Document Security and Research Data Management Suggested Best Practices

For information on data classification, see CU system guidance here.

Research data and documents may be collected in many ways (e.g., questionnaires, audio/video tape, online surveys). Additionally, the data may be stored in a range of formats (e.g., handwritten notes, analog/digital recordings, computer files). Each mechanism for collecting and storing data and documents poses issues regarding security against unauthorized access and use, prevention of accidental loss or damage, and eventual disposal. Maintaining human subject data securely with the appropriate level of anonymity, confidentiality, or de-identification is a key factor in minimizing the risk for research participants, the researchers, and the university.

Best Practices for Electronic Data

- All electronic data collection and storage devices should at a minimum be password protected with a strong passphrase. A passphrase should be at least 12 characters in length. Please visit the Office of Information Technology (OIT) password site for additional information.

- If it is necessary to use portable devices for the initial collection of identifiers, the data files should be encrypted, and the identifiers moved to a compliant system as soon as possible. The system will need to meet the regulatory requirements, if applicable, and be vetted by the Office of Information Security. The portable device(s) should be locked in a location where access is limited and/or logged when not in use. All data collected on portable devices should be transferred to an approved platform as soon as possible after collection and shredded from the portable device.

    - If you are working with HIPAA data, it is recommended that you use specific HIPAA compliant external hard drives, which are encrypted or able to be locked with keypads.

- Access to identifiable data should be on a need-to-know basis.

- Remove all direct identifiers as soon as possible, possibly using codes in lieu of the identifiers, keeping identifiers separated from coded data. Identifiers, data, and keys should be placed in separate password protected/encrypted files, and each file should be stored in a different secure location.

- Ensure that you are storing and/or transmitting files on an approved platform. When transferring files via email, encryption is highly recommended.

- Use accepted practices to protect against indirect identification, such as aggregate reporting or pseudonyms.

- Ensure that the research data, when destroyed, is done so in a manner that protects the participants. Electronic data files should be deleted from all storage devices including any recycle bins. OIT has resources for helping with remediation of electronic data to learn more visit UCCS OIT.

Best Practices for Hard Copy Data

- Hard copies of items containing any identifiable information should be kept secure with limited access. These items may include, but are not limited to, consent forms, payment forms, data forms, keys/codes for data, and some types of questionnaires.

- Hard copies of identifiable information can be stored in locked file cabinets, locked offices, and locked storage facilities. They should not be stored in open access areas, vehicle trunks, or exposed areas. If hard copies of data are to be transferred and stored off campus, please contact the IRB to discuss resources and security requirements before doing so.

- Hard copies of the data should be shredded at the appropriate time interval after research is complete. This is typically 3 years from the end date of the research. For the majority of protocols, the regulations allow for electronic copies of records to be kept and hard copies to be destroyed at any time. However, certain funders have different requirements. **If your study is funded, please consult IRB staff before destroying any hard copies.**

Key Definitions

The UCCS IRB often finds incorrect use of anonymous, confidential, and de-identified data on applications. Understanding the correct use of these terms can assist in determining appropriate data management and security for the protocol and minimizing application revisions.

- *Anonymous Data:* Data originally collected without identifiers. Data are anonymous if no one, including the researcher, can connect the data to an individual. No identifying information is collected from the participant. Be aware that the collection of indirect identifiers (i.e., information regarding other unique characteristics) might make it possible to identify an individual from a pool of subjects. For example, a participant that is a member of a minority ethnic group might be identifiable from the pool of subjects.

- *Confidential data:* A link between the data and the individual that provided the data exist. The research team is obligated to protect the data from disclosure outside the research according to the terms of the research protocol and informed consent document. The best defense against a breach of confidentiality is multiple layers of security.

- *Identifiable data (identifiable private information):* Private information for which the identity of the subject is or may readily be ascertained by the investigator or associated with the information.

- *De-identified data:* Although collected in an identifiable manner, any direct or indirect identifiers or codes linking the data to the individual participants are removed and destroyed.

- *Encryption:* Using one or more mathematical approaches to obscure or make data unreadable without the associated key or password