Internet.
www.aaas.org/spp/dspp/sfrr/projects/intres/main.htm
- NESH. Guidelines for Research Ethics
<http://www.etikkom.no/NESH/eretn.htm>
- HSFR. Etik. Forskningsetiska principer för humaniora och samhällsvetenskap. (Ethics. Research ethical principles for the humanities and social sciences. Uppsala: HSFR, 1996 (revised edition).
- King, S.A. (1996). Researching Internet Communities. Proposed Ethical Guidelines for the Reporting of Results. The Information Society, 12: 119-127.
- Natural Sciences and Engineering Research Council of Canada.
<http://www.nserc.ca/programs/ethics/english/intro03.htm>
www.funredes.org/mistica/
www.funredes.org/olistica/

Acknowledgement.

The research here presented has been supported by a grant from the Swedish Board for Research in the Humanities and Social Sciences.