Data Management and Security Guidance

***Page 1** describes the minimum core data security procedures that should be in place for the protection of subject data. These procedures are triggered by a "yes" response to item 11.1 in the eResearch application. The nature of the research, sensitivity of the data, etc., can influence whether the core data security procedures are adequate or if additional steps should be taken to ensure data and subject safety.*
***Page 2** presents the template for a required data management and security protocol document (final document uploaded to sec. 44)*
***Page 3** is a table that lists protected health information, personal identifying information, and other sensitive information*

Required Minimum Data Security Controls for Collection of Personally-identified Data

Minimum data security controls are intended to establish and maintain a low risk threshold. Failure to implement the data security "best practices" could result in increased risk to subjects. As part of the eResearch application, the PI must demonstrate that all of the core data security control elements have been met. The core controls are:

1. All data collection and storage devices must be password protected with a strong password. A strong password requires a level of complexity. Please follow the link for crafting a strong password.
2. All data/research files must be encrypted.
3. Access to identifiable data should be limited to members of the study team.
4. Identifiers, data, and keys should be placed in separate, password protected/encrypted files and each file should be stored in a different secure location.
5. For secure data transmission, Transport Layer Security (TLS) (a.k.a. SSL) or equivalent, and a minimum key length of 128 bits must be used for any data that is transmitted electronically.
6. If it is necessary to use portable devices for collection of identifiable data, the data files should be encrypted and the identifiers moved to a secure system as soon as possible after collection. Additionally, the portable device should be locked up in a secure location when it is not in use. The PI should consult with their departmental IT Security Liaison to discuss how to correctly configure desktop computers, laptops, and other devices for safe use in the collection and storage of research data.
7. M+Google Mail and Calendar services may not be used to collect, store, or transmit confidential or sensitive human subjects research data. For a list of allowed and restricted services when storing/transmitting sensitive identifiable data, see http://safecomputing.umich.edu/dataguide/?q=node/65
8. No protected health information or other sensitive information should be transmitted via email, except within the U-M Health System and Medical School.
9. If utilizing any cloud-computing services, the PI must follow the UM guidelines found at http://www.safecomputing.umich.edu/cloud/ and at http://www.safecomputing.umich.edu/google/

Additional Required Data Security Controls – if data are of a higher sensitivity (see Table 1, page 3)

1. All data should be downloaded from local devices to a secure UM server as soon as possible after collection.
2. Passwords should be built in at multiple levels on each local machine that is used for the collection and storage of research data (e.g. at BIOS and at login).
3. If research includes sensitive identifiable data, outside consultants or vendors should be required to sign a confidentiality agreement.
4. If the research design allows, the PI should delete or destroy identifiable information as soon as possible after collection.

Data Management and Security

For applications that may require Full Board review: *The following questions serve as a template. The questions must be presented and followed by your responses within a stand-alone Data Management and Security document.*

For Non-full board applications: *Use this page as a reference to assist in responding to the questions about data security, identifiable data, etc., within the eResearch application.*

1. **What is the nature of the data?**
   a. Electronic (text, audio, video, binary), hardcopy files, or biological specimens?
   b. Do the data contain protected health information, personal identifying information or other sensitive information? If yes, please precisely describe what these are (see Table 1, page 3).
   c. Are identifiers retained and linked to the data? Who will have access to the data? Who will have access to the identifiers?
   d. Are the data stripped of identifiers and the identifiers destroyed (anonymized data)? When will this take place?
   e. Are identifiers de-linked from the data and managed by use of a code? How are the identifiers, data files and key managed and secured? Who will have access to the identifiers, data files and key?
2. **Where and how will the data be stored and what security measures will be used for each?**
   a. Personal computer or laptop? University computer or laptop; location? Office file cabinet? Thumb/jump drive? Departmental or other U-M server; name and/or location?
   b. What security measures will be used with each (password protected; encryption; locked file cabinet in locked office, 128 bit encryption, etc.)?
   c. Who will have access to the computer/laptop/server/or files?
3. **How will data be transmitted or transported?**
   a. How will electronic files be transmitted? What measures are in place for secure transmission of data?
   b. How will hardcopy files be transported?
   c. How are the files and data protected while in transmission or when transported?
4. **When and how will data or records be deleted or destroyed?**
5. **Will cloud-computing resources be used?** (refer to UM policies at http://www.safecomputing.umich.edu/cloud/ and at http://www.safecomputing.umich.edu/google/)
   a. What is the resource and what is the privacy policy for the resource?
6. **Will online data collection services be used?**
   a. What is the service/host? How is the survey accessed? How are data accessed by the study team? Will any non-secure services be used to access, collect, or transmit data (e.g., public portals, administrator logins, public WiFi networks, or public computers)?
   b. How are data moved/transmitted from the online host to the local storage device (computer, laptop, server, thumb drive, etc.)?
   c. Will the data be purged from the online host once downloaded to the local device? How and when?
   d. If the data are identifiable and sensitive, are confidentiality agreements in place with outside consultants or vendors?
7. **Will any datasets be used?**
   a. Is there a Memo of Understanding (MOU) or Data Use Agreement (DUA) associated with the use of these data?
   b. Does your security plan include all requirements contained in the MOU/DUA?

Table 1: Protected Health Information, Personal Identifying Information and Sensitive Information^

Protected Health Information (PHI): An individual's personal and health information that is created, received, or maintained by a health care provider or health plan and includes at least one of the 18 personal identifiers listed below in association with the health information:

- Name
- Street address
- All elements of dates except year
- Telephone number
- Fax number
- Email address
- URL address
- IP address
- Social security number
- Account numbers
- License numbers
- Medical record number
- Health plan beneficiary #
- Device identifiers and their serial numbers
- Vehicle identifiers and serial number
- Biometric identifiers (finger and voice prints)
- Full face photos and other comparable images
- Any other unique identifying number, code, or characteristic

Limited Data Set - a limited data set can include the following identifiers: a unique number code, or characteristic that does not include any of the above listed identifiers, Geographic data (without street address), and/or dates.

Private Personal Information (PPI): Information about an individual which includes any of the identifiers below:

- Name
- Street address
- All elements of dates except year
- Telephone number
- Fax number
- Email address
- URL address
- IP address
- Social security number
- 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
- Driver's License numbers or other identification card number
- Device identifiers and their serial numbers
- Vehicle identifiers and serial number
- Biometric identifiers (finger and voice prints)
- Full face photos and other comparable images
- Any other unique identifying number, code, or characteristic (e.g., student identification number)

Certain categories of sensitive information may require additional considerations due to regulatory or other requirements (e.g., FERPA and student information, GLBA and customer information, employee information, and donor information).

Other Sensitive Information

An individual's first name (or first initial) and last name in combination with any of the following:

- Social Security Number
- Driver's License Number or California ID card number
- Financial account information such as a credit card number
- Medical Information

Note: Identifiers in combination with data about illegal behaviors, physical/mental health information, or other information that poses a risk to subject reputation, insurability, employability, or legal status will heighten the level of sensitivity and require additional corresponding security measures.

^ Borrowed from *Guidance and Procedure: Data Security in Research*, UCLA Office of the Human Research Protection Program (OHRPP), February 24, 2011