DATA SECURITY

This document provides guidance to investigators on protecting human subjects' identifiable data. See CPHS Data Security for policy information, including relevant definitions. Should you need additional assistance please contact the Office for the Protection of Human Subjects (OPHS) at 510-642-7461 or at ophs@berkeley.edu.

Table of Contents:

A. General Information

People who volunteer to participate as subjects in research do so with the understanding that the researcher(s) will protect their identity and the information obtained about them from inadvertent or inappropriate disclosure. The principle that the Institutional Review Board (IRB) upholds in assessing the benefits and risks of the research is expressed in the Belmont Report as "beneficence"— to maximize possible benefits and minimize possible harms to human subjects. As benefits and risks may be reflected in protections for privacy and confidentiality, all human subject research protocols must have in place an acceptable and documented procedure for the protection of identifiable and/or confidential information before the protocol will be approved, amended, granted continuing approval, or determined exempt from IRB review. See CPHS Data Security for policy information, including relevant definitions.

B. Physical Data

Investigators should be careful to protect confidential physical records. Confidential paper records should be kept in locked file cabinets when not in use and physical access to any facility that contains confidential information should be restricted. Access control measures include smart card swipes, PIN keypads, and locked doors. Investigators should be aware of who has access to keys, and should consider this when storing data. Confidential information should not be left on copiers, fax machines, or other shared devices.

C. Electronic Data

1. Secure passphrases should be used to protect electronic data files, including digital audio recordings. Passphrases should be sufficiently complex and adhere to ISO's Passphrase Complexity Guidelines.

2. When encrypting data, investigators should consult with UCB's Information Security Office (ISO) or comparable centralized campus technology experts, when determining which encryption software to use. For reference, see guidelines on Data Encryption on Removable Media.

3. In general, investigators should encrypt identifiable data before it is transferred over a network or over email. Consent forms should also be encrypted before being transferred over a network or over email when the forms themselves have the potential to reveal information that could place the subject at risk of criminal or civil liability or be damaging to the subject's financial standing, employability, educational advancement, or reputation. Campus storage systems (bDrive, and Box) are compliant with P3 security requirements, and so P3 level data and below need not be first encrypted if using these systems, unless these data are subject to additional external requirements, such as CA State data. For more information on data protection levels, see the below matrix.

4. When using an online data collection site (e.g. Amazon Mechanical Turk, Qualtrics, etc.), investigators should carefully review the site's data security policy. If the site stores identifiable information and/or links survey responses to individual participants, this must be made clear in the investigator's eProtocol submission and in the corresponding informed consent document(s).

5. Generally, an electronic data file may be stored online (e.g., on a cloud storage system) only if the file does not contain identifiable information or has first been encrypted so that, should there be a security breach, the data cannot be linked back to individual participants. However, UC Berkeley's Box may be used for non-encrypted P3 level data and below, unless the data are subject to additional external requirements, and CalShare (SharePoint Online) may be used for non-encrypted P4 level data and below. If considering storing data on a cloud, investigators should first consult ISO or centralized campus technology experts to determine which cloud computing service to use. Important considerations include: (1) data storage location; (2) backup policy; (3) deletion policy; (4) rights that the cloud provider claims for the data; (5) isolation guarantees that the provider offers. Investigators may wish to consider using local hosting options.

D. International Research

1. As *CPHS Data Security* policies apply regardless of location, investigators need to make sure that they have appropriate security measures in place while in the field, in transit, and back at their permanent residence.

2. Depending on a number of factors, including political climate and availability of secure storage locations, investigators may find it difficult to maintain data security while in the field. In such circumstances, investigators are encouraged to upload their data to a cloud storage system. (See C5 above for further guidance.)

3. When traveling across U.S. borders, investigators should be aware that the U.S. government could, at their discretion, take an electronic device, search through all the files, and keep it for further scrutiny. While this current policy is evolving and may change in the future, investigators should still take special care to encrypt all confidential information on their electronic devices when traveling across U.S. borders. It should be noted that confidential information should always be encrypted when stored on a removable medium, regardless of border crossings.

E. Questions for Investigators and Research Staff to Ask Themselves:

- Am I collecting or retaining any identifiable data beyond what is absolutely necessary for the study? Have I destroyed identifiers that are no longer needed for my research?

- Have I replaced all personally identifiable information in my research records with a code and kept the key to identifiers separate from the records? (For example, the key to identifiers is kept in a separate encrypted file or in a separate locked file cabinet.)

- Do I routinely and regularly review and update my data security procedures?

- If I'm collecting sensitive information, have I consulted with information security experts to make sure my research and/or clinical data are secure from both physical and electronic theft?

F. Human Research Data Risk Assessment Matrix and Security Measures

Failure to meet security objectives can cause harm to an individual and/or the University. This document provides impact assessments that describe the level of harm that such failure may cause. Required security measures are listed below the matrix (see next pages).

Human Research Data Risk Assessment Matrix

Impact/Risk Assessment	Description/Examples
Minimal/UC P1 *(Public Information)*	Information intended for public access, but whose integrity is important. For example, published research.
Low/UC P2 *(Non-sensitive information and de-identified information)*	• De-identified human subject information with negligible re-identification risk and no Notice-Triggering data elements, that a subject has been told would remain confidential, even though no harm is expected if this information were to be disclosed. • Information of which unauthorized use, access, disclosure, acquisition, modification or loss could result in minor damage or small financial loss, or cause minor impact on the privacy of an individual or group. • Non-public research using publicly available data. • Public Directory Information for faculty, staff, and students who have not requested a FERPA block.
Moderate/UC P3 *(Moderately sensitive individually identifiable information)*	• Individually identifiable information that, if disclosed, could reasonably be expected to be damaging to a person's reputation, cause embarrassment, have a moderate impact on the privacy of a group; result in moderate financial loss; or could require legal action. This classification level also includes lower risk items that, when combined, represent increased risk. • Student record information protected by FERPA. • Personal information as defined in the General Data Protection Regulation (GDPR). • Medical devices supporting diagnostics (not containing extremely sensitive subject information). • Low risk export controlled data or technology (EAR/ITAR). Contact the Export Control Office for a determination. • Information that could cause harm to an individual if disclosed, including, but not limited to, risk of criminal liability, psychological harm or other injury, loss of insurability or employability, or social harm to an individual or group.

Impact/Risk Assessment	Description/Examples
High/UC P4 (*Extremely sensitive individually identifiable information*)	• Data elements with a *Statutory Requirement for Notification* to affected parties in case of a confidentiality breach, for example: o Social security number (SSN) o Driver's license number o California State identification number o Financial account numbers, credit or debit card numbers and financial account security codes, access codes, or passwords o Personal medical information, including protected health information (PHI) covered under HIPAA. o Personal health insurance information o A username or email address, in combination with a password or security question and answer that would permit access to an online account • General Data Protection Regulation (GDPR) special categories (Article 9 'sensitive') of identifiers. • Federal Controlled Unclassified Information (CUI) • Financial aid and student loan information • Financial, accounting, and payroll systems • Passport documentation (images and numbers) • High-risk export controlled data or technology (DoE 10 CFR Part 810, high-risk EAR/ITAR). Contact the Export Control Office for a determination. • Any confidential or personal information that requires the highest level of access control and security protection, whether in storage or in transit. E.g., human subject genetic information, etc.

For more information on the above categories, see ISO's Data Classification Standard and How to Classify Research Data.

Security Measures:

As applicable, employ the following security measures. (Note that there are no specific requirements for the protection of public research information, but researchers should protect such data for other reasons (e.g., intellectual property issues, sponsor requirements, protect against unauthorized modification, etc.).)

- Collect the minimum identity data needed.
- Whenever possible, de-identify and/or separate data elements into a coded data set and an identity-only data set.

- Limit access to personally identifiable information.
- Encrypt data if identifiable information is: (1) stored on a networked computer or device, (2) stored on or transmitted via the internet, (3) stored on a computer or removable medium, which is not permanently located in a secure location. Moderately sensitive or non-sensitive data from publicly available sources, which also contain identifiable or potentially identifiable information, do not require encryption except to conform with terms of use from data providers or sponsors.
- Employ Minimum Standards for the Security of Networked Devices.
- Employ Minimum Security Standards for Electronic Information.
- For research involving High Risk/UC P4 data, supplementary security requirements apply. Contact UC Berkeley's Information Security Office for further guidance.