Purpose:

This is not meant to be a governing document and, as with all research involving human subjects (hereafter referred to as HSR), the Indiana University Institutional Review Board (IRB) has independent authority to approve or disapprove HSR. For questions regarding specific protocols, please contact IU Human Subjects Research Office in Indianapolis at (317) 274-8289 or in Bloomington at (812) 856-4242.

I. INTRODUCTION

The use of the internet for research is becoming more and more prevalent. It provides a quick way to gain access to a large number of potential subjects and information without expending too many resources. While this is good news for investigators, this raises many unresolved questions concerning recruitment, privacy, confidentiality, and informed consent. The current HSR regulations do not address many of the unique issues associated with internet-based HSR. This guidance is intended to shed light on a variety of issues surrounding research involving the internet, social media, and mobile devices and to help investigators design studies that are in line with the currently regulatory and ethical landscape.

The broad and overarching term “internet research” includes both the internet as a tool for research and the internet as a locale or venue of research. For example, research employing survey instruments, search engines, databases, or databanks would constitute using the Internet as a tool for research. Such research may not involve direct interaction with human subjects, but identifiers or personally identifiable information may be generated, collected, and/or analyzed. In contrast, using the Internet as a medium or locale of research entails qualitative or quantitative studies of various Internet spaces, such as chat rooms, virtual environments, or social media sites. Some of the most common uses of the internet for research:

- Research studying information that is already available 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 recruiting or interacting, directly or indirectly with subjects (self-testing websites, survey tools, Amazon Mechanical Turk®, etc.).
- Research about the Internet itself (use patterns or effects of social media, search engines, email, etc.; evolution of privacy issues; information contagion, etc.).
- Research about Internet users – what they do, and how the Internet affects individuals and their behaviors.
- Research that utilizes the Internet as an interventional tool, e.g. interventions that influence subjects’ behavior.
- Others (emerging and cross-platform types of research and methods, including m-research [mobile]).
- Recruitment in or through Internet locales or tools, for example social media, push technologies.
II. DEFINITIONS

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

b. **Data Mining**: Research involving observation and reporting of online behavior and/or sorting through data to identify patterns and establish relationships.

c. **Identifiable Private Information**: For purposes of this guidance and as defined in the regulations, the identity of the subject is or may readily be ascertained by the investigator or associated with the information.

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

e. **Private Information**: For purposes of this guidance and as defined in the regulations, information about behavior that occurs in a context in which an individual can reasonably expect that no observation or recording is taking place, and information which has been provided for specific purposes by an individual and which the individual can reasonably expect will not be made public.

III. CONFIDENTIALITY OR ANONYMITY

a. Researchers conducting web-based research should be careful not to make guarantees of confidentiality or anonymity, as the security of online transmissions differs depending on the mode. Investigators need to address how they intend to assure confidentiality of the data collected, keeping in mind that the degree of concern and protection is directly related to the sensitivity of the data.

- Data transmitted via e-mail cannot be anonymous without employing additional steps.
- Data transmitted over the web can only be anonymous if software is used to store the information directly in a database without identifiers; otherwise, identifiers are attached to the data. Web servers automatically already store information about visitors to a web site and that information can be accessed by others.

IV. PUBLIC VS PRIVATE

A fundamental question is whether the Internet should be considered a public or private space. Members of Internet communities do not have the expectation that they will be research participants.

V. DATA COLLECTION

Internet-based data collection methods can range from the use of existing data and observations to interventions and survey/interview procedures. Each is discussed below.

a. **Existing Data**
Guidance for Research Involving Human Subjects

Research utilizing data that are both existing and public is not considered human subjects research and does not require IRB review. Data only accessible through special permission are generally not considered public. However, if steps are required to access data (e.g., registration/login, payment, etc.) but access is not restricted beyond these steps (e.g., anyone who creates a username and password can access the data) the data may qualify as publicly available. When determining whether or not data are public, the investigator must decide if there exists an expectation of privacy. If it’s 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. Researchers accessing data which are identifiable or that when combined could readily be identifiable must obtain IRB review and approval.

b. Observations

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 support regarding substance abuse and may believe that all discussions and personally identifiable information will remain private.

c. Chatrooms

When navigating in a chatroom, 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.

d. Surveys

Survey research is one of the most common forms of internet-based research. Researchers are advised to format survey instruments in a way that will allow participants to refuse to answer specific questions. For example, the list of responses can include an option such as, “Decline to answer.” In addition, participants must always be given the option to withdraw from a study, even while in the middle of a survey.

Use of SurveyMonkey.com, Mechanical Turk, etc. 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 IRB requires, use of the given survey company may not be approved. Research participants also need to be informed of data security measures.
Guidance for Research Involving Human Subjects

- Researchers should have a familiarity with the survey software being used, the types of information being collected, e.g. IP address, email address, the options the survey software provides regarding what information to collect, the ways in which information will be stored, and how any identifying information will be de-linked from survey data.

e. 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, Skype, 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. Voice intonation and facial expressions are often used to convey meaning. Thus, investigators may need to ask clarifying questions in order to accurately interpret responses, and provide additional information in order to be sure that participants understand the questions.

VI. Social Media as a Recruitment Tool

Investigators are using social media and other internet forums to identify and contact potential subjects for participation in all kinds of research, from surveys to clinical trials. Since subjects self-identify with social media and internet sites and choose which sites and forums to connect with, investigators can target subjects with specific interests, making recruitment efforts more effective and efficient. Internet recruitment is conducted via many different forums and methods, including electronic flyers posted on social media sites, social media groups, websites devoted to clinical trials, and even blog posts and discussion groups/boards. Regardless of the specific recruitment method, investigators should consider the following guidelines when using internet forums for recruitment purposes.

a. Advertising content

- Recruitment materials available on the internet should follow the same guidelines applicable to traditional recruitment methods.
- Recruitment is the first step in the informed consent process; as such, all materials presented to potential subjects must be reviewed and approved by the IRB.
- Recruitment materials should include investigator contact information, information about the purpose of the study, any eligibility criteria, benefits to the subject, and time commitment required for participation in the study.
- Materials should never promise free medical treatment, imply unanticipated benefits, or emphasize payment.
- Websites and recruitment materials for clinical trials and FDA-regulated research may not make claims which are inconsistent with approved FDA labeling, must indicate when drugs and/or devices are investigational, and may not offer post-approval discounts on drug/device costs in return for participation in the study.

b. Initial contact of potential subjects
Despite the public nature of social media and the internet, investigators should carefully consider ethical recruitment practices and subjects' privacy when determining how subjects will be contacted.

- Contact methods must be clearly described in IRB documentation.
- "Cold contact" of individual potential subjects via social media is strongly discouraged. The IRBs would generally expect that initial contact with a potential subject be conducted by someone whom the potential subject would recognize, either because the individual has friended or followed the study team or has otherwise agreed to be contacted, or because the recruiting individual is a part of the potential subject's care team.

c. **Sharing of potential subject data**

- The recruitment process logistically necessitates the sharing of data between the potential subject and the study team. During initial IRB review, investigators must describe provisions for protecting confidentiality of potential subjects and subjects. Such provisions are especially critical when recruitment is conducted via the Web.
- Investigators should never request that potential subjects provide identifiable information via public forums, including tweeting, private messaging, posting, etc.
- Potential subjects interested in participation should be directed to contact the study team via non-public means.
- Health information should never be shared via email or social media contact methods.

VII. **USING SOCIAL MEDIA FOR SUBJECT CONTACT AND FOLLOW-UP**

Internet forums may be an effective way to stay in contact with subjects after they've agreed to participate in the study; however, subjects should be informed of the possibility of electronic contact.

a. The informed consent process should clearly describe any anticipated contact via social media or other electronic forum.

b. Subjects should provide consent to be contacted via social media.

c. If they agree, subjects should provide their Facebook profile names, Twitter handles, etc, to the study team for contact purposes.

d. If contact via social media and other electronic means is anticipated, the HIPAA authorization should clearly request authorization for such contact.

e. If subjects do not prospectively consent to contact via social media and other internet forums, the study team may request permission to do so from the IRB by requesting a waiver of informed consent for this purpose and justifying it appropriately. Such a request should include:

- Plan for ensuring that the study team identifies the correct individuals
- Copy of the planned contact language. Language must not imply or share potential confidential information, including protected health information.
VIII. RESEARCH WITH MOBILE DEVICES/APPS

Many research projects utilize mobile apps, either as a tool for collecting research data, or as the object of the research. Some mobile apps are regulated by the FDA as medical devices. In general, if a mobile app performs the same function as a medical device (i.e. intended for diagnosis of disease or other conditions, or the cure, mitigation, treatment, or prevention of disease), it may be subject to FDA regulations and requirements. IRB documentation should clearly describe the purpose of any mobile apps in order to assist HSO staff in determining whether FDA regulations may apply.

a. Secondary Subjects

Mobile device studies are particularly likely to capture information about persons who are not the intended (and consented) subject of the research.

- Researchers should prepare themselves for this possibility and take steps to limit the amount of information gathered.
- Can subjects turn the device off to avoid recording information in inappropriate locations?
- Can subjects effectively advise other people of their participation?
- Does the study require the subject to download an app? Was the app created specifically for this study, or was it created by a third party?
- For third party apps, the researcher needs to be familiar with the terms and conditions and direct subjects to review it. The researcher should also know what information the app sends to the app developer.
- Does the study rely on the subject’s data plan to transmit data?
  - The research should advise the subject that participation could lead to increased costs, and if possible provide an estimate for data use.

IX. CONSENT ISSUES:

a. As with research conducted in more traditional settings, researchers employing web-based surveys still have a responsibility to inform prospective subjects about the research and any potential risks associated with their participation. For most web-based surveys, it is not practical to obtain signed consent from subjects. As such, for non-exempt studies, the researcher should request from the IRB a waiver of documentation of informed consent.

b. Many times it is appropriate to simply state in the information provided to potential subjects that completing and submitting a survey, or completing a research task, consent is implied. Other times, researchers may wish to include an “I agree” button that must be clicked before proceeding to the research procedures.
c. Research may be taking part via third party sites that retain data provided by subjects, which should be noted as appropriate.

d. A written consent form may not be the best way to ensure that subjects are informed about the study. While still needed so that subjects can save or print a form for their own records, the consent process may involve videos, gifs, etc. that more effectively convey the needed information.

X. CHILD ISSUES:

a. Obtaining parental consent for online research with minors will rarely be practical. Even if a researcher attempts to obtain parental consent, it is likely not verifiable.

- A waiver of parental consent will likely be needed for any internet research involving children, however, depending on the nature of the research, the researcher should suggest that potential child subjects inform their parents of the research and their involvement, as appropriate.

b. Contacting adolescents for research purposes is different than contacting younger children. Adolescents are considerably more likely to already have an online presence and have much greater freedom and decision making capacity than younger children.

- The minimum age/grade level for inclusion in online/mobile device research (without parental consent) is thirteen (13), which is in compliance with COPPA.

c. Researchers should consider the possibility that a parent may discover that their child is taking part in a research study for which he/she was not previously notified. This could lead to increased risk for child participants, as their confidentiality could be more at risk and they could be subject to punishment for participating without their parents’ knowledge. The researcher and University could also be put at risk.

- Researchers should note these risks and take appropriate measures to protect against them.

d. COPPA: Operators of commercial websites and online services directed towards 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](http://www.ftc.gov/privacy/privacyinitiatives/childrens.html).

e. Research conducted online meant to exclude children should consider including a statement in the study information that the research is intended only for individuals 18 years of age and older.

XI. INTERNATIONAL ISSUES:
Guidance for Research Involving Human Subjects

a. As with all international research, activities and behavior that are acceptable/legal in the United States may not be in other countries.

- Researchers should carefully target their intended audience.
- Research on topics likely to be problematic (e.g. sexual behavior in Saudi Arabia, attitudes toward government censorship in China) should include information regarding said risks and plans to protect against them.

a. Because the internet reaches across borders, and it can be difficult, if not impossible, to exclude individuals from outside of the United States from participating in research conducted online, researchers should consider the inclusion of a statement in the study information that the research is intended for US subjects only.

XII. GENERAL ISSUES:

a. Many websites used to recruit research subjects or conduct research procedures include terms and conditions to which members and/or visitors of the website are bound. Researchers should familiarize themselves with these terms and conditions before recruiting from or conducting research on them to ensure subjects are not asked to violate the service agreement.

b. Some web sites may record research activities/communications.

c. Recruitment via social media or other sites may unintentionally reveal information about the subject to a third party unless ads are specifically targeting individuals with traits already disclosed (e.g. outing a Facebook user).

d. Researchers should consider informing subjects that they should be aware of the terms and conditions of the website to understand what, if any, data may be used/maintained by the website itself.

XIII. DATA COLLECTION AND STORAGE:

a. Data should not be stored on third party servers*, regardless of the type of research.
   *exception for collaborating institutions, funding agencies

b. Links to IU Knowledge Base regarding data storage should be provided. (I see no reason for us to try to explain these when UITS already has, and has done a better job than we could.)

Additional Resources:

https://protect.iu.edu/cybersecurity/data/surveys