Ethical Guidelines for Internet Research

Contents

Introduction

The nature of the internet and the rapid pace of change are giving rise to new and distinctive questions, obliging researchers to reflect on different, often conflicting considerations and norms in research ethics. In 2014, the National Committee for Research Ethics in the Social Sciences and the Humanities (NESH) therefore revised its "Research Ethics Guidelines for Internet Research" from 2003. The purpose of the guidelines is to provide guidance and advice for Internet research; they do not replace current legislation in the area.

The general framework for the specific guidelines is the NESH's *Guidelines for research ethics in the social sciences, law and the humanities* of 2006 (the text refers to the relevant guidelines as required). The current guidelines for research ethics apply to all research in the social sciences, the humanities, law and theology. It is the responsibility of the researchers to meet this research ethics obligation. When conducting research on/using the Internet, it is often unclear if or how these guidelines are adequate or relevant because the research can raise completely new issues of a methodological, epistemological or ontological nature. The ethical guidelines for Internet research are therefore intended as additional guidance for researchers. This is particularly true when the research directly or indirectly involves people. Internet research is a very comprehensive topic and can intersect with (and fall between) different academic disciplines and fields of research. Internet research can also raise questions of research ethics outside the field and the NESH guidelines. In such cases, consulting the *General guidelines for research ethics* may be useful.

Definitions

"Research" is broadly defined here, and covers research-based teaching and dissemination. Research also covers student work in the form of theses at the Master's and PhD level.

"Internet" basically refers to a global data network for communication and sharing of information. In a broader sense, "Internet" may refer to technological systems and units (things), different forms of use and application (tools), or a place of activity and interaction (social space). The various approaches raise different theoretical and methodological questions, which in turn open up new areas of research ethics assessments and considerations.

"Internet research" consequently covers: 1) research *on* the Internet as a phenomenon in its own right, its structure and technology (production systems, technical design, programs, codes, etc.); 2) *use of* the Internet as a research tool (collection of data and information using informants, surveys, archives, logs, algorithms, etc.); and 3) research on production, communication and interaction *using* the Internet (e-mail, social media, blogs, Internet forums, discussion pages, comments fields, or the use of computer games, search engines, web pages, etc.). Internet research is therefore not defined by the use of new methods.

In the following we will review key research ethics issues related to Internet research. They are particularly related to the researcher's obligation to base his/her work on fundamental respect for human dignity. This is elaborated on in the NESH's *Guidelines*, where researchers have an obligation to respect integrity, freedom and participation (point 6), to avoid injury and severe burdens (point 7), to have regard for third parties (point 11), to secure children's right to protection (point 12), to respect individuals' privacy (point 13), to respect posthumous reputations (point 17), and to respect the values and motives of others (point 18).

Public or private?

According to the NESH's *Guidelines*, researchers can in general use material from open forums freely without obtaining consent from the parties covered by the information (point 8). This must always be balanced against the requirement of respect for individuals' privacy and close relationships (point 13). It may be difficult to distinguish between the private and the public arenas in relation to information about behaviour that is disseminated and stored on the Internet. Researchers must therefore remember that there may be a difference between what people view as private and what technically is publicly-available information about people's use, communication, production and interaction.

For example, a researcher may freely use information from online newspapers' coverage of a case. In other situations, however, the distinction between the private and the public spheres may be less clear. One reason is that individuals who publish information on the Internet themselves may regard the information as private, even though it is technically publicly-available. Others may be of the view that the information is public, while also having strong reservations against the information being used by others in new settings and contexts, such as research. Some people view a personal blog as a public arena, while others consider the blog as publicly-available, but with private content. Different participants on an Internet forum may have differing views of what is private and public, and their patterns of communication and behaviour will be influenced by this view. Researchers have special responsibility for safeguarding the research subjects' integrity. In many situations, much can be gained by taking the "integrity of the context" into consideration: the context in which the information or communication takes place is important when assessing what is private and what is public. Researchers must not assume that all Internet users have a conscious view of or knowledge of which information will actually be made public and which information will not be made public.

This applies particularly, but not exclusively, to the use of social media, where the degree of publicness and users' ability to set personal data settings and publish location data varies.

Some Internet forums restrict access. Information provided here is not clearly public or private, but because the forum is not open to everyone, it cannot simply be used for research purposes without the researcher obtaining the consent of the subjects. The stricter the forum's restrictions on access, both in terms of the opportunity to acquire information/communication and the opportunity to participate in the forum, the greater the participants' expectations of protection or anonymity. Researchers must be aware of the access requirements of the forum being studied and how strict they are. Researchers must show greater care when there is stricter access restriction.

Free and informed consent

According to the NESH's *Guidelines*, research projects that include persons must generally only be initiated after securing the subjects' free and informed consent (point 9). In other words, researchers must weigh relevant considerations and determine whether it is necessary to obtain consent or not. This general rule also applies to research on/using the Internet.

There are many difficulties in connection with obtaining consent in digital forums. First, a request for consent to participant observation may be destructive for the very interaction the researcher wishes to study. Second, participation itself is ephemeral, while the information the researcher wants to use may be stored and available. As a result, contacting the people from whom consent is desired after the event may be difficult. Third, in some cases people pretend to be someone other

than themselves. This means that a researcher cannot be sure that the consent has been obtained from the person from whom consent was desired.

In cases where consent is necessary, these practical problems make demands on planning the research. Obtaining consent over the Internet also requires greater efforts to guarantee the quality of the consent. This means that greater precautions must be taken to ensure that persons are not recruited who should not be participating in the study, for example children in studies targeted at adults. Another problematic issue may be to ensure that the subjects adequately and correctly understand the information provided concerning the research and why the researcher requests consent, if this information is communicated in writing only, and over the Internet. In some cases, problems obtaining consent may require that the researcher should refrain from studying the forum altogether.

There is a distinction between researchers merely mining data and actively participating in a forum. If they actively participate in a forum with access restriction, researchers should present their intentions. Such disclosure should take place before starting the research. Similarly, there may also be reasons to make oneself known if information is obtained without active participation in the forum. Researchers must nevertheless take due consideration of ethical norms and any rules regarding behaviour that may apply to the forum. Researchers are responsible for clearly explaining to research subjects their role and the expectations, limitations and requirements that pertain to their role as researcher (point 19).

The consideration of respect for the subjects calls for the researcher to also inform subjects in open forums about systematic registration (e.g. recording) or reporting of information when possible.

Children's right to protection

Children and youths are valuable informants, and are entitled to be heard, also in the context of research. This also applies to Internet research on children and youths. Such research raises special challenges because it is part of children's activities, which are often beyond the control of adults, and the boundaries between the worlds of children and of adults are not very clear. However, this does nothing to diminish the requirement to obtain parental consent for children's participation in research, even though this may prove more difficult in practice (point 12). Here it is also important to verify the respondents' actual biological age. The children themselves must also grant informed consent to participate in the research.

In some cases, a better way to respect children's right to protection may be to use established research methods that do not involve the Internet as a research tool, even when collecting data on the Internet.

Personal data, confidentiality and anonymity

NESH's *Guidelines* stipulate that researchers must treat all information about private matters confidentially. The material must usually be anonymised to protect privacy and to prevent harm and a severe burden on the persons being researched (point 14).

Internet research is often based on interaction in digital forums (Twitter, Facebook, blogs, etc.), which is of an ephemeral nature, contributing to the expectation that the information is protected. In such situations, people may be willing to volunteer personal and sensitive information in an Internet forum that in principle is available to all people, without intending for the statements to be shared or studied (points 17 and 18). When a person's personal or sensitive data is available in an open forum and it is used for research purposes, they are entitled to the data being used and communicated appropriately (point 13).

The purpose of anonymity is to protect an individual's privacy, so that the information is treated properly and identifiable data is processed by as few people as possible.

Pseudonyms ("'nicks" or nicknames) are an important aspect of certain forms of interaction on the Internet. Pseudonyms are often used on different forums and in different contexts. Researchers should take into consideration that the use of the subjects' pseudonyms does not necessarily mean that their personal data is anonymised.

Big Data raises a whole set of new problems and challenges. Everyone leaves vast quantities of digital data – whether we use social media, swipe our cards, use our phones, buy apps or search the Internet – all of which can potentially be linked by researchers (or others) in ways the individual has no knowledge of or has granted informed consent to. It is also easier to track the informants' identity when they use digital forums than when they use other information channels. One might go as far as saying that linking *Big Data* and tracking IP addresses makes the concept of "anonymity" impossible in practice. Researchers may find it difficult to guarantee both that the data is, in fact, anonymous and that personal data will be deleted after the research project. Nevertheless, doing so is the researcher's responsibility. Potential informants must be informed as far as possible about these challenges and the possible consequences of the research. This is particularly true when the research leads to information being linked and generating new, sensitive information about identifiable individuals.

Regard for third parties

Different forms of interaction on the Internet often lead to the direct or indirect collection of data on persons who are not informants or whose data is used, so-called third parties. Regard for third parties means that researchers should consider and anticipate effects on third parties (point 11). Research on, using and with the Internet often includes data from other persons than the informants. For example, research that uses or looks at interaction in social media will provide access to third parties because they are closely linked to the primary informants in social or technical terms. This information may also be of a sensitive nature, and must be handled accordingly. The research may have an impact on the privacy and close relationships of individuals who are not included in the research, but who are drawn in as parties closely related to the informants, also through technical access to their information. The consideration of a burden on third parties should be weighed against the consideration of the critical function and the search for the truth of the research (see also point 7).

Use of quotes from the Internet

Research on/using the Internet can also use quotes. Quotes are understood here in a broader sense, and may include data feeds, location data, photos, videos, etc. When using quotes from the Internet in connection with publication of research results, researchers should carefully consider the fact that it is possible to trace the informants' and third parties' identity by performing a full-text or partial-text search on the quotes. At the same time, researchers are also responsible for the data and content remaining unchanged. Both these factors must therefore be kept in mind when considering whether to use a quote, and any assessments should be explicitly accounted for.

Reporting of results

Many social and cultural movements on the Internet are based on the norms of openness, freedom of information and sharing results. In connection with research on such digital communities, a special obligation arises to convey the results of the research back to the participants (point 47).

References

Association of Internet Researchers (AOIR), "Ethical Decision-Making and Internet Research", Recommendations from the AoIR Ethics Working Committee (Version 2.0)

Elgesem, Dag: "What is special about the ethical issues in online research?", *Ethics and Information Technology 4*, 3 (2002), pp. 195–203

Ingierd, Helene and Hallvard Fossheim (Eds.): *Internet Research Ethics*, Oslo: Cappelen Damm Akademisk, pp. 98–121

Livingstone, Sonia, Kjartan Ólafsson and Elisabeth Staksrud: "Risky Social Networking Practices Among 'Underage' Users: Lessons for Evidence-Based Policy", *Journal of Computer-Mediated Communication 18* (2013), pp. 303-320 [DOI: 10.1111/jcc4.12012]

Staksrud, Elisabeth: "Counting Children. On research methodology, ethics and policy development", in Helene Ingierd and Hallvard Fossheim (Eds.), *Internet Research Ethics*. Oslo: Cappelen Damm Akademisk, pp. 98–121