# Data Security and Privacy Checklist

**Purpose:** To provide a planning tool primarily for use by researchers to think through the research plan and prepare for submitting an application or grant proposal. The checklist is intended to strengthen project plans, alerting researchers to potential vulnerabilities, and to prompt additional planning to reduce information risks, to the extent necessary and feasible.

Research involves increasingly complex arrangements for the storage and transmission of research data. Robust data privacy and security planning is necessary to protect the privacy of research subjects and to secure sensitive, personally identifiable information (PII).

After completing the checklist, researchers are encouraged to contact their institutional compliance areas such as the IRB or IACUC, and/or IT departments as appropriate. The checklist is not intended as an audit tool; it does not certify compliance and expresses no opinion as to the adequacy of any given plan.

**Scope:** All students, faculty, or staff that conduct a research project

|  |  |
| --- | --- |
| **Project Information** | |
| 1.1 Project Details | Project Name:  Single Site Study:  Yes  No  Will there be a coordinating center?  Yes  No  Will data be shared between centers?  Yes  No |
| 1.2 Principal Investigator (PI) Information | PI Name:  PI Institutional Affiliation: |
| 1.3 Data Manager/Data Custodian  *(Individual responsible for data, other than PI)* |  |
| 1.4 Study Coordinator Name: |  |
| 1.5 Other Persons at the Institution with Access to the Data (indicate role/title) |  |

|  |  |
| --- | --- |
| **Receiving and Collecting Data** | |
| 2.1 Will data be obtained from a source outside the study?  *(e.g., a vendor, a company, a collaborator from a different institution or department, a government agency)* | Yes  No  If yes, please specify: |
| 2.2 Will data be produced by the study? | Yes  No  If yes, please describe the datasets: |
| 2.3 Where, and in what format, will data be stored? | Data will be stored:  Yes  No  Format of data: |
| 2.4 Will this project involve secondary use of data?  *(i.e., reuse of data from another project)* | Yes  No  If yes, list the project name and investigator who originally obtained the data: |
| 2.5 Is there an approval letter from the original data owner for this reuse? | Yes  No |
| 2.6 Is this research funded by an outside sponsor? | Yes  No  If yes, please specify: |
| 2.7 Do the terms of the award or the research agreement limit how the data may be used, maintained, or shared? | Yes  No  If yes, please specify: |

|  |  |
| --- | --- |
| **Data Information** | |
| 3.1 Data type | Lab data  Survey data  Imaging data  Claims and enrollment  Service  Clinical data  Genetic information  Media (Video , photo , audio )  Other, please specify: |
| 3.2 Will the data contain any HIPAA identifiers?  *(Refer to the Additional Checklist Guidance section for examples of HIPAA identifiers)* | Names  Geographic subdivisions smaller than a state (except the first three digits of a zip code)  Elements of dates (except year) directly related to an individual, including birth date, admission date, discharge date, date of death  Ages over 89  Telephone number  Fax numbers  Email addresses  Medical record numbers  Health plan beneficiary numbers  Account numbers  Certificate/license numbers  Vehicle identifiers and serial numbers  Device identifiers and serial numbers  Web Universal Resource Locators (URLs)  Internet Protocol (IP) address numbers  Biometric identifiers, including finger and voice prints  Full face photographic images or any comparable images  Any other unique identifying number |
| 3.3 Does this research involve identifiable human subject data? | Yes  No  If yes, has an IRB reviewed this study?  Yes  No  Pending  IRB Name:  IRB Approval date (if any):  IRB approval number (if any): |
| 3.4 Does this research involve animal research? | Yes  No  If yes, has an IACUC reviewed this study?  Yes  No  Pending  IACUC Name:  IACUC Approval date (if any):  IACUC approval number (if any): |
| 3.5 Will datasets received or created be “limited datasets”?  *(Refer to the Additional Checklist Guidance section for the definition of limited datasets.)* | Yes  No  Coded?  Yes  No  If yes, who has the link? |
| 3.6 Will data be de-identified … | De-identified?  Yes  No  If yes, will a third-party de-identification service be used? |
| 3.7 For what purpose will data be used? | For student research  Yes  No  For post-doctoral research  Yes  No  For publication?  Yes  No  For external collaboration?  Yes  No  Other, please describe: |

|  |  |
| --- | --- |
| **Data Storage, Access, Collection, and Security** | |
| 4.1 Does this study have a Data Management Plan (DMP) or Data Security Plan? | Yes  No  If yes, who approved the plan? (i.e., IRB, IACUC, IT Services) |
| 4.2 Describe how data will be stored while the study is active: | Data storage:  If data will be collected, transmitted, and/or analyzed via an internet application or cloud service, include the security plan for this data, if any: |
| 4.3 Where will the data be accessed from? | Data will be accessed from:  Marquette network (hard drive, local server, SharePoint)  Internet/web application  Cloud service  Other  If accessed from a source not within the Marquette Network, please specify: |
| 4.4 Will the data be accessed from a remote device *(e.g., e-tablet, smart-phone, non-Marquette managed personal/home computer)*? | Yes  No  If yes, please specify: |
| 4.5 Will data from this study be stored electronically? | Yes  No  If yes, please specify:  Local server (SharePoint or Teams)  Third party servers  Hard drives  Portable devices  Other, please describe: |
| 4.6 Does the researcher have a reporting plan in the event of intentional or unintentional loss, alteration, or destruction of data? | Yes  No  If yes, please specify: |
| 4.7 Will the researcher keep paper-based records? | Yes  No  If yes, please specify: |
| 4.8 Does the researcher have a plan for maintaining backup copies of the data? | Yes  No  If yes, please specify: |
| 4.9 Does the researcher have means to notify institutional departments or data vendors about material changes to the data plan? | Yes  No  If yes, please specify: |
| **Data Sharing and Data Transport** | |
| 5.1 Will this data be shared with individuals outside of MU staff, faculty, or students within the research group (e.g., external collaborators)? | Yes  No  If yes, please specify: |
| 5.2 Will data be submitted to publicly accessible repositories during the research? | Yes  No  If yes, please specify: |
| 5.3 If the project involves Protected Health Information, are appropriate agreements in place? | Yes  No  If yes, please specify:  Has the IRB approved the plan to share the data?  Yes  No |
| 5.4 Is there a plan for encryption of data when transferred electronically from site to site or safeguarding of data if physically transported? | Yes  No  If yes, please specify:  Has IT Services approved the transfer plan:  Yes  No |
| 5.5 Will data be collected, analyzed, stored on an internet application or remote third party service? | Yes  No  If yes, please describe the security protocols for the application: |
| 5.6 For data that is being shared, in what format is the data being sent? | Identifiable  Coded  De-identifiable  Other, please describe: |

|  |  |
| --- | --- |
| **Data Retention and Destruction** | |
| 6.1 How long will the data be stored?  Additionally, if data is stored beyond 7 years, what value will the data have after 7 years? | Storage duration:  Value of data stored beyond 7 years, if applicable: |
| 6.2 Is there a plan for post-study disposal / destruction of data? | Yes  No  If yes, please specify: |
| 6.3 How will data be returned to the original owner (either the funding institution, principal investigator (PI), or the research participant), if applicable? | Describe process: |
| 6.4 Will the data be populated within the MU data repository upon research being completed? | Yes  No  If no, please specify where the data is being stored: |

Additional Checklist Guidance

Definition of PHI

Any individually identifiable health information, whether oral or recorded in any form or medium that

* Is created or received by a health care provider, health plan, public health authority, employer, life insurer, school or university, or health care clearinghouse, and
* Relates to the past, present, or future physical or mental health or condition of an individual; the provision of health care to an individual; or the past, present, or future payment for the provision of health care to an individual

HIPAA Identifiers

* Names
* Geographic subdivisions smaller than a state (except the first three digits of a zip code if the geographic unit formed by combining all zip codes with the same three digits contains no more than 20,000 people and the three digits of a zip code for all such geographic units containing 20,000 or fewer people is changed to 000).
* Elements of dates (except year) directly related to an individual, including birth date, admission date, discharge date, date of death; and all ages over 89
* Telephone numbers
* Fax numbers
* Email addresses
* Medical record numbers
* Health plan beneficiary numbers
* Account numbers
* Certificate/license numbers
* Vehicle identifiers and serial numbers
* Device identifiers and serial numbers
* Web Universal Resource Locators (URLs)
* Internet Protocol (IP) address numbers
* Biometric identifiers, including finger and voice prints
* Full face photographic images or any comparable images
* Any other unique identifying number, characteristic, or code

Limited Datasets

DO NOT include direct identifiers (see above), but datasets may include the following indirect identifiers:

* Town or city, state, zip code
* Ages in years up to 90 years (must aggregate all ages 90 or older)
* Dates directly related to an individual—such as birth date, date of death, admission date, discharge date, visit date, diagnosis date, etc. (Month/Year is preferred [no exact day]). Sometimes vendors or agencies provide a study number with the data. To be labeled as a limited dataset these study numbers CANNOT be an encoded identifier such as a scrambled birth date, patient initials, last four digits of the social security number, etc.