ISSUE
01
OCTOBER
2020

Compliance

Responsible Conduct in Research

this issue

Data Management Plan

The plan provides a guiding document to refer to as you perform your research and serves as a standard operating procedure for a research group. This is especially helpful for collaborative projects or long-lived projects with short-term or often changing research assistants. These data management plans are longer and more detailed than the plans created for grant applications. Although there may be some differences depending on the research domain, in general, a good data management plan will address the following aspects:

- Define the roles and responsibilities of research team members
- Data types and formats
- Capture methods and file naming
- Ethics and Intellectual Property
- Access, Data Sharing, and Reuse
- Short-Term Storage and Data Management
- Deposit and Long-Term Preservation.

Research Data Security Guidelines

Researchers in the social and behavioral sciences are expected to be proactive in designing and performing research to ensure that the dignity, welfare, and privacy of individual research subjects are protected, and that information about an individual remains confidential. The protection of research data is a fundamental responsibility rooted in regulatory and ethical principles and should be upheld by all data stewards.

The Research Data Security Guidelines pertain to researchers and research team members who obtain, access, or generate research data, regardless of whether the data is associated with funding or not. These guidelines help Alabama A&M University (AAMU) researchers understand the sensitivity of the information they are collecting and develop appropriate data protection plans, know the proper mediums and places to store data, understand how and when to dispose of data, prepare their research data for public use, understand how to keep research data secure while traveling, and what to do in the event of theft, loss, or unauthorized use of confidential research data. These guidelines can also be used as part of the **data management planning** process to meet federal funding agency requirements and prepare research data for public use.

- **Investigator Responsibilities**

Anyone who conducts research with human subjects at AAMU is responsible for protecting the data collected and used for their study. This is especially important when the data (a) contain personal identifiers or enough detailed information that the identity of participating human subjects can be inferred, (b) collect information that is highly sensitive, or (c) is covered by a **restricted use agreement**. The guidelines are intended to help researchers understand when and how to use the most effective and efficient methods for storing and analyzing confidential research data to adequately protect those data from theft, loss, or unauthorized use.

As a general practice, researchers working with human subjects should avoid collecting **Personally Identifiable Information (PII)** whenever possible. Perhaps the best way to protect a research subject's identity is by not knowing that identity in the first place. However, in many cases, the collection of PII is necessary for carrying out a research project.

PII is defined as information that is uniquely associated with a person. The **HIPAA** privacy rules identify items such as name, mailing address, email address, social security number, etc.) that are considered forms of PII.

Restricted Use Agreement

Many researchers at AAMU receive data from outside agencies or institutions that are subject to restricted use agreements (also called data sharing agreements). These are legal contracts that impose restrictions on the researchers' use of the data and sometimes include detailed procedures for secure storage, restricted access, and data analysis. As part of the agreement, certain government agencies may also visit the researcher (or "licensee") to conduct a compliance audit. In other cases, restricted use agreements may prevent public release of the data or sale of the data to a third party. But in cases where an agreement does not specify data security procedures, researchers must consider the need to keep their data secure so that the potential for harm to any individuals or organizations is minimized. When faced with two sets of data security requirements (e.g., one from the AAMU IRB and one from a restricted use agreement), the researcher should always default to the requirements with higher standards for data protection.

Data Protection Standards in Restricted Use Agreement

In cases where a Restricted Use Agreement requires additional or wholly different data storage and security measures, the researcher should default to whichever standards are more secure concerning the data's sensitivity level. In cases where the requirements conflict with the recommended standards outlined above, the researcher should consult with the AAMU Office of Research Compliance.

AAMU Office of Research Compliance (ORC)

Director:

James O. Bukenya, Ph.D.

Phone: 256-372-5729

Fax: 256-372-5911

Email:
research.compliance@aamu.edu

What Type of Research Data do I have?

As part of planning their research design, researchers must develop a data protection plan that adequately protects their data from unintended disclosure or possible theft. Researchers are required to submit a data protection plan as part of their IRB protocol. To support the development of data protection plans, the guidelines define a three-level categorization system for research data and describe the minimum data protections recommended for each level. These storage technologies and data handling practices are designed to be consistent with AAMU's standards for protecting student records and other types of administrative data. A researcher can always elect to use more secure data management methods than the minimum recommended (for example, one could manage Level 1 data as if they were Level 3 data). However, the benefits of additional security layers beyond the minimum recommended need to be weighed against the additional burdens they can impose on the research team members.

➢ LEVEL 1 - Beginning information about individually identifiable people

Level 1 data contain PII on human subjects who have been given an assurance of confidentiality. Level 1 data files do not contain sensitive information but need some protection due to confidentiality assurance. Accidental or unintended disclosure is unlikely to result in harm to the study subjects. The risks to the research subject may be considered no greater than those associated with everyday life.

➢ LEVEL 2 - Sensitive information about individually identifiable people

Level 2 data include PII that, if disclosed, could reasonably be expected to present a non-minimal risk of civil liability, moderate psychological harm, or material, social harm to individuals or groups. The risks to the research subject may be considered greater than those associated with everyday life.

➢ LEVEL 3 - Very sensitive information about individually identifiable people

Level 3 data include PII that could cause significant harm to an individual if exposed, including, but not limited to, serious risk of criminal liability, serious psychological harm or other significant injuries, loss of insurability or employability, or significant social harm to an individual or group. The risks to the research subject may be considered greater than those associated with everyday life.

As a general practice, PII that is needed for project management but not needed for analysis should be separated from the data to be used for analysis at the earliest possible phase of the project. In practice, this usually means splitting the data into two files: one containing all of the PII not needed for the analysis and a unique ID variable, the other containing the same unique ID variable, and all of the data collected for the analysis. The common identifier in both data sets enables the researcher to re-link the PII and non-PII data at a future date, as the project's needs may require.

Removal of all PII (temporarily or permanently) significantly reduces the risk of harm to study participants but does not entirely eliminate the potential for harm resulting from loss, theft, or unintended disclosure. However, researchers should continue to use the minimum data protection standards outlined below when working with de-identified data.

Source:

The guidelines are adopted and modified with permission from the Office of Research Integrity and Assurance at Princeton University. *https://ria.princeton.edu/research-data-security.*

Image source:

https://carpedatumlaw.lexblogplatform.com/wp-content/uploads/sites/190/2016/10/shutterstock_454777225.jpg.

How to Store and Protect Data that Contain PII

Level 1 data that contain PII may be stored on any of the following devices as long as they are at a minimum configured to require users to authenticate themselves using a login ID and password and only allow access to authorized project team members and system administrators: Acceptable storage media include:

- The hard drive of a server or workstation as long as it is configured in a manner that is consistent with the University security practices
- Centrally-managed network file storage
- A secure cloud storage system approved by the Office of Information Technology.
- Any of the following devices that are managed using an audited, check-out/check-in system:
  - An external drive
  - A piece of removable storage media (e.g., USB drive)

When not in use, storage media must be kept in a locked drawer or cabinet in a secured space (e.g., a central storage area or an authorized project team member's office), with key access required for both the office and the storage location. Similarly, data collected or stored on paper forms that have PII (such as signed consent forms or questionnaires) should be stored in a locked file cabinet.

Devices used to access Level 1 data may include any workstation or server that:

- Houses the hard drive on which the data are stored,
- Is physically connected to the external hard drive or the removable medium on which the data are stored,
- Is authorized to access a shared network drive on which the data are stored, or
- Is authorized to connect to a physically-secured and firewall-protected server that has access to the data.

The practices for managing Level 1 data should be reviewed by the researcher at the start of the data set's lifecycle, annually during the lifecycle of the data set, and at the end of the data set's lifecycle. **Note** that the IRB requires this information as part of the annual review process.

The practices for managing **Level 1 data** should be reviewed by the researcher at the start of the data set's lifecycle, <u>annually</u> during the lifecycle of the data set, and at the end of the data set's lifecycle. **Note** that the IRB requires this information as part of the annual review process.

Level 2 data are subject to all of the standard practices listed for Level 1 data with the following adjustments:

- Level 2 data should not be copied to and/or stored on a personal workstation's hard drive unless the Level 2 data is stored on the workstation's hard drive in an encrypted form using encryption technology.
- If Level 2 data is stored on an external hard drive or piece of removable media, the media must be managed to utilize a check-in/check-out mechanism.
- Level 2 data transmitted across the network must be encrypted utilizing an approved encryption protocol and key length.
- When maintaining network-based storage systems containing Level 2 data, system administrators must authenticate themselves using an authentication mechanism that requires the administrator to provide a second factor to validate his or her identity in addition to a password, such as a code sent to a specified mobile device.
- It is recommended that users of Level 2 data not access the data directly from their personal device through network file services but through a server that requires an authentication mechanism that requires the user to provide a second factor to validate their identity.
- Only project team members may be given access to materials related to Level 2 data (derivative results, output, etc.). These individuals should be identified in the IRB protocol associated with the proposed research. For electronic data, access to the related data must be actively managed via system access controls. When not in use, any physical or removable media containing these materials must be stored in a locked drawer in a project team member's office, with key access for both the office and the storage location. It is recommended that Level 2 data should be stored in an encrypted manner using an approved encryption protocol.

Level 3 data are subject to all of the standard practices described for Level 2 data with the following adjustments:

- All authorized users of Level 3 data, including project team members and system administrators, are strongly encouraged to access the data through a server that requires an authentication mechanism that requires the user to provide a second factor to validate his or her identity in addition to a password.
- Client access to the data set should require a file encryption key to decrypt the data. Level 3 data must be stored in an encrypted manner using an encryption protocol and key length.
- Level 3 data must not be copied to and/or stored on a personal workstation's hard drive.

The practices for managing **Levels 2 and 3** data should be reviewed by the researcher at the start of the data set's lifecycle, <u>four times a year</u> during the data set's lifecycle, and at the end of the data set's lifecycle. **Note** that the IRB requires this information as part of the annual review process.

Highly Sensitive Data

Research data are considered highly sensitive when there is a heightened risk that disclosure may result in embarrassment or harm to the research subject. Data on topics such as sexual behavior, illegal drug use, criminal behavior, crime victimization, or mental health are considered highly sensitive. Information that could have adverse consequences for subjects or damage their financial standing, employability, insurability, or reputation should be adequately protected from public disclosure, theft, loss, or unauthorized use, especially if it includes PII.

Public Use Data Files

Public use data files are files from which all PII has been removed, and the data are coded in such a way as to make identification of research subjects extremely unlikely. Researchers who work with the public use data sets that do not contain PII may not meet the regulatory definition of research involving human subjects. However, some restricted use agreements nevertheless require local IRB review. As such, researchers are encouraged to consult the Institutional Review Board to determine if their proposed research requires IRB review.

What to Do in the Event of Theft or Unauthorized Use of Research Data

A data security breach occurs when there is a loss or theft or other unauthorized access to sensitive personally identifiable information that could result in the potential compromise of the confidentiality or integrity of data.

Anyone at AAMU who knows or suspects that their confidential research data has been lost, stolen, or used inappropriately is advised to contact the Office of Research Compliance for immediate assistance.