HARVARD CATALYST | THE HARVARD CLINICAL AND TRANSLATIONAL SCIENCE CENTER

HARVARD CATALYST
DATA PRIVACY AND SECURITY PLANNING CHECKLIST

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.

This checklist is a planning tool primarily for use by investigators as they think through their research and prepare an IRB application. The checklist is intended to strengthen project plans, alerting investigators to potential vulnerabilities, and to prompt additional planning to reduce information risks, to the extent necessary and feasible..

After completing the checklist, investigators are encouraged to contact their institutional compliance support staff 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.

In addition to investigators, this tool may also be useful to IRBs, as a supplement to application forms, and to Institutions, to adapt into institutional policies and procedures. This document was created by the Harvard Catalyst Data Protection Subcommittee and is available for use across all Harvard Catalyst Institutions.

I. 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)		

II. RECEIVING AND COLLECTING DATA

2.1 Will data be obtained from a source outside the study? **(i.e.. 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 data sets:**	
2.3 Where, and in what format, will data be stored?	**Data will be stored:** **Format of data:**	
2.4 Will this project involve secondary use of data (i.e. re-use 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 re-use?	**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 ☐**	

III. 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?	**☐ Names** **☐ Geographic subdivisions smaller than a state (except the first three digits of a zip code (See Appendix)** **☐ Elements of dates (except year) directly related to an individual, including birth date, admission date, discharge date, date of death;**	

	☐ 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	
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 data set contain any HIPAA Identifiers or contain any personal information? **See Appendix A: List of HIPAA identifiers and MA Data Security law definition of Personal Information**	**Yes☐ No ☐** **If yes, please specify:**	
3.5 Will data sets received or created be "Limited Data Sets"? **See Appendix A: Definition of Limited Data Set**	**Yes☐ No ☐**	
3.6 Will data be …	**Coded? Yes☐ No ☐** **If yes, who has the link?** **De-identified? Yes☐ No ☐** **If yes, will a third-party de-identification service be used?** **Yes☐ No ☐**	

	If yes, please specify:	
3.7 For what purpose will data be used?	**a. For student research? Yes☐ No ☐** **b. For post-doctoral research? Yes☐ No ☐** **c. For publication? Yes☐ No ☐** **d. For external collaboration? Yes☐ No ☐** **e. Other, please describe:**	

IV. DATA STORAGE, ACCESS, COLLECTION, AND SECURITY		
4.1 Does this study have a Data Management Plan or Data Security Plan:	**Yes☐ No ☐** **If yes, who approved this plan? (i.e. IT Department, IRB):**	
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 From where will the data be accessed?	**Data will be accessed from:** **If from an internet/web application or cloud service, please specify:**	
4.4 Will the data be accessed from a remote device (i.e. e-tablet, smart-phone, home computer)?	**Yes☐ No ☐** **If yes, please describe:**	
4.5 Will data from this study be stored electronically?	**Yes☐ No ☐** **If yes, please specify:** **☐ Local servers** **☐ Third party servers** **☐ Hard Drives** **☐ Portable devices** **☐ Other, please describe:**	
4.6 Do you 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 you keep paper based records?	**Yes☐ No ☐** **If yes, please specify:**	
4.8 Do you have a plan for maintaining backup copies of your data?	**Yes☐ No ☐**	

4.9 Do you have means to notify institutional departments or data vendors with regard to material changes to your data plan?	**Yes☐ No ☐** **If yes, please specify:**	

V. DATA SHARING AND DATA TRANSPORT		
5.1 Will this data be shared with individuals outside of your research group (i.e., collaborators)?	**Yes☐ No ☐** **If yes, please describe:**	
5.2 Will data be submitted to publicly accessible repositories (i.e. GWAS, DbGAP)?	**Yes☐ No ☐** **If yes, please specify:**	
5.3 If the project involves Protected Health Information, are appropriate agreements in place? has the IRB approved the plan to share this data? (See appendix for definition of "PHI")	**Yes☐ No ☐** **Has the IRB approved the plan to share this data?**	
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 ☐**	
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 protocol for the application:**	

VI. DATA RETENTION AND DESTRUCTION		
6.1 How long will the data be stored?		
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, if applicable?		

APPENDIX A

I. 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."

II. 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

II. Limited Data Sets: (LDS) do NOT include direct identifiers (see above) but 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 data set 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.

http://www.nist.gov/customcf/get_pdf.cfm?pub_id=50835

http://www.hhs.gov/ocr/privacy/index.html

III. Personal Information (MA Data Security Law)

First name and last name or first initial and last name in combination with any one or more of the following data elements that relate to such resident: (a) Social Security number; (b) driver's license number or state issued identification card number; or (c) financial account number, or credit or debit card number, with or without any required security code, access code, personal identification number or password, that would permit access to a resident's financial account.

Attribution, Sharing and Adapting the Data Privacy and Security Planning Checklist

We encourage you:

- **To request** — to email us and request the materials
- **To share** — to copy, distribute, and transmit the work
- **To adapt** — to adapt the work to suit your needs

Under the following conditions:

- **Attribution**: In freely using the materials, we require that you acknowledge Harvard Catalyst as the publisher and that you give appropriate credit to the individual authors.

- **Suggested citation:** *This material is the work the Harvard Catalyst Data Protection subcommittee. The Data Protection subcommittee is a subcommittee of the Regulatory Knowledge & Support Program and affiliated with Harvard Catalyst | The Harvard Clinical and Translational Science Center. This work was conducted with support from Harvard Catalyst | The Harvard Clinical and Translational Science Center (National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health Award 8UL1TR000170-05 and financial contributions from Harvard University and its affiliated academic health care centers). The content is solely the responsibility of the authors and does not necessarily represent the official views of Harvard Catalyst, Harvard University and its affiliated academic health care centers, or the National Institutes of Health.*

With the understanding that:

- **We might contact you**: We are interested in gathering information regarding those who are using the material and how they are using it. We may contact you by email to solicit information on how you have used the materials or to request collaboration or input on future activities.
- **We ask that you share your adaptation:** If you adapt the tool, please share them with us so that we may support and improve the checklist. All contributors will be appropriately acknowledged. Please send your requests, questions and comments to *regulatory@catalyst.harvard.edu* and visit the Harvard Catalyst Regulatory Knowledge and Support web page.
- **You respect the rights of others**: In writing this material, Harvard Catalyst is indebted to the work of others. In no way are any of the rights of others in the work itself or in how the work is used affected by our adaptation.

When reusing or distributing, make clear the above terms: For any reuse or distribution, you must make clear to others the terms of this work. The best way to do this is with a link to the web page containing this checklist.