Packet Overview

Dear Data Owner. This packet can be used to help navigate the regulatory and data security requirements for hosting and vending high risk data. Please let us know if you have any questions. You can reach Isabella Chu at itaylor@stanford.edu

|  |  |
| --- | --- |
| File Name | Purpose of Document |
| i\_ Data\_Risk\_Assessment | Early in the process, you should complete a Data Risk Assessment with your institution’s data security office. In addition to personally identifiable information, there are many other elements (proprietary concerns, security concerns) which can merit additional layers of protection. |
| 1\_Protocol\_51290  2\_Protocol\_51296  1B and 1C: Boilerplate waiver of HIPAA, Consent and Assent language | Data Core Protocols:  Protocol 51290 covers receipt, cleaning, storage, linkage, de-identification and risk reduction and manipulation of almost any variable imaginable.  Protocol 51296 allows data exploration, pilot work, variable identification and preliminary code. Also all research on “deidentified” datasets using PHS hosted data.  These protocols can be copied for your use.  Depending on the risk classification, specific, high value datasets are likely to require their own protocol for any work beyond ingest, curation and pilot exploration. |
| 3\_ Institutional\_DUA\_Template\_15Apr19 | This document can be used if you are sharing copies of moderate or high risk data with another institution or there will be substantive data exchange or research collaboration with another institution. It is designed for the case of data exchange between two collaborative, non-profit institutions. It is not intended to cover contractual obligations that involve fee for service or procurement of existing commercially available datasets. |
| 4\_DMP\_General\_01Mar19 | This describes procedures for managing high risk data which you can modify according to the standards at your institution. |
| 5\_permission\_group\_SAMPLE\_16Apr19 | You should consult with your data security office to set permissions for your dataset. There may be special requirements you have for your dataset. Our platform is able to automate much of the tracking and collection of these requirements but this can also be done manually. |
| 6\_DUA\_Individual\_15Apr19 | For moderate or high risk data, we recommend that you have every individual accessing data sign a DUA. You can customize this for your organization. |
| 7\_terms\_citation\_costs\_template\_15Apr19 | This should be customized to each high or moderate risk dataset and covers things like:  1) **Citation language. You should make sure there is a paper that can be cited every time someone uses your data!**  2) Who can use the data/PI requirements  3) Linkage restrictions  4) Data overlay restrictions  5) Cell size suppression and other non-intuitive nuances of use  6) Fees to use data |
| 8\_Data Download Policy | This is the Stanford PHS data download policy. We recommend you do not allow downloads of raw data for high risk datasets. You can customize this for you own institution. |
| 9\_[PHS Data Security Training](https://drive.google.com/file/d/0BzXdI855xhjrU0owUzVlaE9GS0E/view) | This 15 minute training is specific to Stanford’s Center for Population Health Sciences. However, this can be updated for general use and other types of high risk data. If you wish to receive the slides for this training to modify for your purposes please email [phsdatacore@stanford.edu](mailto:phsdatacore@stanford.edu). We will be coming up with a new and improved training in 2020. |