Ethical Decision-Making and Internet Research (version 2.0)
Recommendations from the AoIR Ethics Working Committee

AUTHORS

Annette Markham, Guest Professor, University of Aarhus, Denmark
Elizabeth Buchanan, Endowed Chair, University of Wisconsin-Stout, USA
with contributions from the AoIR Ethics Committee, including: Maria Bakardjeiva, (Canada), Andrea Baker (USA), Heidi Campbell (USA), Charles Ess (Denmark), Radhika Gajjala (USA), Mark Johns (USA), Steve Jones (USA), Heidi McKee (USA), Jim Porter (USA), Soraj Hongladaram (Thailand), Susannah Stern (USA), Leslie Tkach-Kawasaki (Japan), Leslie Regan Shade (Canada), Michele White (USA).

Additional Acknowledgments: David Brake (UK), Abbe Foreman (USA), Jeremy Hunsinger (Canada), Stine Lomborg (Denmark), Asa Rosenberg (Sweden), Lois Scheidt (USA), Michael Zimmer (USA).

HOW TO REVIEW THIS DOCUMENT

This version of the AOIR Ethics Guidelines is posted here for AOIR member review. As this WORKING DOCUMENT is not yet approved by the AOIR membership, please do not distribute.

This document has been approved by the AOIR Ethics Committee. It will be posted for general AOIR member review from September 5 through September 30, 2011. If you prefer to write your comments via email, please send directly to Elizabeth Buchanan, Ethics Committee Chair. If you prefer a printable version of the entire document, please click here.

The working document has undergone multiple revisions. Although it remains a draft, we believe it represents the general AOIR community stance toward Internet research ethics. We appreciate any feedback you have, knowing that it will improve the quality, scope, and eventual utility of this document. Thank you kindly for your time and energy toward this effort.

The committee’s goal in putting this together was to augment, rather than replace, the first AOIR Ethics Guidelines document, published in 2001. That initial document remains relevant and useful. Please review this version 2.0 with our overall goal in mind: The broader research community relies on AOIR to supply an updated set of guidelines regarding ethics and internet research. This document serves to guide ethical practice but cannot be all inclusive or comprehensive of all disciplinary, institutional, and cultural stances.

While we recognize that the diversity of our viewpoints and cultures will prohibit complete consensus on the details of this document, we must balance this with the need for new researchers and research governing bodies to have a public statement from us. After six weeks of online review, we will host a feedback session at the AOIR conference in Seattle, Washington. After the conference, the committee will confer, make additional revisions based on member feedback, and produce a final version. The AOIR community will have two weeks to review the final document and vote on whether or not it should be accepted.

Again, thanks for your help in making this document as strong as possible. Please make comments in the blog version of this document [http://aoirethics.ijire.net] or email comments to Elizabeth Buchanan [buchanane@uwstout.edu] and Annette Markham [amarkham@gmail.com], and they will share those comments with committee members.

There will also be a feedback session at AoIR 12 in Seattle about the draft guidelines. If you are attending AoIR, please consider attending on Tuesday, October 11, at 4:00 (room to be announced) to share your thoughts with committee members.
INTRODUCTION AND BACKGROUND

The work represented in this version emerges from theoretical, empirical and field research conducted by members of the Association of Internet Researchers, including members of the AoIR Ethics Working Group.

The first release of the AoIR Ethical Decision-Making document occurred in 2002, after two years of international and cross-disciplinary collaboration. The document and its guidelines emerged from a series of extensive dialogues among researchers experienced with facing and resolving ethical issues in Internet research, philosophers, and other members of AoIR’s international, interdisciplinary community. The intention was to develop guidelines from the bottom up -- i.e., out of the day-to-day practices of researchers in a wide range of disciplines, countries and contexts, in contrast with a more usual top-down approach that tries to apply norms, principles, practices, etc., in a deductive way. This approach was crucial because the enterprise of Internet research is expansive -- that is, globally informed -- but also situated in innumerable locales.

Workshops conducted annually at the AoIR Internet Research conferences, from 2005 through the present, have provided extensive groundwork from which this document emerged. Given the interdisciplinary and intercultural research of AoIR, more than one set of (often conflicting) norms, principles, values, policies, or laws can be seen to apply to a specific case/issue. The document has subsequently received much use, and has been cited and used in a wide range of publications by a diverse number of disciplines (see Appendix 1). The AoIR Guidelines document has also been used by research ethics boards (REBs) and institutional review boards (IRBs) when making decisions about Internet research-based protocols (Buchanan & Ess, 2009).

Ethical Pluralism

As with the original AoIR document, this revision stresses ethical pluralism in its approach and advocates for guidelines, not recipes. Ethical conundrums are complex and rarely decided along binary lines. There is much grey area in ethical decision-making. More than one set of norms, values, principles and usual practices can be seen to legitimately apply to the issue(s) involved. Hence, it becomes difficult to make judgments as to which set(s) apply, especially when one set conflicts with another in some way. This forces the researcher to determine which
is more relevant in a given context. In contrast with those ethical judgments that usually do not
give the researcher any trouble (because they are more or less unambiguous, algorithmic, “top-
down” deductive applications of accepted norms and principles), ethical conundrums require us to
first choose which norms, values, principles, and/or practices apply - and for this there are only
guidelines, not strict rules. Multiple judgments are possible, and ambiguity and uncertainty are
part of the process and difficulty. This document strives to maintain a stance of ethical pluralism.

**Guidelines, not Codes**

We advocate guidelines rather than a code of practice so that research can retain a flexible
quality over time, be responsive to diverse contexts, and be adaptable to continually changing
 technological contexts. When one considers that ethical assessments are always operationalized
via some sort of practice (method), it becomes clearer that an adaptive, inductive approach can
yield potentially more ethically legitimate outcomes than a simple adherence to a set of
instantiated rules.

Such an approach takes on board precisely the fine-grained details and context of each specific
case or issue. The emphasis on a dialogical approach highlights the notion that there exists a
range of possible, ethically legitimate judgments and decisions. In contrast with the comparatively
straightforward and unambiguous top-down deduction from a given principle, this approach thus
highlights the researcher’s responsibility for making such judgments and decisions within specific
contexts – a responsibility that is often all the more uncomfortable as it often requires choices
between a range of competing values and norms. This approach not only privileges the often
subtle details of a specific context, but also preserves the freedom and integrity of the researcher
as the one ultimately responsible for his or her ethical choices. Hence, this revised version of the
AoIR Ethical Decision Making document adheres to the same underlying principles as the first: A
dialogic, case-based and inductive approach to ethics.

**Limitations, or What This Document is Not**

There are limitations to the current document, as expected. We do not reiterate the many specific
rules of governmental or institutional research policies, as these are well documented and widely
available. We also do not include specific disciplinary best practices or codes of ethics, as this
document seeks to function at the macro-level.\(^1\)

Although we identify current Internet technologies and contexts, we acknowledge that
technologies themselves change rapidly. Therefore, this document is designed to emphasize
processes for decision-making and questions that can be applied to ever-changing technological
contexts.

At its most fundamental level, we recognize that ethical decision making interweaves one’s
fundamental world view (ontology, epistemology, values, etc), one’s academic and political
environment (purposes), and one’s defining disciplinary assumptions. Decision making occurs
across a range of possible junctures in the cycle of inquiry, including research design, research
conduct, and research production and dissemination. Because of the complexity of ethical
decision making in the individual case, this document focuses on general principles and only a
fraction of possible ethical dilemmas that may arise.

**Purpose and Audience**

While the first AoIR document thus enjoyed extensive use, much has changed in the field of
Internet Studies since 2002. The scope and contexts of Internet research have been dramatically

---

\(^1\) Such as the Forum for Ethical Review Committees in Asia and the Western Pacific
expands through the continuing global diffusion of the Internet into nearly every country in the
world, as facilitated through a growing array of devices (including game consoles, Internet-
enabled phones and other mobile devices) and ever-increasing bandwidth; rapidly expanding
suites of new communication applications; and the increasingly seamless interweaving of online
and offline activities and experiences. Alongside these developments, the literature of Internet
research ethics has grown considerably, providing us with a far more extensive range of
theoretical resources and practical examples to help recognize and guide ethical reflection.

The AoIR ethics working committee, with extensive feedback from AoIR membership, has
developed this version 2.0 document in an effort to recognize and respond to the array of
changing technologies and ongoing developments that affect the ethics of Internet research.

This document aligns with the purpose of the first AoIR document and serves to extend, not
replace, the previous document: It represents a series of considerations designed to support and
inform those responsible for making decisions about the ethics of Internet research; it provides a
resource for a wide audience of researchers, review boards, ethicists, and students by providing
a current discussion of important ethical issues and pertinent literature in the field; and it connects
these individuals to a larger researcher community through interactive fora and social media
resources.

This document can provide support for organizations and related groups that commission, fund or
have overall responsibility for or an interest in Internet research practices in national and
international contexts and can be used to help inform any such bodies of the ethical issues that
might be considered and possible ways of resolving ethical problems.

INTERNET RESEARCH

Internet research ethics literature has explored the question: Are there unique ethical issues
within Internet research itself or are they simply “traditional” issues emergent in a new media? We
take the stance that there are novel ethical issues and dilemmas in Internet research that exist
precisely because of its particular qualities. Many ethical challenges may be successfully
resolved through more traditional and established approaches: but when the limits of those
approaches are reached, Internet research calls for new models of ethical evaluation and
consideration.

Definitions

The Internet is a social phenomenon, a tool, and also a (field) site for research. Depending on the
role the Internet plays in the research project or how it is conceptualized by the researcher,
different epistemological, logistical and ethical considerations will come into play. The term
"Internet" originally described a network of computers that made possible the decentralized
transmission of information. Now, the term serves as an umbrella for innumerable technologies,
devices, capacities, uses, and social spaces.

Because the types of interaction and information transmission made possible by the Internet vary
so widely, researchers find it necessary to define the concept more narrowly within individual
studies. This is complicated by the fact that the studies of and on the Internet cut across all
academic disciplines. There is no single set of methodological or theoretical guidelines that can
be assumed to hold universally.

This document uses the following working definition: Internet research encompasses inquiry that
(a) utilizes the Internet to collect data or information, e.g., through online interviews, surveys,
archiving, or automated means of data scraping; (b) studies how people use the Internet, e.g.,
through collecting and observing activities or participating on listservs, web sites, blogs, social
network sites, gaming, or other online environments or contexts; and/or (c) utilizes datasets,
databanks, repositories available via the Internet. Internet research is not machine specific or dependent, and we recognize the impact of smart devices, cellular, and space-less Internet activities.

It is important to note that definitions of and experiences with these contexts vary widely. Also, technological convergence merges many of these categories in evolving and sometimes surprising ways. The Internet mediates everyday life in industrialized cultures, whether or not we are actively using a "browser" on a computer. Thus, Internet research should be considered in its broadest sense.

The Distance Principle

As knowledge production becomes more dependent on information derived from some sort of Internet research, we recognize that the "objects" of scientific research are more person-based than not. This has become clear to our group over the past twenty years of Internet development. For instance, a researcher cannot talk about discursive or physical participation in the online information sphere without acknowledging that this necessarily involves a person somewhere in the process—a person who is thinking and behaving within his or her own cultural and moral stances.2

The Internet complicates the fundamental research ethics question of personhood (Is an avatar a person? Can we assume a person is wholly removed from large data pools?) and harm (does the connection between an online identity and his or her real-life person enable psychological, economic, physical, harm?). Due to the complexity of Internet contexts, harm may not be immediately visible, but may emerge at any point in the research process.3 The fluidity of Internet contexts requires that researchers attend to ethical issues throughout the process, not simply at the beginning. Researchers and reviewers should be prepared to address this reality.

One way of evaluating the extent to which the ethical dilemmas may be hidden is to focus on the way that procedures for data collection or analysis extract data from the context. As the conceptual or experiential distance between the researcher and author/participant decreases, we are more likely to naturally define the research scenario as one that involves "humans." For example, an email or virtual worlds interview produces data that is experientially near the participant. In this situation, it is likely that the researcher would recognize that this information gathering process involves a human participant.

As the conceptual or experiential distance between the object of research and the person who produced it increases, there may be a tendency to define the research scenario as one that does not involve "human." For example, a data set containing thousands of tweets or an aggregation of surfing behaviors collected from a bot is conceptually distant from the persons who engaged in these activities. In this latter scenario, however, there is a greater possibility for forgetting that there was ever a person somewhere in the process that could be directly or indirectly impacted by the research.

It is apparent that that Internet technologies have the capacity to effectively blur the boundary between a private person and a public text, and that persons hold varying standpoints in relation to their production and consumption of information. Considering this distance principle can help one recognize that the relationship between persons and data is not always clear. For example, considering the question: “Is this a text or a person?” in a particular case involves a process of

---

2 This does not necessarily equate to definition of the person as a “human subject” as defined and by the U.S. regulatory model, but may suggest consideration of the premises underlying this categorization and adoption of practices suggested within this model.

3 See for example Gaijala’s experience with South Asian listserv communities (2004); Stern’s experience with minors (2003).
analyzing the conceptual distance between the object of research and the persons whose activities created this data, consulting a range of perspectives from previous studies, identifying and assessing individual or community expectations and definitions, and making an ethical judgment.

GUIDING RESEARCH ETHICS PRINCIPLES

Principles of research ethics and ethical treatment of persons are codified in a number of policies and accepted documents, such as the UN Declaration of Human Rights, the Nuremberg Code, the Declaration of Helsinki, the Belmont Report, and the African Charter on Human and Peoples’ Rights. At their core, the basic tenets include the fundamental rights of human dignity, autonomy, protection, safety, maximization of benefits and minimization of harms, and knowledge, or, in their most currently accepted phrasing, respect for persons, justice, and beneficence. While originally stemming from the biomedical contexts, these principles have been adapted beyond these early contexts and rise above disciplines and methodologies. We accept them as basic to any research endeavor. In addition to these principles, researchers must be attendant to extant legal requirements in the countries implicated in the research.

Ethical research is that which seeks to minimize harm. Defining and identifying potential ‘harm’ in Internet research can be difficult, as technologies blur boundaries of traditional, more easily identifiable dichotomies, such as public-private, researcher-researched, subject-object. The ethical dilemmas that arise in Internet research often resists yes-no binaries, thus making decision-making flow charts used by some research boards inapplicable. For example, the U.S. Department of Health and Human Services offers Human Subjects Regulations Decision Charts where one of the first questions is “Does the research involve intervention or interaction with the individuals?” But what constitutes intervention or interaction on the Internet, particularly in social networking sites, virtual worlds, and massively multiplayer online games, is often not so easy to determine, calling for careful consideration of the particular circumstances of the context and careful consideration of other relevant cases (see bibliography for sources grouped by topic).

The following sections are designed to help guide one’s consideration. We begin with a brief list of broad ethical considerations that ground all research, followed by a list of questions more attuned to Internet research. These detailed sets of questions may enable a researcher to become aware of ethical dilemmas and challenges that may go otherwise unnoticed. We then provide a heuristic chart that lays out a similar array of questions arising in studies of particular contexts/venues or with particularly types of information or data produced or retrieved from those contexts.

As emphasized in the casuistic, or case-based approach to Internet research ethics developed by McKee and Porter (2009), reviewing other researchers’ approaches can provide examples of the range of judgments possible, to guide one’s own judgments in relevantly similar cases. To aid in further consideration, we provide in the bibliography a comprehensive listing of references that will give researchers and reviewers a range of studies that encounter or discuss specific types of ethical challenges in Internet settings.
Broad Ethical Considerations

The greater the vulnerability of the community/author/subject, the greater the obligation of the researcher to protect the community/author/subject.

Because 'harm' is defined contextually, ethical principles are more likely to be understood inductively rather than applied universally. That is, rather than one-size-fits-all pronouncements, ethical decision-making is best approached through the application of practical judgment attentive to the specific context (what Aristotle identified as phronesis).

Because all data at some point involves individual persons, human subject considerations may be necessary even if it is not immediately apparent how and where persons exist in the information stream.

When making ethical decisions, researchers must balance the rights of subjects (as authors, as research participants, as people) with the benefits and rights to research of researchers and of research. In different contexts the rights of subjects may outweigh the benefits of research.

Ethical issues arise and need to be addressed during all steps of the research process, from planning to publication.

Ethical decision-making is a deliberative process, and researchers are well-served to consult as many people and resources as possible: fellow researchers, people participating in or familiar with contexts/sites being studied, research review boards, ethics guidelines, published scholarship (within one’s discipline but also in other disciplines), and, where applicable, legal precedent.

Internet Specific Ethical Questions

The list below begins with common questions asked in the course of a research project. The bullet points represent more specific considerations that arise in Internet-related contexts.

What is the context or venue?

- How is the context defined and conceptualized in the project? Does this match the way owners, users, or members might define it? (e.g., parameters such as 'culture,' 'person,' 'data set,' and 'public text' each carry different ethical expectations for researchers)
- What are the ethical expectations of the venue, particularly around issues of privacy? Both for individual participants as well as the community as a whole?
- How are subjects/authors situated in the context?
- How is the researcher situated in the context?

Who is involved in the study?

- What are the ethical expectations of the community/subjects/authors?

---

4 For further discussion of phronesis as applied to Internet research ethics and to digital media ethics, see for example Ess (2009) and McKee and Porter (2009).
5 The AoIR discussion list and archives (http://listserv.aoir.org/listinfo.cgi/air-l-aoir.org) provide many examples of researchers seeking advice from a world-wide community.
6 See for example Hudson & Bruckman (2004); Gajjala (2004); Ikonomidis Svedmark & Nyberg (2009); Baker (2009).
7 See for example Markham (2009).
• What is the ethical stance of the researcher? (E.g., a mismatch between the ethical stance of the researcher and the community/subject/author may create ethical complications).8
• What are the ethical traditions of researchers’ and/or author/subjests’ cultures or countries?

What is the primary object of study?

• What are the ethical expectations commonly associated with this type of data? (E.g., working with aggregated, de-identified data carries different ethical expectations than working with interview data.)
• What other questions might arise as a result of the particular context from which this data was collected?

What are the methods and methodological approaches of the researcher?

• What are the ethical expectations of the research community associated with the particular approaches (e.g., ethnographic, survey, linguistic analysis)?
• What potential harms might arise as a result of one’s choice of methods? (E.g., would quoting directly from a blog cause harm to the blogger and if so, could another method of representation be less risky?)9

What are the potential harms or risks associated with this study—for individuals, for online communities, for researchers, for research?

• How are the concepts of ‘vulnerability’ and ‘harm’ being defined and operationalized in the study? How are risks to the community/author/subject being assessed?
• Would a mismatch between researcher and community/subject/author definitions of ‘harm’ or ‘vulnerability’ create an ethical dilemma? If so, how would this be addressed?
• What harms—to life, to career, to reputation—may occur from the research? (e.g., would the research “out” an LGBTQ individual who is not publicly out and perhaps cause them to lose their jobs? Would the research cause someone to face criminal or civil penalties?)
• What possible harms to privacy may occur? (e.g., online groups disbanding or individuals ceasing to use an online support group or withdraw from blogging activities because of the presence of researchers;10 individuals feeling that real or perceived privacy has been violated;11 individuals being made anonymous against their wish to remain known and public in any published results)
• Who or what else could cause harm to the author/participant beyond the researcher?
• Are we acting in ways that minimizes risk?
• Does our research adequately protect the researcher as well as the community/author/subject?

8 See for example Ayers (2003).
9 See for example Ikonomidis Svedmark & Nyberg (2009).
10 See for example Bromseth (2002); Kraut et al., (2003).
11 The complexity of this situation has been discussed and studied extensively. Sveningsson-Elm (2004) provides a useful model for mapping distinctions between actual and perceived privacy; boyd & Marwick (2011) offer recent discussions with teens to complicate the notion of privacy, including a keen example of how teens felt violated by the use of their profiles in a school-wide presentation by administrators, even as they knew their information was public (p. 6). Nissenbaum (2010) and Nippert-Eng (2010) write extensively about the complications of privacy. Also, the Association of Internet Researchers list archives contain extensive discussion of this issue.
What are potential benefits associated with this study?

- Who benefits from the study--do the potential subjects? If not, what greater benefit justifies the potential risks?
- Is the research aiming at a good or desirable goal?

How are we recognizing the autonomy of others and acknowledging that they are of equal worth to ourselves and should be treated so?

- Will informed consent be required from subjects?
- If so, what procedures to obtain consent will be followed? (E.g., print or digital signatures, virtual consent tokens, click boxes or waiver of documented consent)
- Will consent be obtained just from individuals or from communities and online system administrators?
- What harm might result from asking for consent, or through the process of asking for consent?
- How will the researcher ensure that participants are truly informed?
- Are vulnerable persons being protected? How is vulnerability determined in contexts where this categorization may not be apparent?
- How will parental or guardian consent be obtained in addition to assent where required by research regulations?

What issues should be considered because of the technologies involved in the conduct of this study?

- In situations whereby consent is desired but written informed consent is impossible (or in regulatory criteria, impracticable) or potentially harmful, will procedures or requirements be modified? What other ethical concerns might arise if informed consent is not obtained?
- If access to an online context is publicly available, do members/subjects/authors perceive the context to be public? What considerations might be necessary to accommodate ‘perceived privacy’ or the notion that individuals might care more about the appropriate flow of information as defining it as public or private?
- In situations where identity, age, and ability of the participant is unknown or hidden, and harm cannot be determined as an a priori category based on known vulnerability of participant, how will harm be considered as an ethical concern and operationalized in the study?
- If information collected in the course of a study can be linked back to an individual by means of Internet search or other technology, what process will the researcher use to determine how that information will be treated? (E.g., many challenges surround the responsible use of images and video). Is this data considered by the participant to be personal and private or public and freely available for analysis and re-publication?

---

12 Although in some cases, traditional informed consent procedures may be warranted, in other cases, these may not be the most ethically sensitive approach. As recommended by the 2010 AOIR Ethics workshop: “Informed consent may happen at different points throughout a study, sometimes it may be more ethical to get informed consent at the end when you want to present a specific case study or quote an individual or focus on a particular element. Therefore, informed consent should be always an inductive process” (Buchanan, Markham, & Ess, 2010).
13 See for example Stern (2003); Ikonomidis Svedmark (2011).
14 For further discussion of this, see for example Sveningsson (2003), Sveningsson-Elm (2008) and McKee & Porter (2009).
15 For further discussion, see Nissenbaum (2010).
16 See for example Sweeny (2003).
• What method is being used to secure and manage potentially sensitive data? What unanticipated breaches might occur during or after the collection and storage of data or the production of reports?¹⁷ (E.g., if an audience member recorded and posted sensitive material presented during an in-house research presentation, what harms might result?)

**Charting Ethical Questions By Data And Type**

This following chart is intended to provide a reference guide for researchers and research reviewers to recognize where a particular research project might fit in the Internet research continuum or in what specific type of venue a project exists, what types of information/data may be derived from these contexts, and what common ethical questions have been asked within these contexts.

This should be viewed as a conceptual heuristic comprised of overlapping and fluid categories. Notably, there are varying ways to frame these contexts and varying ways to perceive one's research in or of these contexts. Two people can be utilizing the same technological context and yet have very different perspectives on it. These distinctions are also blurred as technologies converge and become less obvious as a part of everyday life.

Researchers and review boards can refer to this chart as a guide or checklist of considerations when writing or reviewing research proposals, developing methodological approaches for a particular study, designing specific protocols, and so forth. Ideally, a researcher should be able to answer these questions and justify those responses within the context of his or her research question, discipline, specific Internet venue or context, and region or jurisdiction.

¹⁷ See for example Zimmer (2010).
This chart provides a useful starting point for internet researchers to consider ethics. Convergence of technologies and capacities continues to break down the strict boundaries between these categories. While not intended to provide answers, it promotes consideration of a range of issues and questions that may become relevant in the course of any internet related research.

<table>
<thead>
<tr>
<th>Types of Data collected</th>
<th>Types of Venues/Contexts</th>
<th>Commonly asked questions about ethical practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct communication</td>
<td></td>
<td>How is protection of autonomy of participant/author achieved through informed consent or protection of vulnerable persons? How can researcher ensure that author/participant understands and agrees that content or interaction may be used for research purposes? Is the communication archived or easily searchable and retrievable? Is the data subject to open data laws or regulations? How long does the third party provider or ISP preserve the data and where? Could privacy be achieved through anonymization of email content and/or header information?</td>
</tr>
<tr>
<td>Special Interest Forums</td>
<td>(e.g., email- or web-based conversations and archives, e.g., threaded discussion forums, chatrooms)</td>
<td>How do terms of service (TOS) articulate privacy of content and/or how it is shared with 3rd parties? Regardless of TOS, what are community or individual norms and/or expectations for privacy? Does the author/participant consider personal network of connections sensitive information? Is the data easily searchable and retrievable? If the content of a subject’s communication were to become known beyond the confines of the venue being studied – would harm likely result? Is the conversation thread or forum perceived as public or private by the author(s)/subject(s)? How is profile, location, or other personally identifying information used or stored by researcher? Is the data easily searchable and retrievable? How is informed consent or protection of privacy achieved? How are vulnerable persons identified and protected? If non-active archives are used, how is vulnerability or harm defined and how are potential or actual subjects protected?</td>
</tr>
<tr>
<td>Social Network Sites</td>
<td>(e.g., LinkedIn, googles+, Facebook, Myspace, Flickr, FourSquare)</td>
<td>How do the terms of service articulate privacy of content and/or how it is shared with 3rd parties? Does the author/participant consider personal network of connections sensitive information? Is the profile or location information used or stored by researcher? Does author/participant understand and agree to interaction that may be used for research purposes? Does research purpose and design balance possible conflicts between participant and researcher perceptions of public/private and sensitive/nonsensitive? Does the dissemination of findings protect confidentiality? Is the data easily searchable and retrievable? If the content of a subject’s communication was ever linked to the person, would harm likely result?</td>
</tr>
<tr>
<td>Personal sites/blogs</td>
<td>(e.g., homepages, blogs, youtube, and all forms of multimedia presentation)</td>
<td>Could analysis, publication, redistribution, or dissemination of content harm the subject in any way? If the content of a subject’s communication were to become known beyond the confines of the venue being studied would harm likely result? Does the author/participant consider personal network of connections sensitive information? Does author/participant consider the presentation of information or venue to be private or public? Does the terms of service conflict with ethical principles? Is the author/subject a minor?</td>
</tr>
<tr>
<td>Avatar-based social</td>
<td>spaces , virtual worlds, and online gaming spaces (e.g., Second Life, SIMs, MUDS/MOOS, MPORPG)</td>
<td>Should these virtual worlds be considered “public”? What constitutes “privacy” in such places? Should avatars be considered as persons and afforded the same protections as human subjects? Will the process of requesting consent itself cause harm? How and when should consent be sought? What requires consent? To what extent do users perceive their interactions and communication to be private in these spaces? How Do Terms of Service specify researcher presence, anonymity of users, and privacy/ confidentiality? To what extent and in what ways could research activities interfere with or compromise a user’s play or outcomes in the game? How should researchers juggle their own multiple roles? Could data be used to identify a user’s physical location and other sensitive demographic information?</td>
</tr>
<tr>
<td>Commercial Web Services</td>
<td>(e.g., Google, AOL, Yahoo, Bing, MSN, SurveyMonkey, Cloud Storage)</td>
<td>What are the participant/author’s expectations of privacy? Is the data easily searchable and retrievable? Is the data subject to open data laws or regulations? Does the service’s privacy policy contradict ethical principles? What measures safeguard data at the site of data collection? How long will the data be stored on the servers? Does this contradict the time frame indicated by the researcher or institutional policies? What happens to the data after the researcher completes work on the service? How are the data destroyed? How will cross-border data be handled if IP addresses are considered by one country to fall under privacy regulations?</td>
</tr>
<tr>
<td>Databanks/Repositories</td>
<td></td>
<td>Where is the data stored? How long will the data exist in the repository? What consent is needed for subsequent data use? Does the remixing/mashing of data enable identification of individual or group identities or enable any additional risks to participants? In the case of shared data, what conditions were placed on data use by the original researcher, if any? Regardless of conditions, what ethical responsibilities may require consideration by later users? What mechanisms are in place to ensure appropriate data provenance and ownership? How will images/audio be effectively anonymized?</td>
</tr>
</tbody>
</table>

Locations and movements
- Physical locations (GPS)
- Physical movements
- Surfing behaviors

Interactions, behaviors, transactions
- Hyperlinks
- Comments or Recommendations
- File or Information Sharing (file or snippet)
- Forwarding /Replying
- Interpersonal Interactions, conversations
- Networks (e.g., maps visualizing communication flow or strength of relations between persons)

Production, presentation, performance
- Texts (e.g., authored texts, naturally occurring discourse, interview transcripts)
- Images (presented or produced by user or captured by researcher)
- Video (presented or produced by user or captured by researcher)
- Audio (presented or produced by user or captured by researcher)
- User motions and movements (in online training, virtual spaces, produced or presented by user or captured by researcher)

Archived information
- Demographic information
- Bookmark collections
- Discussion archives
- Data banks
- Transaction logs
- Clickstream data
- Trace data

Created by Annette Markham. Draft for AOIR Guidelines for Internet Research Ethics (version 2.0). DRAFT ONLY
Potential Areas For Unethical Or Misuse Of Internet Data

At various points in the research process, opportunities for unethical behaviors or acts arise, regardless of context or methodology. In reviewing the literature specific to Internet research ethics, and based on Buchanan and Ess's study of US-based IRBs (Buchanan & Ess, 2009), we have identified the most common forms of violations that can occur in Internet research.

- Failing to respect expectations of privacy within particular contexts
- Failure to obtain consent for research from individuals, groups, or service providers
- Failure to adequately anonymize data
- Using deceptive practices when unwarranted by research approach or question
- Failure to preserve the integrity of subject/actor’s intellectual property or creative works
- Causing or facilitating psychological, economic, informational, or reputational harms
- Dismissing, ignoring, or violating terms of service
- Failure to adhere to local/regional, as well as cultural laws and/or directives

CONCLUSION

This document provides a basic overview of ethical Internet research considerations. It advocates an emergent approach to ethics, which considers the specific needs of each case. Therefore, rather than prescribing a set of approved practices, it suggests a characteristic range of questions that Internet researchers and those responsible for oversight of such research should consider. But of course, this document is neither complete nor final. Over two decades of research practice have provided us with multiple examples of ethical complications and unique solutions. A strong knowledge of previous studies is advocated and we provide as comprehensive a bibliography as possible to aid in that endeavor. As well, we recognize that as Internet researchers encounter new venues, contexts, inter/actions, etc., additional questions and responses will inevitably arise, either as variations of those mentioned in this document or as distinctly new. Our intent is that this document will remain useful in those new contexts as it at least suggests starting points and possible analogies for raising new questions and developing new responses.

At the same time, we anticipate that as this document comes into use, it may help researchers further identify its limits as they encounter new contexts and ethical issues that emerge in the ongoing evolution of both the Internet and our multifaceted efforts to research the communicative engagements these technologies make possible. Hence, we intend this document to serve both as a first way of addressing contemporary ethical challenges and as an occasion for inviting further critical evaluation and ongoing dialogical development of Internet research ethics.

In sum, we hope this document will prove helpful to researchers, ethicists, and others concerned with the important ethical challenges of Internet research, and the Association of Internet Researchers remains grateful for further critical comment and examples that help to carry our dialogues onward from this point.
WORKS CITED


**NOTE ABOUT FORMATTING AND STYLE:**
This document is a working copy. It will undergo thorough copy editing. Please feel free to make suggestions and recommendations to the authors.