Research Design:

Ethical Issues in Designing Internet-Based Research: Recommendations for Good Practice

Shikha Gupta
Faculty of Health Sciences, Queen’s University
Kingston, Ontario K7L 3N6, CANADA
15sg48@queensu.ca

Abstract

This article presents an overview of internet-based research, highlighting the absence of a standard terminology to define and classify such research. The label internet-based research or online research can cover a diverse range of research designs and methods, involving different degrees of ethical concern regarding privacy, transparency, confidentiality, and security. Although the basic principles of human research ethics (such as respect, integrity, justice, and beneficence) are still applicable in this context, interpreting and applying these principles correctly and protecting the interests of the research participants effectively are not easy to ensure. While the nature of the internet poses challenges of user authentication and confidentiality, the diversity of national laws and codes of ethics poses additional challenges. The article refers to relevant Canadian laws, with which the author is familiar. Finally, a set of recommendations are offered to mitigate the ethical challenges of internet-based research. These include ethical practices such as ensuring transparency while recruitment, considering participants’ expectations about privacy, ensuring legal compliance, using secure communication protocols, obtaining informed and knowledgeable consent, offering participants the opportunity to withdraw from the research and retract their data, and ensuring that data are not used for subsequent non-research purposes.

Index Terms: online research; internet-based research; internet research ethics; human research ethics; ethics review; social media

1. Background

The role of internet has been expanding in the overall research landscape (Gelinas et al., 2017). Researchers around the world are using internet-based methods for some parts of their research cycle (such as planning, recruitment, data collection, knowledge dissemination), or for accomplishing the entire research project (Lafferty & Manca, 2015). The number of internet users increased from 16 million in 1995 (i.e., 0.4 % of the world’s population) to 3,696 million in 2016 (i.e., 49.5 %) (“Internet Growth Statistics,” 2017). This increasing internet penetration has provided an immense opportunity to researchers to overcome geographical divide, build global research communities, and use technological advancements for conducting and reporting research online (Lafferty & Manca, 2015).

Internet-based research now constitutes one of the most common data collection methods around the world (Maronick, 2009). Buchanan and Hvizdak (2009) reported that in a survey of 750 university Human Research Ethics Boards, internet research involving online or web surveys were the most common type of method that was proposed and reviewed by the boards. Not only online and web surveys, a new range of approaches are emerging such as data collection through virtual observation in interactive spaces such as blogs, websites, chat rooms, social media sites; geographical mapping; and so on (Warrell & Jacobsen, 2014). Social networking platforms such as Facebook, Twitter, or LinkedIn are now commonly used for gaining access and recruiting potential participants. This trend is common in business and health care fields (Horevoorts, Vissers, Mols, Melissa, & Van De Poll-Franse, 2015; Hunter, 2012; Taylor, Kuwana, & Wilfond, 2014; Touvier et al., 2010).

While these online platforms provide an opportunity for researchers to reach large number of people in a relatively short time, their use also raises considerable ethical concerns (Buchanan & Hvizdak, 2009; Flicker, Haans, & Skinner, 2004; Keller & Lee, 2010; Warrell & Jacobsen, 2014). In the following sections of the article, I discuss two key questions related to online research methods:

Q1. What is online or internet-based research and how many types of online research are there?

Q2. What are the key ethical concerns that researchers need to consider while employing online research methods?

2. Online Research Methods: Definition and Classification

In the past few years, many researchers and organizations have attempted to define and classify online or internet-based research (Eynon, Fry, & Schroeder, 2011; Kitchin, 2007; Markham & Buchanan, 2012). According to the Association of Internet Researchers, internet research encompasses enquiry 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 and access the internet, e.g., through collecting and observing activities or participating on social network sites, web sites, blogs, games, virtual worlds, or other online environments or contexts;

(c) utilizes or engages in data processing, analysis, or storage of datasets, databanks, and/or repositories available via the internet;

(d) studies software, code, and internet technologies;

(e) examines the design or structures of systems, interfaces, pages, and elements;

(f) employs visual and textual analysis, semiotic analysis, content analysis, or other methods of analysis to study the web and/or internet-facilitated images, writings, and media forms;

(g) studies large scale production, use, and regulation of the internet by governments, industries, corporations, and military forces.

(Markham & Buchanan, 2012, p. 3f)

Moreno et al. provided a simple classification and divided online methods into observational, interactive, and survey/interview research. They explained that observational research includes collection of public information where participants are not aware of the research, interactive research consists of cases where the researcher contacts the participants for permission to view their content, and survey/interview includes a method where researcher employs mailed surveys and consent forms to accomplish data collection (Moreno, Goniu, Moreno, & Diekema, 2013).

Kitchin (2007) distinguished between online research and web-based research (Kitchin, 2007, p. 16). She defined online research as “methodological practices that begin with a traditional research design and use cyberspace to find suitable samples or to facilitate a research initiative” (Kitchin, 2007, p. 15). Examples of online research include participatory observational ethnographies, focus groups, interviews, surveys, experimental research, breaching experiments, and action research. She defined web-based research as “methods that capture data—for the purpose of analysis—through web sources such as Internet sites, online discussion groups, news server, search engines, and blogs” (Kitchin, 2007, p. 13). She divided web-based research into non-intrusive and engaged categories. Non-intrusive techniques of data collection do not interrupt the naturally occurring state of the site or cyber community, nor interfere with it using premanufactured text. On the other hand, engaged techniques reach into the site or cyber community and involve engagement with the participants of the web source (Kitchin, 2007, p. 15).

Eysenbach and Wyatt categorized internet research into three groups: passive analysis, in which researchers do not involve themselves and just observe discussions, chat rooms, websites, or interactions; active analysis, in which researchers may or may not disclose their identity as researchers, but participate in communications to clarify or determine the accuracy of information; and when researchers identify themselves, seek participants and gather information in the form of online semi-structured interviews, online focus groups, online surveys, or use the internet to recruit subjects for traditional research (Eysenbach
& Wyatt, 2002, pp. 1-2). Mann and Stewart suggested four ways in which online methods can be used for qualitative data collection. These are standardized interviews in the form of email and web-based surveys, non-standardized forms of online one-to-one interviewing, “observation” of virtual communities, and the collection of personal documents online (Mann & Stewart, 2000, p. 66).

It is clear that there is no standard or uniform terminology, definition, or classification used for online research methods. While some authors divide online research methods on the basis of level of interaction among participants and researchers (Eysenback & Wyatt, 2002; Kitchin, 2007), others classify it based on the modes used to collect online data (Mann & Stewart, 2000; Moreno, 2013; Markham & Buchanan, 2012). This heterogeneity in the terminology and typology of online research methods makes it difficult for new researchers to find clear guidelines for ethical practice. Therefore, in the next section, I describe general principles that guide ethical practice, key ethical concerns, and possible solutions to mitigate these concerns.

3. Online Research Ethics: Key Concerns and Solutions

The general principles that guide ethical practice in online research are essentially the same as those that guide any research involving human beings and include respect for autonomy, justice, and beneficence (Kitchin, 2007). Autonomy refers to the notion that each individual has the right to privacy and dignity that should be protected at all times (Flicker, Haans, & Skinner, 2004). In other words, every participant should be able to make their own decisions to participate in research and the persons who are unable to make these decisions should be protected (Kitchin, 2007). Within the context of online research, it requires researchers to protect the personal information of internet users and refrain from disclosing anything that would allow their personal information to be inferred (Gelinas et al., 2017). The significance of the ethical principle of autonomy is emphasized in the Declaration of Helsinki (World Medical Association, 2017) and operationalized through the process of informed consent (Flicker, Haans, & Skinner, 2004).

The principle of justice refers to the notion that all research participants should be treated fairly, equally, and nobly during the entire course of the research process (Kitchin, 2007). In procedural terms, it requires that the researcher’s identity and the research methods are transparent, and that no segment of the community is unfairly burdened or faces discrimination. It also imposes obligations towards the individuals who are not able to protect their own interests, therefore should be protected from any exploitation for the sake of research and the progress of knowledge (World Medical Association, 2006). For example, while doing online recruitment, the principle of justice dictates that the investigators engaged in recruitment activities should be truthful, transparent, and honest when describing the aims, details, risks, and benefits of their studies (Gelinas et al., 2017).

The principle of beneficence requires researchers to evaluate all physical, social, psychological or medical harms or risks that their participants may face by virtue of being in the project, and make every possible attempt to minimize these harms and maximize
the benefits to them (Kitchin, 2007). Within the context of online research, the risk of harm arises when there is a disclosure of participant’s identity or any other sensitive information that may expose them to the risk of embarrassment, reputational damage, or legal prosecution (Townsend & Wallace, 2016).

Despite being guided by the same ethical principles, the evolving nature of internet-based strategies used for research necessitates researchers, ethicists, policy makers, and institutional review board members to address new ethical challenges that are emerging, and develop solutions in a context-dependent manner (Eynon, Fry, & Schroeder, 2011). The most common concerns that arise in the research studies employing online methods include ensuring anonymity, maintaining confidentiality, and obtaining informed consent in a virtual setting (Eynon, Fry, & Schroeder, 2011; Keller & Lee, 2010; Moreno et al., 2013).

Researchers suggest that, as the level of involvement and interaction increases, the level of risk might also increase (Kitchin, 2007; Warrell & Jacobsen, 2014). Additionally, whether the internet research comes into the ambit of human-based research determines if the proposal is subject to review by Human Research Ethics Boards, and the level of scrutiny required. According to the Canadian *Tri-Council Policy Statement: Ethical Conduct For Research Involving Humans* (Panel on Research Ethics, 2014), research that is non-intrusive, and does not involve direct interaction between the researcher and the participants through the internet, and the one that involves collection of online documents available in the public domain (such as images, videos, or other archival materials to which the public is given uncontrolled access, for which there is no expectation of privacy) are not considered under the remit of human subject research for the purpose of ethics review (Panel on Research Ethics, 2014; Warrell & Jacobsen, 2014). However, despite these explanations provided, there is an ongoing debate regarding what is considered public or private space on the internet and to what extent privacy should be expected or protected (Mackenzie, 2017; Warrell & Jacobsen, 2014). One major example of this controversy is the case of Facebook data mining inadvertently leading to a loss of anonymity for users (Zimmer, 2010). This case demonstrated that the nature of the internet makes anonymity difficult to protect despite researcher’s efforts to protect it.

Overall, it has been suggested that each type of online research method (i.e., observational, interactive, or survey/interview research) is highly contextual and involves different levels of engagement and interaction between the participant and the researcher, which has implications for ethics. Recently, Bender et al. (Bender, Cyr, Arbuckle, & Ferris, 2017) recommended a “privacy by design” framework for online recruitment. This framework recommends a proportionate approach to ethics assessment, which advocates for risk mitigation strategies that are proportional to the magnitude and probability of risks. However, who decides or what determines the “magnitude and probability” of risk is unclear. In addition, approaches to these issues vary internationally. Whether the Canadian guidelines are adequate and up-to-date given public perceptions and the new technologies and modes of research that are emerging, is an open question.

The *Personal Information Protection and Electronic Documents Act* of 2000 is Canada’s federal law that governs the collection, use and disclosure of personal information in a
manner that recognizes the right of privacy of individuals with respect to their personal information and the need of organizations to collect, use or disclose personal information for purposes that a reasonable person would consider appropriate in the circumstances (Office of the Privacy Commissioner of Canada, 2018).

Similarly, there are provincial legislations for each province in Canada. For example, the Personal Health Information Protection Act of 2004 is a provincial legislation of Ontario that was specifically established to govern the collection, use, and disclosure of personal health information, and also provide guidelines for researchers or research-based organizations that intend to collect, obtain, or use health information for scholarly or academic purposes. Thus, keeping these legislations in consideration, I now discuss elements that are of particular importance for each type of online research and provide some suggestions for good ethical practice.

3.1. Observational or Non-intrusive Web-Based Research

In this type of online research, participants are not aware of the research, except when information obtained is retrieved or recorded in such a manner which identifies the people, directly or indirectly, and/or puts them at risk of civic liability, or be damaging to their reputation (Moreno et al., 2013). This particular category applies when investigator observe social websites such as Facebook or LinkedIn, access public blogs, or download and/or analyze YouTube videos (Lafferty & Manca, 2015).

Some elements that are of particular importance for the researchers undertaking an online observational study are users’ privacy, the website privacy policy, and other legal considerations (Moreno et al., 2013). The users of social media platforms have the choice to make their posted content private or public, which in turn becomes the deciding factor whether the study is exempted from institutional ethics review or otherwise. In situations where content is public (i.e., does not need password or permissions for access), the researcher must determine whether the use of content (e.g., quotations from blogs) reveal the identity of the author/owner of the content. If this is the case, the researcher must seek informed consent, minimize the use of direct quotes/content, and consider a certain level of disguise (Eastham, 2011). Most of the social websites such as Twitter, Facebook, and LinkedIn have their privacy policy which is intended for their users. If researchers intend to use these websites, they should either certify compliance with the terms of use on the sites or alert the institutional ethics review board if their proposed techniques fail to comply with the relevant terms of use (or if it is ambiguous whether the techniques comply with the terms of use or not) (Gelinas et al., 2017). Some of the legal considerations relevant to online observational studies relate to the question of what constitutes a privacy violation (Moreno et al., 2013).

3.2. Interactive Research Involving Engagement With Participants

In this type of online research, researchers identify themselves and directly contact the participants, with an intent to access their online media content (such as social interactions) that is not publicly available, or interact with a person (e.g., through friending or following), or gain membership in a closed group (such as a chat room or
Facebook group) (Kitchin, 2007; Moreno et al., 2013). Some authors refer this type of research as virtual or visible ethnography (Hine, 2008; Kitchin, 2007). Hine (2008) and Kitchin (2007) suggests that if a direct communication is to be made with the participant, the study is to be treated as human research and subjected to institutional ethics review. Thus, the investigator is obligated to ensure anonymity of the subjects and gain informed consent without any coercion. The researcher should consider giving full right to the participants to withdraw from the study, either by the researcher leaving the social media group or by assuring the participants that their responses will not be included in the research (Hine, 2008; Kitchin, 2007). Online ethnographers also need to be sensitive to the privacy expectations of members of the online community, even when their activities are occurring online (Ess & Association of Internet Researchers, 2002). Hine (2008) further suggests that investigators not only need to protect the privacy of their participants, but also the privacy of the website/media (such as chatrooms or blogs) through which they draw their information/data. For example, the researcher should refrain from quoting verbatim, if this can render both the individual and the site readily identifiable (Hine, 2008). However, these are subjected to the expectation of the participants and discretion of the researcher whether they consider their participants as authors or “subjects” (Ess & Association of Internet Researchers, 2002). Warrell and Jacobsen (2014) suggest that, if participants are viewed as authors, the researcher may decide to give acknowledgment where it is due. If however, the participants are regarded as “subjects,” the researcher needs to guarantee their anonymity.

3.3. Surveys and Interviews

Surveys and interviews are common and well-developed methods of online research. In conducting surveys and interviews online, ethical concerns may arise with regard to obtaining informed consent, ensuring anonymity of the participants, and maintaining confidentiality of the data (Keller & Lee, 2010). The office of the Privacy Commissioner of Canada reminds researchers to adhere to the principles laid by the federal Personal Information Protection and Electronic Documents Act of 2000, and asks researchers to obtain “meaningful and knowledgeable” consent and provide individuals with sufficient details about their privacy practices (Office of the Privacy Commissioner of Canada, 2018). Consent is considered to be knowledgeable when individuals know the purpose of the information collected about them and their right to give, withhold, or withdraw the consent anytime they wish.

Enrolling participants for research involves some open and clear communication between the researcher and the potential participants, however this may be compromised in virtual settings (Flicker, Haans, & Skinner, 2004). Different ways through which an online consent can be obtained include emails or online statements that may require participants to agree to the terms of participation, thus implying consent (Warrell & Jacobsen, 2014). However, it is difficult to verify whether the participant has actually read the details carefully, whether there are any misunderstandings and issues in comprehension, and whether the person who is giving the consent is the authentic participant (Vehovar & Manfreda, 2008). Furthermore, verifying certain information such as age or mental capacity to give the consent can also be difficult (Keller & Lee, 2010).
Moreno et al. (2013) suggested that researchers should provide complete details of the study, including a readily available link that provides contact information, study aims, data collection procedure, potential benefits and harms, and steps taken to maintain anonymity and confidentiality of the participants. These steps ensure that the participants have opportunities to reach out to the investigators and clarify if they have any questions or concerns. Other ways to ensure participants read and understand the study details in online settings include use of multistage consent forms, enhancing readability of the document by reducing the amount of text, use of subheadings, or colors (Eynon, Fry, & Schroeder, 2011; Social Sciences and Humanities Research Ethics Special Working Committee [SSHWC], 2008). Flicker and colleagues propose that “concerns about identity fraud could be addressed by asking the same questions in multiple formats (e.g., both date of birth and age of users) and then checking for discrepancies” (Flicker, Haans, & Skinner, 2004, p. 128). Prior or pilot testing the consent forms and survey questionnaires are other important ways to minimize the risk of unintended consequences at later stages of the research (Eynon, Fry, & Schroeder, 2011).

3.3.1. Maintaining Anonymity or Confidentiality in Online Survey Research

Eynon et al. argued that though it is the prime responsibility of the researcher to ensure anonymity and confidentiality of the data which is collected and stored, “the extent to which a researcher should be concerned about confidentiality depends on the nature of the data being collected” (Eynon, Fry, & Schroeder, 2011, p. 8). They suggest that if the data are not contentious, or if privacy can be ensured, then this is less of a concern compared to controversial research topics or research where it is necessary to obtain personal information. In case of any sensitive information, researchers need to ensure that participants’ perceptions of anonymity are met, or if not, made explicit to the participant (pp. 8-9). The final report of the Advisory Panel on Online Public Opinion Survey Quality (Public Works and Government Services Canada, 2008) and SSHWC (2008) provide the code of conduct and list the steps that should be taken to ensure ethical practice:

(a) Survey respondents’ participation must be voluntary at all times. Misleading or deceptive statements should be avoided and respondents should be made well aware about the nature of the research and the purposes for which their data will be used.

(b) Personal information should not be sought from or about the respondents without their prior knowledge. Their anonymity should be maintained and the researcher must ensure that data are not used for subsequent non-research purposes. Personally identifiable data and other research related data should be kept and stored separately.

(c) Respondents should be informed about the duration of the survey and must be given links to data protection, privacy policy, cookie policy statements, or information about the use of a software at the beginning of the survey, which should be available in simple language.
(d) Any software used for the survey should not be installed on the respondent’s computer without their consent and even when consent is provided, respondents should be able to remove any such software easily.

(e) Respondents should be given entitlement to ask the researcher to delete their records and researcher should conform to such requests where reasonable.

(f) In case of surveys where an email list is acquired from some agency, the researcher should provide a clear statement of where the email addresses came from and ensure that the individuals listed have consented before to be contacted in future for any research purposes.

3.3.2. Data Protection in Online Survey Research

Personal Information Protection and Electronic Documents Act of 2004 lays out the principles under which any personal information should be collected and secured. The act suggests that appropriate administrative, physical, and technical safeguards should be put in place to protect the personal information against theft, loss, and unauthorized collection, use, copying, modification, or disclosure. The Advisory Panel on Online Public Opinion Survey Quality (Public Works and Government Services Canada, 2008) along with other researchers and committees (Eynon, Fry, & Schroeder, 2011; Eynon, Schroeder, & Fry, 2009; SSHWC, 2008) provide certain measures that can be undertaken by the researchers to ensure integrity, security, and confidentiality of the information collected through online surveys. These measures include the following:

(a) Personal data collected or stored on the websites or servers should be protected using measures such as data encryption or Secure Socket Layer (SSL). For ensuring physical protection of the data, researchers should keep the data in password protected computer directories, use data labels that are meaningless to anyone but the researcher, and code the data in a way that reduces the possibility of people being able to trace the data to a specific individual.

(b) In case of temporary storage of data on the server, the researcher is obliged to take necessary precautions to prevent unauthorized access to the data from the server or during data transfer, and ensure that the temporary storage is terminated at the earliest.

(c) In their privacy policy or statement, researchers should always state the use of cookies or other log files and inform the respondent that they can disable use of cookies in their computer by using their own system settings.

(d) When data are collected through the server that does not belong to Canada, researchers must ensure that compliance to Canadian privacy laws can still be ensured.

(e) Researchers must observe all relevant laws and national codes, specifically if their projects involve participants from various parts of the world. Also, they need to ensure that data transfer is permissible and is safeguarded against any breaches to privacy.
(f) In case where emails are sent in batches for the purpose of recruitment or inviting the participants for surveys, investigators need to ensure that the emails of the respondents are not revealed. If possible, the use of emails should be minimized.

(g) In case of any breach to data safety or anonymity, the participants should be informed immediately with all the details, so that they can decide on actions to be taken.

In Box 1, I summarize good ethical practices researchers can adopt for internet-based research.

<table>
<thead>
<tr>
<th>Box 1. Ethical Practices for Internet-Based Research</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Ensure transparency while recruitment, especially while recruiting using an acquired email list or through closed or open social media groups.</td>
</tr>
<tr>
<td>(b) Consider participants’ expectations, perceptions, and awareness about privacy.</td>
</tr>
<tr>
<td>(c) Ensure compliance to national data protection laws and the applicable ethical codes.</td>
</tr>
<tr>
<td>(d) Use a password identity system to ensure authentic participants.</td>
</tr>
<tr>
<td>(e) Ensure that consent is informed and knowledgeable.</td>
</tr>
<tr>
<td>(f) Ensure that a privacy statement is provided which explicitly mentions purposes for which the data will be used, participants’ rights, and strategies adopted by the research team for maintaining anonymity, confidentiality, and data protection.</td>
</tr>
<tr>
<td>(g) Assign each participant a unique identification number/code to track progress and maintain security.</td>
</tr>
<tr>
<td>(h) Provide opportunities to participants to reach the research team or clarify any concerns.</td>
</tr>
<tr>
<td>(i) Honor participants’ request to delete their records.</td>
</tr>
<tr>
<td>(j) Ensure encryption and Secured Socket Layer protection systems.</td>
</tr>
<tr>
<td>(k) Use password protected computer directory to store the data.</td>
</tr>
<tr>
<td>(l) Remove all identifiers before data analysis.</td>
</tr>
<tr>
<td>(m) Ensure no unauthorized person is able to access the data.</td>
</tr>
<tr>
<td>(n) Ensure that data are not used for subsequent non-research purposes.</td>
</tr>
</tbody>
</table>

4. Conclusion

While online research methods have several advantages, the ethical challenges associated with them are many. In addition, lack of homogeneity in the terminology and typology of online research methods often leads to confusion and makes it difficult for new
researchers to find uniform guidelines to practice these methods. While the guiding ethical principles applicable to online research are the same as those applicable to any research involving human participants, application of these principles can be different depending on the type of the online method and its purpose. Although there are a few grey areas, detailed appraisal of the scholarly literature suggests that maintaining anonymity, confidentiality, transparency, and ensuring data security are of particular importance. Breaches to these can have substantial consequences, especially when it involves collection of personal and sensitive information. Paying attention to the participants’ expectations, perceptions, and awareness about privacy is one of the key elements. With careful design, planning, and implementation, researchers can deal with most of these challenges while ensuring that the participants’ privacy rights are protected and standards for ethical practices are met.

References


Received 8 November 2017 | Accepted 24 January 2018 | Published 29 January 2018

Copyright © 2017 *Journal of Research Practice* and the author