

OSR Fair 2016: Data Use Agreements

DATA USE AGREEMENTS



Memo from Dean Arvin

Dated 2/5/15

Access to large data sets has become a key component of research at Stanford. Often, the data providers, or recipients, require the researcher or Stanford to sign a written or online agreement. This memo clarifies when Stanford researchers may sign these agreements themselves and when to contact a University official to review and sign the agreement.

Stanford's research contracting and procurement offices can answer questions and help you navigate these agreements.

- Agreements with government or non-profit entities are handled by the University's Office of Sponsored Research.
- Agreements with industry should be sent to the Industrial Contracts Office.
- Agreements to purchase or store data are handled by the Procurement Office.

Agreements for Incoming Data:

You may sign a data agreement in your individual capacity under the following conditions, which relate to: (a) the nature of the data, and (b) the proposed terms of agreement. This applies whether the agreement is a letter, non-disclosure agreement, a license, or comes in another form – including online “click” agreements.

The Data:

- Does not include “personal identifying information” (PII), “Protected Health Information” (PHI), identifiable education records, or other personal, private or confidential information that may not be publicly disclosed.
- is not obtained from human subjects, even if de-identified.
- is neither identified as, nor is known to be, export-controlled.
- You and your lab members have no financial interests in the data provider.

The Agreement:

- Includes no data security requirements, citizenship access restrictions or penalties for not complying with the agreement.
- Contains no terms relating to the data provider's future publications, research results, or more extensive or restricted your publications.
- Contains no language that subjects the University to liability, such as insurance coverage, or indemnification terms, or statements that the agreement is governed by foreign law.

When you sign an incoming data agreement, you are making a commitment to the University. If you have any questions, contact the university contracting office above. They will ask for information about the research, such as a project description, your affiliation, and a contact person for approval. They will consult with the Office of General Counsel, the Office of Risk Management, the Privacy Office, the Information Security Office, and the Export Control Office.

Agreements for Outgoing Data:

These agreements are signed when you send out data with the agreement involving use of the data. They are not signed when you receive data from other parties.

- The data is to be shared in support of an approved project (e.g., it is required by a regulatory or contractual agreement, or derived from other data).
- The data is to be used for additional research purposes, and your laboratory is responsible for the appropriate use of personal and non-personal information.

The Stanford Research Computing Center can assist with data security

What is a DUA?

A data use agreement (DUA) is an agreement that must be entered into before there is any transfer, use, or disclosure of data to an outside institution or entity. A DUA might be included in a Sponsored Research Agreement, or you may need a separate agreement entirely.

At a minimum, any DUA must contain provisions that:

- Identify the Data itself
- Establish the permitted uses and disclosures of the Data
- Identify who may use or receive the information
- Prohibit the Recipient from using or further disclosing the information, except as permitted by the agreement or as otherwise permitted by law
- Require the Recipient to use appropriate safeguards to prevent an unauthorized use or disclosure not contemplated by the agreement
- Require the Recipient to report to the Data Provider any unauthorized use or disclosure of which it becomes aware
- Require the Recipient to ensure that any agents (including any subcontractors) to whom it discloses the information will agree to the same restrictions as provided in the agreement

What is NOT a DUA?

A research DUA for OSR is not appropriate when Data is anticipated to be used for traditional business, clinical, financial, or administrative functions in support of University and/or Hospital operations. DUAs are not intended to govern any confidential discussions with outside entities regarding a proposed study (a COA or NDA).

Instructions

All Stanford Faculty and Staff are subject to the limitations and requirements regarding the right to enter into a Data Use Agreement. The OSR decision tree is intended to help determine in tandem with ISR's policy not only which Office at Stanford will handle the Data Use Agreement, but also when a Data Use Agreement is either required or recommended.

In order to determine if a DUA is required or recommended, work your way from the center wheel and determine the status of the other Party. If the status is green, proceed to the middle wheel and determine the Data risk level based on the definitions provided. If the status is green, proceed to the last wheel. If the Data is subject to any of the qualifications under the red outer label, a DUA is required. **Remember:** If you ever land on a red option, you know a DUA is required and you contact your institutional official in either ICO or OSR.



THE MATRIX

Common Issues with DUAs

- Patient or research subject identifiable information places a lot of legal responsibility upon Stanford which may not be appropriate/allowable
- Data security, transfer, and use requirements within a DUA can be extensive/burdensome and may not be something Stanford can comply with
- Data Providers often require some form of publication review/restriction to ensure individual data is not released.
- The use of any Background IP may affect rights in the Data and can require additional licensing or clearances before using the Data.
- Data can be designated as confidential or sensitive by the Provider and may require various security clearances or liability assurances from the Recipient.

Common Data Providers

- The most common Data providers are actually other U.S. educational institutions
- The Database of Genotypes and Phenotypes (dbGaP)
- The National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) Central Repository
- The Wellcome Trust Case-Control Consortium (WTCCC)
- Centers for Medicare & Medicaid Services (CMS)



Common Questions from OSR

- Who owns/generates the Data?
- What will the Data be used for?
- Is the Data being identified/used under a Sponsored Project?
- What is the risk level of the Data?
- Where will the Data be stored? What provisions have been made regarding the system security?
- Who will have access to the Data? Will the Data need to be shared with outside third-party entities?
- If the Data is from human subjects, do you and the Recipient/Provider have an approved Stanford IRB protocol that contains the use of the Data?

Stanford

Office of Sponsored Research

Data Definitions and Guidance

Definitions

Data: Factual information used as a basis for reasoning, discussion, or calculation.

Research Activities: The conduct of scholarly inquiry as it relates to the performance of a study.

Non-Research Activities: Traditional business, clinical, financial, or administrative functions in support of University and/or Hospital operations.

Stanford Memo from Dean Arvin regarding Data Use Agreements

All Stanford Faculty and Staff are subject to the limitations and requirements regarding the right to enter into a Data Use Agreement. The OSR decision tree is intended to help determine in tandem with DoR's policy not only which Office at Stanford will handle the Data Use Agreement, but also when a Data Use Agreement is either required or recommended.

https://doresearch.stanford.edu/sites/default/files/documents/faculty_guidance_for_data_agreements.pdf

Resources and Links

- Stanford webpage for PHI policies and compliance: <https://acp.stanford.edu/hipaa/hipaa>
- Stanford Security Risk Classifications: <https://uit.stanford.edu/guide/riskclassifications>
- Stanford Industrial Contracts Office: <https://sites.stanford.edu/ico/>
- Stanford Office of Sponsored Research: <https://doresearch.stanford.edu/research-offices/osr-office-sponsored-research>

Data Classifications

Data Risk Classifications

Low Risk Data: The data is intended for public disclosure and any loss of the data would have *no* adverse impact on Stanford. Examples: Research data (at PI's discretion), information in the public domain.

Medium Risk Data: The data is not generally available to the public, or the loss of the data could have a *mildly* adverse impact on Stanford. Examples: Unpublished Research data (at PI's discretion), Stanford internal memos or reports, student records.

High Risk Data: Protection of the data is required by law, or the loss of the data or system could have a *significant* adverse impact on Stanford. Examples: SSN's, financial account numbers, export controlled information, donor contact information.

Health-Related Data Definitions and General Classifications

De-identified Data Set (Low Risk). Health information that does not identify an individual.

Limited Data Set (Medium Risk). Health information that includes up to three of the following individual identifying elements: dates (admission, DOB, DOD, service, etc.), location (state, city, zip code), and age.

Protected Health Information ("PHI") (High Risk). Individually identifiable health information consisting of 18 specific identifying data elements that identifies the individual.