

K-State Resources For Developing Data Management Plans and for Storing and Sharing the Data Associated with Them

Dept of Computer and Information Sciences,
K- State Libraries, and
Office of Research and Sponsored Programs

Why save the data?

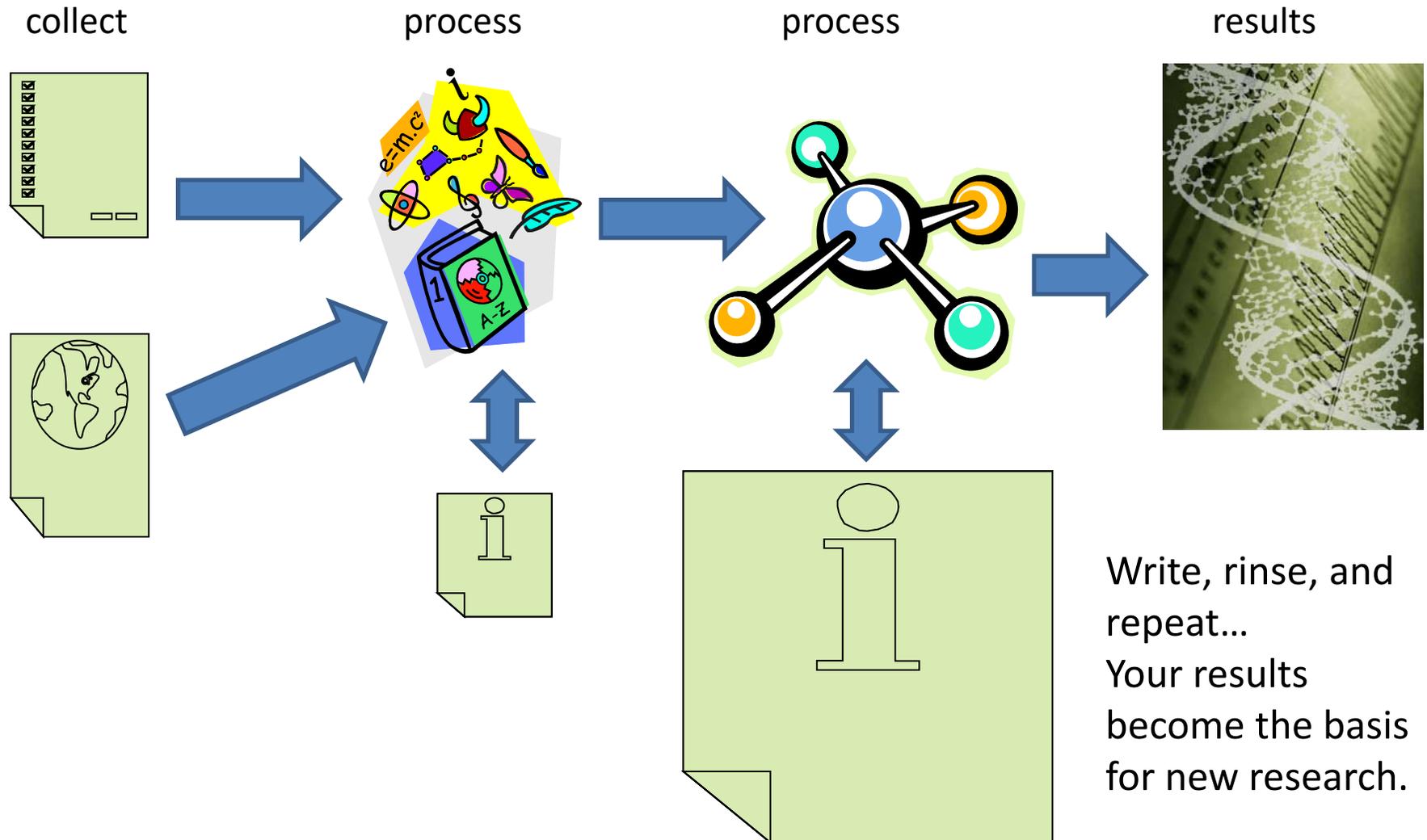
During a decade as head of global cancer research at Amgen, C. Glenn Begley identified 53 "landmark" publications -- papers in top journals, from reputable labs -- for his team to reproduce. Begley sought to double-check the findings before trying to build on them for drug development.

Result: **47 of the 53 could not be replicated.** He described his findings in a commentary piece published on Wednesday in the journal Nature.



Sharon Begley, "In cancer science, many "discoveries" don't hold up,"
Reuters, Mar 28, 2012

Where does your data go?

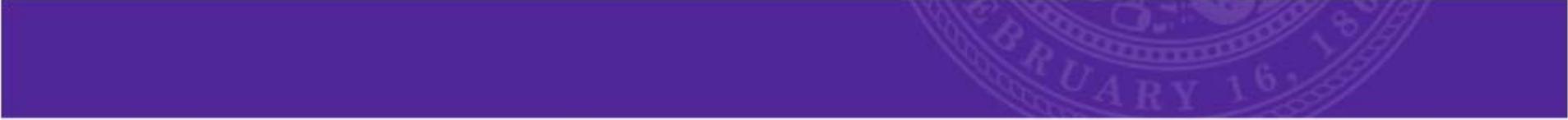


Classifying data

| | Highly active | Largely inactive |
|-----------------------|-------------------------------------------------|---------------------------|
| Small data (KB to GB) | Make searchable and available online | Make available on request |
| Big data (GB to TB) | Make searchable online and available on request | Make available on request |

A few thoughts

- NSF is still being lenient, but it can matter
 - Competition is really, really stiff
- Be specific
 - “We will produce docs and handle them in an domain-standard manner.” is not a data plan.
 - “Data-mentionment” plans...
- Your tool chain matters too
 - Try a workflow tool like VisTrails or Taverna



K-State Research Exchange

Marty Courtois

Scholarly Communications & Publishing

212 Hale Library

(785) 532-4428

courtois@ksu.edu

K-REx is...

- an **institutional repository**. Our goal is to collect the research and scholarly output of K-State
- an **archive**. You can **self-archive** your work to ensure long-term availability
- an **open access** repository. Your work will be freely available on the web

Benefits of Depositing Your Work in K-REx...

- Safe **digital archive**
- Each item is assigned a **permanent URL**
- Full **metadata** description
- Indexed in **Google**, Google Scholar, etc.
- Will likely **increase citations** to your work
- Supplements rather than replaces traditional **scholarly publishing**

Data

- K-REx can accept **any file format**
- User must have **software** to view the file
- Best to use **open formats**
 - Comma Separated Values (.csv)
 - Extensible Markup Language (.xml)
 - Portable Document Format (.pdf)
- Avoid **proprietary formats**
 - Excel (.xlsx, .xls)
 - Access (.accdb, .mdb)
 - Statistical analysis software(.sas, .spss)

Data

- Smaller files (up to 1 GB)
- No restriction on access
- Active or inactive files

To Submit Files to K-REx

- Contact me (courtois@ksu.edu)
- Send me your files
- Archive your data as well as the accompanying journal article
- Library will ensure we comply with journal publisher's copyright and self-archiving policies
- Coming soon: submission form on K-REx homepage (<http://krex.ksu.edu>)

Office of Research and Sponsored Programs



Welcome to the Office of Research & Sponsored Programs!

The Office of Research & Sponsored Programs (ORSP) located on the main floor in Fairchild Hall, Room 102, promotes and facilitates the research, education, and outreach missions of Kansas State University by supporting and administering all activities involving extramurally sponsored programs. The mission of ORSP is to provide the Kansas State University community with a wealth of services (some listed below).

Preaward Services (PAS) located in Room 002 in the basement of Fairchild Hall, has the legal responsibility to review ALL proposals prepared for submission to ensure compliance with the sponsor's programmatic guidelines, state/federal regulations, and University policy and commitments. PAS has been delegated University signatory authority which is required for a proposal to be submitted to a sponsor. When planning to submit a proposal, make sure you contact PreAward Services FIRST.



Data Management Plan(DMP)Tool and Library Resources

Meagan Duever
Data Services Librarian
124 Hale Library
mjduever@k-state.edu
785-532-7447

Data Management Tags: k-state manhattan

Last Updated: Mar 2, 2012 | URL: <http://guides.lib.k-state.edu/content.php?pid=298773> | [Print Guide](#) | [RSS Updates](#) | [SHARE](#) [f](#) [t](#) [e](#) ...

- [Home](#)
- [Organizing Your Data](#)
- [Metadata](#)
- [Archiving/Preservation](#)
- [Sharing Your Data](#)
- [Data Management Resources](#)

[Home](#) | [Comments\(0\)](#) | [Print Page](#) | **Search:** | [This Guide](#) ▼ | [Search](#)

Why Should you Manage your Data?

To Organize your Research

Managing your data increases efficiency, and makes it easier to understand the details and procedures relating to your data and its collection throughout the lifecycle of the project. Sure, you remember every step of how your data was collected and how it's been manipulated today, but will you remember months from now when you are consumed by the next project? Can others follow the notes you've made to repeat your experiment? Or to re-use your data for another purpose?

To Satisfy Funding Requirements

A growing number of agencies that fund research (like NSF and NIH) are requiring the data to be preserved and shared. A data management plan will help you meet these requirements by addressing documentation and preservation issues.

To Preserve your Data

The data you collected is the foundation of your research. Preserving your data ensures that your work can support future research, facilitates new discoveries, and allows other researchers to recognize your contribution. A good preservation plan also makes disaster recovery possible.

[Comments \(0\)](#)

Subject Guide



Meagan Duever

Contact Info

Data Services Librarian
124 Hale Library
Manhattan, KS
[Send Email](#)

Links:

[Profile & Guides](#)

Data Management Planning

- [Data Plan Guide](#)
Questions to ask as you begin to plan your research
- [Elements of a Data Management Plan](#)
What you need to consider as you create your data management plan

[Comments \(0\)](#)

Requirements by Funding Agency

- [National Science Foundation \(NSF\)](#)
- [National Institutes of Health \(NIH\)](#)
- [National Endowment for the Humanities \(NEH\)](#)

[Comments \(0\)](#)

Data Management Tags: k-state manhattan

Last Updated: Mar 2, 2012 | URL: <http://guides.lib.k-state.edu/content.php?pid=298773> | [Print Guide](#) | [RSS Updates](#) | [SHARE](#) [f](#) [t](#) [e](#) ...

- Home
- Organizing Your Data
- Metadata
- Archiving/Preservation
- Sharing Your Data
- Data Management Resources**

Data Management Resources [Comments\(0\)](#) [Print Page](#) **Search:** [This Guide](#)

Checklists

- MIT's Data Management Checklist
- California Digital Library's Data Planning Check List
- Stanford's Questions to ask as you prepare a grant proposal
- ICPSR's Elements of a Data Management Plan
- Digital Curation Centre's Checklist for a Data Management Plan

Comments (0)

K-State Resources

- Office of Research and Sponsored Programs
- University Research Compliance Office
- Research Policies
- Statistics Consulting Lab

Comments (0)

Planning Tools

- California Digital Library's DMP Tool
- Purdue's Data Curation Tool
- Digital Curation Centre's DMP Online Tool

Comments (0)

Data Repositories

- List by Repository
Purdue D2C2 List
- Repository List by Subject
Open Access Directory List

Comments (0)

Templates

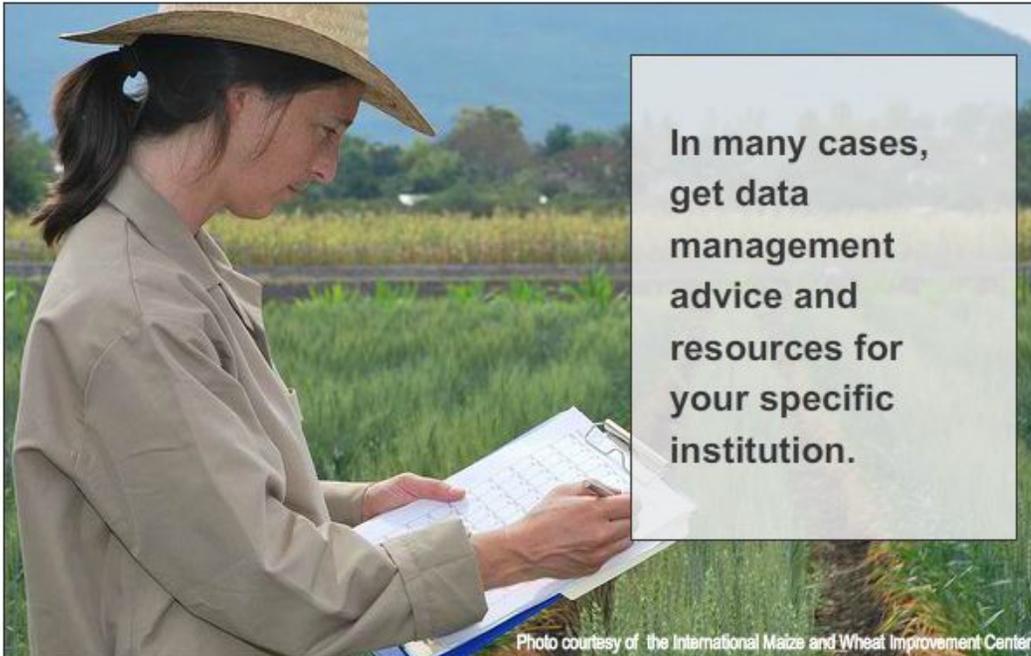
- DataONE
- Data Conservancy
- University of Minnesota
- University of Virginia Library
- California Digital Library

Comments (0)

DMPTool

Guidance and Resources for your Data Management Plan

[Home](#) [About DMP Tool](#) [DMP News](#) [My Plans](#) [Funder Requirements](#) [Help](#) ▾



**In many cases,
get data
management
advice and
resources for
your specific
institution.**

Photo courtesy of the International Maize and Wheat Improvement Center

The DMP Tool allows you to: **1 2 3 4**

[Get Started!](#)

**Data Management Plan: Sample Plan Created at the
DataONE Best Practices Workshop - Santa Fe NM 7/2011
Atmospheric CO2 Concentrations, Mauna Loa
Observatory, Hawaii, 2011-2013**

1. Types of data produced

An samples at Mauna Loa Observatory will be collected continuously from an intake located at five towers, a central tower and four
towers located at various positions. Raw data files will contain continuously measured CO2 concentrations, calibration standards,
reference standards, daily check standards, and blanks. The sample lines located at various positions were used to examine the
influence of source effects associated with wind directions. In addition to the CO2 data, we will record real time wind speed and
direction, temperature, humidity, precipitation, and cloud cover. Site conditions at Mauna Loa Observatory will also be noted and
reported. The final data product will consist of 5-minute, 15-minute, hourly, daily, and monthly average atmospheric concentration of

[See a plan created with the DMP Tool](#)

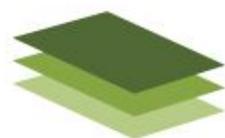
Recent DMP News

[DataBib registry of repositories now linked](#)

[Survey on role of libraries in data management through
April 9](#)

[Support for NSF Emerging Frontiers in Research and
Innovation \(EFRI\) grants](#)

[More news >](#)



DMPTool

Guidance and Resources for your Data Management Plan

[Contact Us](#) | [Get Started](#) | [Login](#)

[Home](#)

[About DMP Tool](#)

[DMP News](#)

[My Plans](#)

[Funder Requirements](#)

[Help](#) ▼

Get Started

Login

If your institution is listed below the DMP tool will provide links to local data management resources and support available to you.

If you're using the DMP tool for the first time you'll be prompted to provide some additional information

None of the above ▼

I am a:

new user

returning user

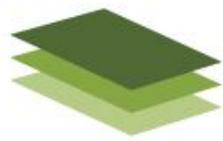
Anyone can use the DMP Tool

Don't see your organization in the list? *You can still use the DMP Tool...* just select "None of the above" and you'll be able to create an account or login.

DMPTOOL is a service of the [University of California Curation Center](#) of the [California Digital Library](#)

Copyright © 2010-2012 The Regents of the University of California

[Privacy Policy](#) | [Terms of Use](#) | [Photo Credits](#)



DMPTool

Guidance and Resources for your Data Management Plan

[Contact Us](#) | [Manage Profile](#) | [Logout](#)

You are logged in as Meagan Duever

[Home](#)

[About DMP Tool](#)

[DMP News](#)

[My Plans](#)

[Funder Requirements](#)

[Help](#) ▼

My Data Management Plans

Create a new plan:

Existing plans: 2

Plan name: something

Solicitation No.:

Funder: Institute of Museum and Library Services

Status: You provided responses for 0 out of 11 questions

Comment:

[\[edit\]](#) [\[view\]](#) [\[delete\]](#) [\[share\]](#) **Export to:** Plain Text Rich Text

Plan name: aerial

Solicitation No.:

Funder: NSF-EAR: Earth Sciences

Status: You provided responses for 0 out of 5 questions

Comment:

[\[edit\]](#) [\[view\]](#) [\[delete\]](#) [\[share\]](#) **Export to:** Plain Text Rich Text

Tips

Choose export to create a plan to save to your local drive.

You can choose to "share" a PDF version of your plans. You will be provided with a URL for the plan to share with others. You can retract a shared plan if you no longer wish it to be available.

Recent DMP News

[DataBib registry of repositories now linked](#)

[Survey on role of libraries in data management through April 9](#)

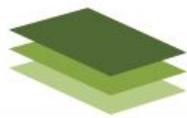
[Support for NSF Emerging Frontiers in Research and Innovation \(EFRI\) grants](#)

[More news >](#)

DMPTOOL is a service of the [University of California Curation Center](#) of the [California Digital Library](#)

Copyright © 2010-2012 The Regents of the University of California

[Privacy Policy](#) | [Terms of Use](#) | [Photo Credits](#)



NSF-CISE: Computer and Information Science and Engineering: Plan description

Progress

Click on a section below to edit it at any time.

✔ = complete

Plan description

- 1. Roles and responsibilities
- 2. Types of data
- 3. Policies for access and sharing and appropriate protection and privacy
- 4. Data storage and preservation of access
- 5. Additional possible data management requirements

The NSF-CISE: Computer and Information Science and Engineering plan will cover the subject areas listed to the left.

You can save a plan in progress and return later to finish or edit.

Plan Name: (required)

Please give your plan a name to help you identify it in the future

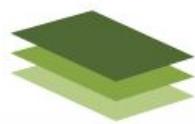
Solicitation Number:

Comment:

Provide any notes you want to appear on your My Plans page. This will not appear in the document

B *I* U abε x₂ x² aa ab ABC ↶ ↷ ↻ ↺ ↻ ↻ ↻ ↻

Resources



NSF-CISE: Computer and Information Science and Engineering: 1. Roles and responsibilities

The Data Management Plan should clearly articulate how the PI and co-PIs plan to manage and disseminate data generated by the project. The plan should outline the rights and obligations of all parties as to their roles and responsibilities in the management and retention of research data, and consider changes that would occur should a PI or co-PI leave the institution or project. Any costs should be explained in the Budget Justification pages.

Progress

Click on a section below to edit it at any time.

✔ = complete

Plan description

1. Roles and responsibilities
2. Types of data
3. Policies for access and sharing and appropriate protection and privacy
4. Data storage and preservation of access
5. Additional possible data management requirements

Help box size: [small](#) | [medium](#) | [full](#)

Explain how the responsibilities regarding the management of your data will be delegated. This should include time allocations, project management of technical aspects, training requirements, and contributions of non-project staff—individuals should be named where possible. Remember that those responsible for long-term decisions about your data will likely be the custodians of the repository/archive you choose to store your data. **While the costs associated with your research (and the results of your research) must be specified in the Budget Justification portion of the proposal, you may want to reiterate who will be responsible for funding the management of your data.** Consider these questions:

- Outline the staff/organizational roles and responsibilities for implementing this data management plan.
- Who will be responsible for data management and for monitoring the data management plan?
- How will adherence to this data management plan be checked or demonstrated?
- What process is in place for transferring responsibility for the data?
- Who will have responsibility over time for decisions about the data once the original personnel are no longer available?

B *I* U abc **X₂** **X²**

Resources

General

- [NSF Data Sharing Policy](#)
- [NSF Data Management Plan Requirements](#)
- [NSF CISE Data Management Guidance](#)

NSF-CISE: Computer and Information Science and Engineering: 2. Types of data

The Data Management Plan should describe the types of data, samples, physical collections, software, curriculum materials, or other materials to be produced in the course of the project. It should then describe the expected types of data to be retained and shared, and the plans for doing so. The DMP should cover how data are to be managed and maintained during the project.

Progress

Click on a section below to edit it at any time.

✔ = complete

Plan description

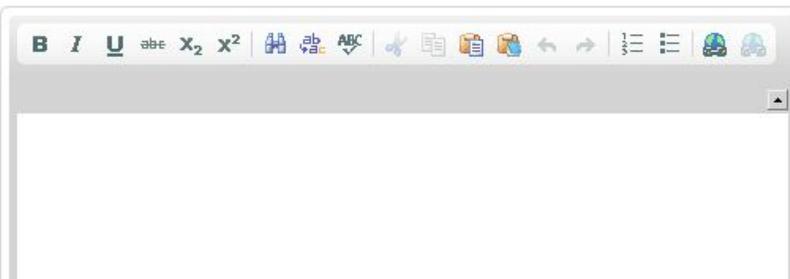
- ✔ 1. Roles and responsibilities
2. **Types of data**
3. Policies for access and sharing and appropriate protection and privacy
4. Data storage and preservation of access
5. Additional possible data management requirements

Help

box size: [small](#) | [medium](#) | [full](#)

Give a short description of what "data" will mean in your research—explain what the contents of each dataset will be, including size and amount if known. It would also help if you can identify your methods for collecting data. *Consider these questions:*

- What data will be generated in the research?
- What data types will you be creating or capturing? (E.g. experimental measures, observational or qualitative, model simulation, processed etc.)
- How will you capture or create the data?
- Which file formats will you use for your data, and why?
- If you will be using existing data, state that fact and include where you got it. What is the relationship between the data you are collecting and the existing data?
- Where (physically) and on what media will you store the data during the project's lifetime?
- How will you back-up the data during the project's lifetime and how regularly will back-ups be made?
- Who is responsible for the storage and backup of the project's data?
- What data will be preserved and shared?



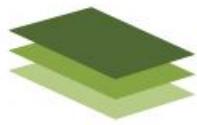
Resources

General

[NSF Data Sharing Policy](#)

[NSF Data Management Plan Requirements](#)

[NSF CISE Data Management Guidance](#)



NSF-CISE: Computer and Information Science and Engineering: 3. Policies for access and sharing and appropriate protection and privacy

The Data Management Plan should describe the period of time the data will be retained and shared; factors that limit the ability to manage and share data, e.g., legal and ethical restrictions on access to human subjects data; and provisions for appropriate protection of privacy, confidentiality, security, and intellectual property.

Progress

Click on a section below to edit it at any time.

✔ = complete

[Plan description](#)

- ✔ 1. [Roles and responsibilities](#)
- ✔ 2. [Types of data](#)
- 3. [Policies for access and sharing and appropriate protection and privacy](#)
- 4. [Data storage and preservation of access](#)
- 5. [Additional possible data management requirements](#)

Help

box size: [small](#) | [medium](#) | [full](#)

In this section you are also asked to account for issues of privacy, confidentiality and ownership that may arise from the dissemination of your data. If the data is of a sensitive nature that public access is inappropriate, address here the means by which granular control and access will be achieved (e.g. formal consent agreements; anonymization of data; level of aggregation; restricted access, only available within a secure network). Think about what you have done to comply with your obligations in your IRB Protocol. *Consider these questions:*

- How and when will you make the data available? (Include the resources needed to make the data available: equipment, systems, expertise, etc.)
- What other types of information should be shared regarding the data, e.g. the way it was generated, analytical and procedural, information?
- What is the process for gaining access to the data?
- Will any permission restrictions need to be placed on the data?
- How will you manage data with sensitive information?
- Are there ethical and privacy issues? If so, how will these be resolved?
- What have you done to comply with your obligations in your IRB Protocol?
- How long will/should data be kept beyond the life of the project?



Resources

General

[NSF Data Sharing Policy](#)

[NSF Data Management Plan Requirements](#)

[NSF CISE Data Management Guidance](#)