

Human Subjects Research Involving the Internet

Introduction

This policy describes general guidance to investigators in planning their research with human participants using Internet-based research methods. Research involving human participants falls under the purview of the NMSU Office of Compliance (OC) through the Institutional Review Board (IRB). The broad and overarching term "Internet-based research" includes both the internet as a *tool for data collection and management as well as the internet as a locale or venue of research*. For example, data collection that takes place over the internet may include methods such as email, web surveys, public listservs, social networking sites such as Facebook®, blogs, chat rooms, electronic bulletin boards, gaming, and similar technologies. IRB approval or Certification of Exemption from IRB review is necessary whenever conducting research involving human participants, regardless of the method of data collection.

What are examples of Internet-based research?

There are multiple forms of Internet-based research. A wide range of Internet based-research where human subjects may be involved include:

- Research studying information that is already on or via the internet without direct interaction with human subjects (harvesting, mining, profiling, scraping -- observation or recording of otherwise-existing data sets, chat room interactions, blogs, social media postings, etc.)
- Research that uses the internet as a vehicle for recruitment or interacting, directly or indirectly, with subjects or through Internet locales or tools, for example, social media, push technologies (e.g., Self-testing websites, survey tools, Amazon Mechanical Turk®, etc.)
- Research about the internet itself (e.g., use patterns, search engines, email, etc.; evolution of privacy issues; etc.)
- Research about Internet users -- what they do, and how the internet affects individuals and their behaviors (effects of social media on people; information contagion, etc.)
- Research that utilizes the internet as an interventional tool, for example, health interventions and their resulting health behaviors
- Others (e.g., emerging and cross-platform types of research and methods, including m-research (mobile))

IRB Requirements

Research involving the collection of data about people through the use of the internet involves many of the same considerations as any other research with human participants. As such, the NMSU IRB will review the use of the internet for research activities under its jurisdiction to ensure that:

- Risks such as violation of privacy, legal risks, and psychosocial stress are minimized;
- Participation is voluntary;
- Informed consent requirements are met; and
- Information obtained from or about human participants is kept confidential

These guidelines for human subject protection are discussed below within each type of Internet-based research medium.

Common Internet-Based Research Concerns

Terms and conditions of use. Researchers should follow the Terms of Agreement (TOA) or Terms of Service (TOS) for any site where data is collected. TOA/TOS outlines the rules a person or organization must observe in order to use a service. For example, some blogs or sites on health issues may require the permission of those who administer the site. Neither the NMSU Office of Compliance nor the IRB can take responsibility for ensuring that the terms and conditions for conducting research on such sites have been met. Failure to ascertain and acquire appropriate permissions could result in consequences that may include sequestration or loss of the data collected, reputational harm to the researcher and the institution, and in the worst case, legal action by the site manager or participants.

Child and Vulnerable Subject Protections. Researchers working with children online are also subject to the Children's Online Privacy Protection Act (COPPA). COPPA prohibits researchers from collecting personal information from a child without appropriate COPPA notifications about how the information will be used and without getting verifiable parental consent. Researchers must ensure safeguards are in place for screening out children, prisoners, and other vulnerable populations unless these populations are the intended participants of their study. *COPPA*: Operators of commercial websites and online services directed toward children under 13 years of age that collect personal information from these children must comply with the Children's Online Privacy Protection Act (COPPA). The goal of COPPA is to protect children's privacy and safety online, in recognition of the easy access that children often have to the web. COPPA requires website operators to post a privacy policy on their website and create a mechanism by which parents can control what information is collected from their children and how such information may be used.

For more information, please see: <http://www.ftc.gov/privacy/privacyinitiatives/childrens.html>.

The IRB must also consider all additional requirements for the approval of research that involves any vulnerable population in the same manner required for non-Internet-based research. As there is no standard for identifying distressed participants online, the IRB must take into consideration potential participant experiences (the sensitive nature of the research) that may be distressing to participants. Such experiences should be considered when evaluating risk/benefit.

Security of Online Transmissions. Collecting data over the internet can increase potential risks to confidentiality due to third-party sites, including the risk of third-party interception when transmitting data across a network and the impossibility of ensuring that internet data are completely deleted or destroyed once the work is complete. All data must be protected as it moves along the communication pathways (e.g., from the participant to the data-collection server, from the server to the investigator). Additionally, all databases storing identifiable information or data must be protected regardless of the source creating the data (e.g., encryption of the database, de-identifying the data). Researchers conducting Internet-based research should inform participants that the security of online transmissions may not be guaranteed. The level of security should be appropriate to the risk. Research involving sensitive topics may require additional protections such as certified digital signatures for informed consent, encryption of data transmission, or technical separation of identifiers and data. Participants should be informed of these potential risks in the informed consent document. For example:

- "Although every reasonable effort has been taken, confidentiality during actual Internet communication procedures cannot be guaranteed."

- "Your confidentiality will be kept to the degree permitted by the technology being used. No guarantees can be made regarding the interception of data sent via the internet by any third parties."
- "Data may exist on backups or server logs beyond the timeframe of this research project."

Informed Consent. The IRB must evaluate the appropriateness of the informed consent process. For Exempt, anonymous Internet-based surveys, it is often appropriate to provide participants with informed consent information and provide, at a minimum, an informational cover letter at the beginning of a survey letting participants know they have a right not to participate in the survey and to drop out of the survey at any time. Ideally, Internet-based surveys are highly advised to include "I agree" or "I do not agree" buttons with which participants would explicitly indicate their active choice of whether or not they consent to participate. In those cases where research is designed to capture private information, such as on Facebook®, a letter of consent for each individual participant should be provided and collected.

If the NMSU IRB determines that signed informed consent is required, the researcher may email the consent form to participants, who may then type their name and the date into the spaces provided on the consent form, and return it to the researcher via email or they may be able to enter this information into a web form.

Concerns Regarding Participant Recruitment through the Internet

The IRB must review and approve all materials used for recruiting participants on the internet. Internet-based procedures for advertising and recruiting potential study participants (e.g., internet advertising, email solicitation, forum post, banner ads) must follow the IRB guidelines for recruitment that apply to any traditional media, such as newspapers and bulletin boards. [Note: here we would like to refer to NMSU IRB recruitment policy] We recommend work on this policy and/or have it on a page that can be easily accessed.

If researchers are providing an incentive, participants should be able to receive the appropriate compensation without revealing his/her identity (e.g., gift certificates from online retailers provided by displaying the unique certificate redemption number to respondents at the completion of a questionnaire) and without being compelled to complete all parts of the study. Participants should be able to withdraw from the study without any penalty.

Concerns Regarding Participant Incentives

Internet research and platforms supporting it create unique concerns regarding participant reimbursement or study incentivization. At the time of this writing (Spring 2021), popular platforms include Amazon's Mechanical Turk (M-Turk) and Qualtrics Panels. "Amazon Mechanical Turk (MTurk) is a crowdsourcing marketplace that makes it easier for individuals and businesses to outsource their processes and jobs to a distributed workforce who can perform these tasks virtually" (www.mturk.com). Such jobs may include participating in research studies. Qualtrics' Panels also provides access to survey populations. A web search will show companies providing similar services.

Rates of participant reimbursement vary dramatically across platforms. For M-Turk, one study found that average earnings amounted to about \$2.00 per hour and that only 4% of all workers earned more than the federal minimum wage (Hara et al., 2018). In contrast, Qualtrics panels reimburse participant time at nearly \$20/hour (R. Palacios, Personal Communication), substantially above minimum wage.

The primary ethical issue is that too little compensation may be exploitive, too much compensation may be coercive, or tempt repeated participation (e.g., ballot box stuffing). When NMSU researchers plan to provide payment or incentives, they must explain *why the compensation is necessary and whether it is reasonable in relation to the experiences of and/or burden on the human subjects*. In this context, the amount what "others" (i.e., previous researchers) have paid is insufficient justification. An alternative rationale is to reimburse people for their time on an hourly basis and according to a standard hourly rate, such as the federal minimum wage (\$7.25 in 2021). Thus, an M-Turk study taking 30 minutes to complete might be reimbursed \$3.75. When calculating the time to complete a study, researchers should include the time needed to review any consent form or information.

Finally, researchers must ensure that payments and partial/prorated payments are made according to the IRB protocol. In the end, the NMSU IRB will evaluate proposed incentives with these issues in mind and on a case-by-case basis. The Board reserves the right to recommend fair/just compensation levels

Concerns Specific to Each Internet Research Category

1) Social Media (Facebook®, Twitter®, etc.)

Internet-based research using social media that is not public or that maintains an expectation of privacy cannot be submitted at the exempt level. Investigators will often be met with hostility if they are not sensitive to the online community's expectations of privacy. Participants of an online community may see the presence of a researcher as intrusive.

In order to determine the privacy expectations of a social media platform, it is important that investigators be familiar with the online space in which they intend to conduct research, especially its Terms and Conditions of Use. If no Terms and Conditions of Use exist, investigators need to determine whether participants have an expectation of privacy, and request the appropriate permissions and consent to conduct the intended research. In spaces that are not considered public, researchers have the responsibility to protect the privacy and rights of participants. If an investigator has prior experience in an online community and is already known to its participants, the researcher may have a better chance of being welcomed into the space.

Researchers may seek to get information not only about and from the individuals specifically recruited for the study, but also about individuals connected to the recruited participant's social network (e.g., his/her "friends") by accessing the information that those individuals have made available to the recruited participant. In this circumstance, the participant population now includes the "friends" who may need to consent before data about them can be included in the study. Information made available by "friends" on the "wall" or another public place on the recruited participant's social network site may

be considered to belong to the participant and can be included without the explicit consent of the "friend," if the study itself is considered to be no more than minimal risk. Researchers must exercise caution to protect the identity of such participants and report results in an anonymous and aggregate form as much as possible.

2) Virtual identities, personas

Online identities (personas or avatars) and their corresponding character names established in online communities should be treated just like real persons. These personas and their reputations can usually be traced back to real individuals. If a researcher wishes to use names of internet personas, or real names in publications, it is normally sufficient to obtain consent from the human controller or to recognize consent from the avatar as a proxy for the controller, although in some cases consenting both the virtual persona and the human controller may be more appropriate.

3) Web-based Surveys

Survey research is one of the most common forms of internet-based research. Consent procedures are discussed above. Researchers must format survey instruments in a way that will allow participants to refuse to answer specific questions. For example, the list of responses may include an option such as "Decline to answer." Online surveys may also be structured so that a participant can skip a question and proceed to answer subsequent questions. In addition, participants must always be given the option to withdraw from a study, even while in the middle of a survey, and still receive compensation if it has been offered. Use of Qualtrics®, RedCAP®, SurveyMonkey.com®, Psychsurveys.org®, Amazon Mechanical Turk®, and other online survey tools is permitted for most minimal risk studies employing online survey procedures. Investigators should review confidentiality measures and data security policies for the given online survey tool and make sure that they are described in the protocol. If security measures are not in line with what the NMSU IRB/OC requires, the use of the given survey company may not be approved. Research participants also need to be informed of data security measures.

4) Interactive communication (emails, Twitter®, Instagram®, gaming)

When navigating in a chatroom and other interactive communication online environments, it is important that those present are able to let the researcher know if they are not comfortable with the researcher's presence and that the researcher respects these wishes. Because access to chatrooms can prove difficult for investigators, and chatroom participants are not always eager to have a researcher in their midst, one suggested technique is for investigators to create their own chatrooms just for research purposes. Investigators can greet individuals joining the chatroom with a message informing them about the study and asking them for their informed consent. This is a good way to be sure that all participants are fully aware of the research and have consented to participate.

5) Existing Data and Online Databases

Research utilizing data that is both existing and public is not considered human subjects research and does not require NMSU IRB/OC review (<http://cphs.berkeley.edu/secondarydata.pdf>). Data only accessible through special permission or registration/login (with username and password) are generally not considered public. When determining whether data are public, the investigator must decide if there exists an expectation of privacy. If it is determined that the data were not intended for public use, even if the data are technically available to the public, the data should be considered private. For example, data available on WikiLeaks® were technically public but included information about individuals who did not authorize the release of such data (<http://en.wikipedia.org/wiki/WikiLeaks>). Researchers accessing

data that contain identifiers and are not publicly available must obtain NMSU IRB/OC review and approval.

6) Observational Research

When online research procedures are employed, the investigator must be sensitive to the definition of public behavior. Despite navigating in a public space, an individual may have an expectation of privacy, and investigators need to be sensitive to that expectation. For example, an investigator wishes to collect data from discussions posted in an online community support group for substance abusers. The online community is technically public in that anyone can view the discussions and join the group, but some group participants are there to provide personal experiences and gain support regarding substance abuse and may believe that all discussions and personally identifiable information will remain private. Researchers should inform participants that "observation" is taking place and that any information exchanged may be used for research purposes when observing a web-based interaction, such as a chat room, that is not open to the public. They should also be reminded that such information will not identify any participants but is being collected to produce aggregate/summary findings.

7) Interviews

Conducting interviews online allows researchers to gather information from respondents who would have been difficult to contact otherwise, such as a very geographically dispersed population. Interviews may be conducted over the internet using email or chat technology such as Google Chat®, AOL instant messenger®, Yahoo! Messenger®, etc. When conversing with a research participant via online chat, investigators should take into account the inability to read visual and auditory cues, which can lead to possible misinterpretation of both questions and responses. Researchers should be careful not to misrepresent findings collected only through text or auditory information.

8) Use of mobile devices and other emerging technology

This type of research conducted from mobile devices such as smartphones may involve the use of existing data and/or interaction with or intervention in the person's environment. Additional considerations apply to research that involves the collection of data via social media applications that are networked with mobile devices, or by installing applications on a person's mobile device to collect data:

- Researchers must not collect location information or other data that is not explicitly described or outlined in the consent form.
- If the research involves installing a mobile application (app) on a person's smartphone or other device for the purposes of data collection, the researcher must describe how the app will be deactivated at the conclusion of the study. This should be done either by making the deactivation part of the study's exit procedures, or by providing instructions to study participants on how to deactivate the app. Additionally, researchers should describe plans to ensure they do not continue to collect data once the study is complete, in case a participant does not effectively deactivate the app.
- If the study involves the use of a mobile device provided by the researcher, the researcher should explain the confidentiality safeguards that are in place (e.g., how s/he will ensure the data is under the research team's control and that third parties do not have access to it), as appropriate to the study.

Regulations

- [45 CFR 46.111](#)
- [Children's Online Privacy Protection Act \(COPPA\)](#)

Glossary of Terms

Blog: A website used as a journal; can be personal or professional in nature.

Chatroom: An online location where individuals can come together to have text-based chat discussions that occur in real time.

Cloud computing: Distant storage or data management servers typically owned and operated by a third party.

Confidentiality: Pertains to the treatment of information that an individual has disclosed in a relationship of trust, and with the expectation that it will not be divulged without permission to others in ways that are inconsistent with the understanding of the original disclosure.

Cookie: A text file placed on user's computer by a website or web server. Often used to keep track of individuals as they navigate a site, and more broadly, the web.

Encryption software: A piece of software that is used to obfuscate information to all of those who do not have the means to decrypt the information.

Gaming: Interactive game worlds

Human Subject: a living individual about whom an investigator (whether professional or student) conducting research obtains (1) data through intervention or interaction with the individual, or (2) identifiable private information.

Interaction: Any communication or interpersonal contact between investigator and subject with any medium.

Internet Protocol (IP) address: A numeric address assigned to every computer that connects to a network, or more commonly, the internet.

Internet Relay Chat (IRC): A protocol used for hosting and participating in chat rooms.

Intervention: (1) Physical procedures by which data are gathered (for example, surveys, focus groups, experiments, venipuncture etc.) and (2) manipulations of the subject or the subject's environment performed for research purposes.

Lurking: A behavior specific to online communities, wherein an individual remains silent, observes, and does not participate in the community.

Online persona: An online character or avatar used by an individual.

Online survey: Any tool used to collect responses to survey questions via the internet.

Private information: Information about behavior that occurs in a context in which an individual can reasonably expect that observation or recording is not taking place, or information which an individual has provided for specific purposes and which the individual can reasonably expect will not be made public. This information may be clearly private (a medical record or personal diary), but may also include a person's Facebook profile that is set so only friends can see messages or photographs. In order for obtaining the information to constitute research involving human subjects, private information must be individually identifiable (i.e., the identity of the subject is or may readily be ascertained by the investigator or associated with the information).

Publicly available: The general public can obtain the data and they are readily available to anyone (without special permission/application) regardless of occupation, purpose, or affiliation.

Secure website: A site that conforms to best current security practices, such as use of the Secure HyperText Transfer Protocol (https), application of relevant patches, and sound auditing and certificate management.

Social media/network services: Web and mobile device-based services that provide a collection of ways for users to interact, such as social networking sites, blogs, discussion groups, or other information sharing or communication services that support messaging, email, video, posting comments, etc.

Virtual community: A group of individuals networked together through association with a virtual environment, homepage, or other Internet medium (e.g., SecondLife®).

References

Use of Social networking sites or mobile devices for human participant research. Retrieved from <https://irb.cornell.edu/documents/IRB%20Policy%202020.pdf>

Research Involving the secondary use of existing data. Retrieved from <http://cphs.berkeley.edu/secondarydata.pdf>

Considerations and Recommendations Concerning Internet Research and Human Subjects Research Regulations. Retrieved from http://www.hhs.gov/ohrp/sachrp/mtgings/2013%20March%20Mtg/internet_research.pdf

Hara, K., Adams, A., Milland, K., Savage, S., Callison-Burch, C., and Bigham, J. (2018). A Data-Driven Analysis of Workers' Earnings on Amazon Mechanical Turk. *CHI '18: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, April 2018, Paper No.: 449, Pages 1–14. <https://doi.org/10.1145/3173574.3174023>

Internet based research. Retrieved from http://cphs.berkeley.edu/internet_research.pdf

UCLA Guidance on Research Involving the Internet. Retrieved from
http://ora.research.ucla.edu/OHRPP/Documents/Policy/8/Internet_Research.pdf