Guidance for Data Security, Storage, and Management

The Common Rule (45 CFR 46, Subpart A) states that Institutional Review Boards (IRBs) must determine that there are adequate provisions for protecting the privacy of subjects and maintaining the confidentiality of data. As such, researchers are required to provide a data security and storage plan to protect the confidentiality of research data. This guidance helps researchers plan for collecting and maintaining human subjects data in a secure manner in compliance with BSU IRB requirements and federal regulations.

1. Definitions

Anonymous data Human subjects data can be describe as anonymous if no identifying information is collected from the individual and thus researchers or any other person cannot know whether or not that individual participated in the study. Some unique individual characteristics (indirect identifiers) could be regarded as identification of a participant. For example, participants who is a member of a certain ethnic group in a specific department might be identifiable from even a large data pool.

Confidential data Human subjects data can be described as confidential when human subjects data can be linked to the person who provided it. Researchers will know that a particular individual has participated in the research even after the research is completed. The researchers are obligated to protect confidential data from disclosure outside of the research team, as indicated in the research protocol and the informed consent document.

De-identified Data: Human subjects data are considered de-identified if the dataset has been stripped of all identifying information and there is no way that it could be linked back to the subjects from whom it was originally collected (through a key to a coding system or by any other means). Even though a dataset may have been stripped of direct identifiers (names, addresses, student ID numbers, etc.), it may still be possible to identify an individual through a combination of other characteristics (e.g., age, gender, ethnicity, and place of employment).

Coded Data: Data are coded when a link will exist between a unique code such as a number, letter, symbol, pseudonym, or any combination and individual subjects' identifiers such as name, medical record number, email address or telephone number. The code should not be a combination of information related to the individual, such as initials, or date of birth.
Note: Coded data are not anonymous.

Protected Health Information (PHI): Individually identifiable health information, held or maintained by a covered entity or its business associates acting for the covered entity, that is transmitted or maintained in any form or medium (including the individually identifiable health information of non-U.S. citizens). This includes identifiable demographic and other information relating to the past, present, or future physical or mental health or condition of an individual, or the provision or payment of health care to an individual that is created or received by a health care provider, health plan, employer, or health care clearinghouse. For purposes of the Privacy Rule, genetic information is considered to be health information.

Private Information: 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 that has been provided for specific purposes by an individual and that the individual can reasonably expect will not be made public (e.g., a medical record).

Sensitive Research Data: Data are considered sensitive when disclosure of identifying information could have adverse consequences for subjects or damage their financial standing, employability, insurability, or reputation.

2. The IRB wants to know the Following Information

All researchers must submit **DATA SECURITY AND STORAGE PLAN** with the IRB application. The following information should be addressed in the form. The list is not exclusive.

- Will University owned or non-University owned hardware, software, and/or equipment be used?
- Who will collect, store, and use data?
- What equipment, tools, and services will be used?

- What data formats will be used (hard copy/paper, digital, audio, video, etc.)?
- Who and how will the data be accessed and/or shared?
- What storage methods will be used?
- What is required for retention and how long does the data need to be kept?
- When and how will the data be destroyed?
- If data will be coded or de-identified, who will do the task(s) and how will they/it be accomplished?

3. IRB recommendation for collecting, storing, and maintaining data securely

1) Use BSU owned equipment and licensed services

The IRB strongly recommends using BSU-owned equipment and licensed services to collect and store research data. It is strongly recommended to use Qualtrics for collecting online survey data and to store research data on BSU secure drives or authorized cloud services like a BSU office 365(OneDrive/SharePoint). Especially when research involves collecting or storing photographic images or voice recordings of research participants and data protected under HIPAA and FERPA, the IRB may require researchers to use BSU-owned devices.
Data should be held on personal devices only for the time necessary to be promptly moved to a secure university-managed location. All personal devices must be password protected.

2) Wearable devices

When using wearable devices, such as an activity tracker, a smartwatch, voice recording devices, location trackers, or other technology to collect research data, information must be included in the informed consent form that states participants will be required to download and agree to terms of service or other agreements applicable to the app if the participant is using their own device and not one provided to them by the researchers. If an app meets the regulatory definition of a mobile medical application as defined by the FDA, additional regulatory determinations may need to be made depending on its intended use.

3) Use Coded Identifiers and a Master Key

To protect the confidentiality of data, it could be a good option to use coded identifiers. You may assign each participant a random unique identifier and use this identifier to label all data collection instruments and sheets.
Note: Do not record any individually identifiable information about the participant as part of the study data.

a. Develop a master key to enable the organization of identifying information and data (preferably in an electronic format and vigorously protected with encryption and passwords). Then, enter the contact or other identifiable information you collect into the master key and record the coded study identifiers in the master key.

b. Once the data are organized and analyzed, the master key (and participant contact information forms, if used) should be destroyed. If it is important to your study to keep the master key, please provide a detailed rationale to the IRB. In your proposal, detail how and when these keys will be destroyed.

c. Data documents should have only the data and the code and all other identifiers must be eliminated. Ideally, the informed consent, data, and the master key should be transported and stored independently (compartmentalization), but reasonable alternatives can be proposed and approved.

4) Plan for Data Transport, Storage, and Security

Transport of data (whether through physical or electronic means) should be limited to reduce the risk of loss or theft. When it is not in transit, data should be stored in a secure location accessible only to authorized research personnel.

a. Data transported physically from a study site to an investigator's office or lab should be locked in a secure container.
b. Data must be transported separately (whether in separate electronic files or physical containers) from consent documentation or master keys to ensure the security. Even if data are lost or stolen, there will be no associated identifiable information at risk of disclosure.
c. Identifiable data and documents should not be stored (except temporarily and out of necessity) at the investigator's place of residence. All identifiable study materials and data should be stored securely on the BSU campus.
d. Electronic data should be stored only on password-protected (and, if possible, encrypted) storage media or computers.
e. Copies of electronic data files should be kept to an absolute minimum. If multiple study personnel need access to the data, storage in a central secure location such as OneDrive is preferable over multiple copies being provided.

f. Electronic data should not be sent over email; but, if necessary, it should only be sent if it is de-identified.

5) Establish a Data Retention Plan

In accordance with federal guidelines, the IRB requires that study data and consent forms must be maintained securely for, at minimum, three (3) years after the completion of a study (this applies only to non-exempt research) unless a specific retention period is established by law, regulation, policy, or contractual agreement. Regulations, best practices, and ethical guidelines in your specific discipline (e.g., those related to data covered by HIPAA) may dictate a longer retention schedule. The IRB requires the following data retention practices.

a. Following the minimum three-year retention period, individually identifiable information (including the master key and any combination of indirect identifiers that could reasonably identify a subject) must be destroyed, if it has not been already. De-identified data may be retained indefinitely.

b. During the retention period, data, signed consent forms and other documentation related to human subjects must be stored in the manner described in the IRB-approved protocol. Access must be limited to those identified in the approved protocol as having access to study data.

6) Use BSU approved Web Conferencing for Colleting Research data

The use of web conferencing to conduct research interviews and/or to collect research data has increased significantly. It is recommended that researchers use BSU approved software or services when conducting these activities. It is researchers' responsibility to ensure their data collection activities are properly secured.

Some web conference software allows the researcher to record sessions, share screens, and automatically transcribe the recording. When recording sessions, researchers must ensure that the recordings are stored on a university secure server or BSU licensed cloud service (OneDrive/ SharePoint).