DATA SECURITY

This document provides guidance to investigators on protecting human subjects' identifiable data. See CPHS Data Security for policy information. 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.

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 take this into account when storing data. Confidential information should not be left on copiers, fax machines, or other shared devices.

C. Electronic Data

1. When password-protecting documents and computers, secure passwords should be created. Passwords should contain at least 10 characters, a mix of upper- and lower-case letters, and combinations of numbers and symbols. To further protect against a possible security breach, real names, birthdates, phone numbers, street or email addresses, or other personally identifiable information should not be included as part of a password.

2. When encrypting data, there are many different encryption software options, including Truecrypt for Windows and MEO Free Encryption Software for both Mac and Windows. Investigators should be sure to consult with UCB's Information Services and Technology (IST), or comparable centralized campus technology experts, when determining which encryption software to use.

3. When collecting data online, investigators should be cautious about stored IP addresses and data that could be accessed by a third party. Investigators should make sure to encrypt identifiable data before it is transferred over a network or over email.

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. A dataset may be stored online (e.g., on a cloud storage system) only if the dataset 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. If considering storing data on a cloud, investigators should first consult IST 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 can, 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 near 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?

Human Research Data Risk Assessment Matrix

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 and appropriate security measures relative to the risk (see next pages).

Human Research Data Risk Assessment Matrix

Impact/Risk Assessment	Description	Security Measures
De-identified and non-confidential information	De-identified research information about people and identifiable information which subject has consented to make publicly available.	There are no specific requirements for the protection of de-identified research information or for other non-confidential research information, but researchers may want to protect such data for other reasons (e.g., intellectual property issues, sponsor requirements, etc.).
Low *(Non-sensitive individually identifiable information)*	Individually identifiable information which subject has been assured would remain confidential, even though no harm would be expected if this information were to be disclosed.	• 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 web, (3) stored on a computer or removable medium which is not permanently located in a secure location.

Impact/Risk Assessment	Description	Security Measures
Medium *(Moderately sensitive individually identifiable information)*	• Individually identifiable information that, if disclosed, could reasonably be expected to be damaging to a person's reputation or to cause embarrassment. • Student record information protected by FERPA. • 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.	The previous measures apply.
High (*Extremely sensitive individually identifiable information*)	• California State Law "notice triggering information" which includes first name OR first initial and last name in combination with one or more of the following: (1) social security number, (2) driver's license number, (3) California identification number, (4) financial account number, credit or debit card number, in combination with any required security code, access code, or password that would permit access to an individual's financial account, (5) medical information, (6) health insurance information. • Health information protected by HIPAA. • Any confidential or personal information that requires the highest level of access control and security protection, whether in storage or in transit.	The previous measures apply, in addition the following: • Employ high security protective measures described in Minimum Security Standards for Electronic Information. • Employ Minimum Standards for the Security of Networked Devices (if applicable).