



Research Data Management & NIH DMSPs

Jenny Muilenburg

Interim Director, Health Sciences Library

Research Data Services Librarian

December 6, 2024

Today's Session

- Data management (aka, stewardship)
- Resources for NIH DMSP
- Tools & on-campus support
- Q&A



Got data?

- Data definition(s)
- Data needs
 - management
 - stewardship
 - sharing

Ten Simple Rules for the Care and Feeding of Scientific Data (2014)

RDM Workshop

- Protect it from loss
- Find it when you need it
- Secure it
- Reuse it
- Share it

The screenshot shows a Canvas LMS interface. On the left is a navigation menu with links: Home, Announcements, Modules, Collaborations, Syllabus, Pages, Files, People, Assignments, Grades, BigBlueButton, Quizzes, Outcomes, Discussions, Rubrics, and New Analytics. Each link has a small icon to its right. The main content area is titled 'Research Data Management Class [SUM2024]' with an 'Assign To' button, an 'Edit' button, and a three-dot menu button. Below the title is a purple banner that reads 'Data Management Planning Workshop'. The text below the banner says: 'Welcome to Data Management Planning, an asynchronous online workshop for UW community members engaged in research with data. Data is the center of most research. This course will teach you how to keep your data accessible and useable when, where, and how you need it. Good data management practices will let you:' followed by a bulleted list: 'Protect your data from loss', 'Find it when you need it', 'Secure your data', 'Reuse your old data', and 'Share your data'. At the bottom, it says: 'Over the next four days (Monday-Thursday), we will complete [four distinct modules](#) related to the management of research data. Modules typically consist of resources, an activity, and a peer discussion forum. To get the most out of this workshop, we ask that you complete all modules within a four day period, at an

RDM Workshop



Illustration credit: Ainsley Seago.



UNIVERSITY LIBRARIES
UNIVERSITY of WASHINGTON

Roles & Responsibilities

- Who is the project director?
- Who collects the data?
- Who knows where all of the data is stored and makes sure it's organized correctly?
- Who generates the metadata?
- Who analyzes the data?
- Who is responsible for the backup schedule?

What is a DMP/DMSP?

A data management plan (DMP) is a document which defines how data handled throughout the lifecycle of a project—that is, from its acquisition to archival.

Example from Nature

Reporting standards and availability of data, materials, code and protocols

An inherent principle of publication is that others should be able to replicate and build upon the authors' published claims. A condition of publication in a Nature Portfolio journal is that **authors are required to make materials, data, code, and associated protocols promptly available to readers without undue qualifications**. Any restrictions on the availability of materials or information must be disclosed to the editors at the time of submission. Any restrictions must also be disclosed in the submitted manuscript.

After publication, readers who encounter refusal by the authors to comply with these policies should contact the chief editor of the journal. In cases where editors are unable to resolve a complaint, the journal may refer the matter to the authors' funding institution and/or publish a formal statement of correction, attached online to the publication, stating that readers have been unable to obtain necessary materials to replicate the findings.



NIH DMSP Resources

Data Management and Sharing Policy

NIH has a longstanding commitment to making the results of NIH-funded research available. Responsible data management and sharing has many benefits, including accelerating the pace of biomedical research, enabling validation of research results, and providing accessibility to high-value datasets.



About Data Management & Sharing Policies

NIH encourages the sharing of data whenever possible. Learn about the 2003 NIH Data Sharing policy and the 2023 NIH Data Management and Sharing policy as well as how they apply to NIH funded research and data.



Planning & Budgeting for Data Management and Sharing

Find out what NIH expects in a Data Management & Sharing plan and what costs are allowed in a request.



Data Management

Proper data management is crucial for maintaining scientific rigor and research integrity. Learn about best practices for scientific data management.



NIH DMSP Resources

- [Example DMPs](#)
- [DMPTool](#). Provides text-based templates, generates a 2-page text document.
- UW membership with [data repository Dryad](#), free to UW researchers to use to openly share data, and one of NIH's recommended generalist repositories.
- [Qualitative Data Repository](#) (trial)

DMPTool

NIH DMSP test plan

Project Details Collaborators Write Plan Research outputs Finalize Download

Project title *

☒ mock project for testing, practice, or educational purposes

Project abstract

B *I* 12pt ▾ A ▾ ☰ ☷ ☹ ☰

Press Alt 0 or Option 0 for help using the rich text editor with keyboard only.

Research domain

- Please select one - ▾

Select Guidance

To help you write your plan, MyApp can show you guidance from a variety of organizations.

Select up to 6 organizations to see their guidance.

☒ DMPTool
☒ University of Washington (uw.edu)
(washington.edu)

Find guidance from additional organizations below

[See the full list](#)

Save

DMPTool

NIH DMSP test plan

Project Details

Collaborators

Write Plan

Research outputs

Finalize

Download

This plan is based on the "NIH-GEN DMSP (Forthcoming 2023) " template provided by National Institutes of Health (nih.gov) - (ver: 7, pub: 2023-01-06).

expand all | collapse all

6/12

+ Data Type (0 / 3)

+ Related Tools, Software and/or Code (0 / 1)

+ Standards (0 / 1)

+ Data Preservation, Access, and Associated Timelines (0 / 3)

+ Access, Distribution, or Reuse Considerations (0 / 3)

+ Oversight of Data Management and Sharing (0 / 1)



DMPTool

NIH DMSP test plan

Project Details

Collaborators

Write Plan

Research outputs

Finalize

Download

New research output

View all research outputs

Type *

✓ - Please select one -

Audiovisual

Collection

Data paper

Dataset

Event

Image

Interactive resource

Model representation

Physical object

Service

Software

Sound

Text

Workflow

Abbreviation

▼

≡

≡

≡

≡ ▼

≡ ▼

🔗

📄 ▼

Press Alt 0 or Option 0 for help using the rich text editor with keyboard only.

☐ May contain sensitive data?

☐ May contain personally identifiable

DMPTool

UW provides access to [DMPTool](#), which has a number of funder templates.

- When you log in, **select University of Washington** as your research institution.
- Log in with your **NetID** and create your **profile**.
- Click "Create New DMP" **OR** copy an existing DMP.

Repository Selection

Selecting a Data Repository

- For some programs and types of data, NIH and/or Institute, Center, Office (ICO) policy(ies) and funding opportunities identify particular data repositories (or sets of repositories) to be used to preserve and share data.
 - For data generated from research subject to such policies or funded under such opportunities, researchers should use the designated data repository(ies).
- For data generated from research for which no data repository is specified by NIH, researchers are encouraged to select a data repository that is appropriate for the data generated from the research project. Be sure to consult the list of [desirable characteristics](#) and the following guidance:
 - Primary consideration should be given to data repositories that are discipline or data-type specific to support effective data discovery and reuse. For a list of NIH-supported repositories, visit [Repositories for Sharing Scientific Data](#).
 - If no appropriate discipline or data-type specific repository is available, researchers should consider a variety of other potentially suitable data sharing options:
 - Small datasets (up to 2 GB in size) may be included as supplementary material to accompany articles submitted to PubMed Central ([instructions](#) [↗](#)).
 - Data repositories, including [generalist repositories](#) [↗](#) or institutional repositories, that make data available to the larger research community, institutions, or the broader public.
 - Large datasets may benefit from cloud-based data repositories for data access, preservation, and sharing.

See [Repositories for Sharing Scientific Data](#) for a listing of NIH-supported data repositories.

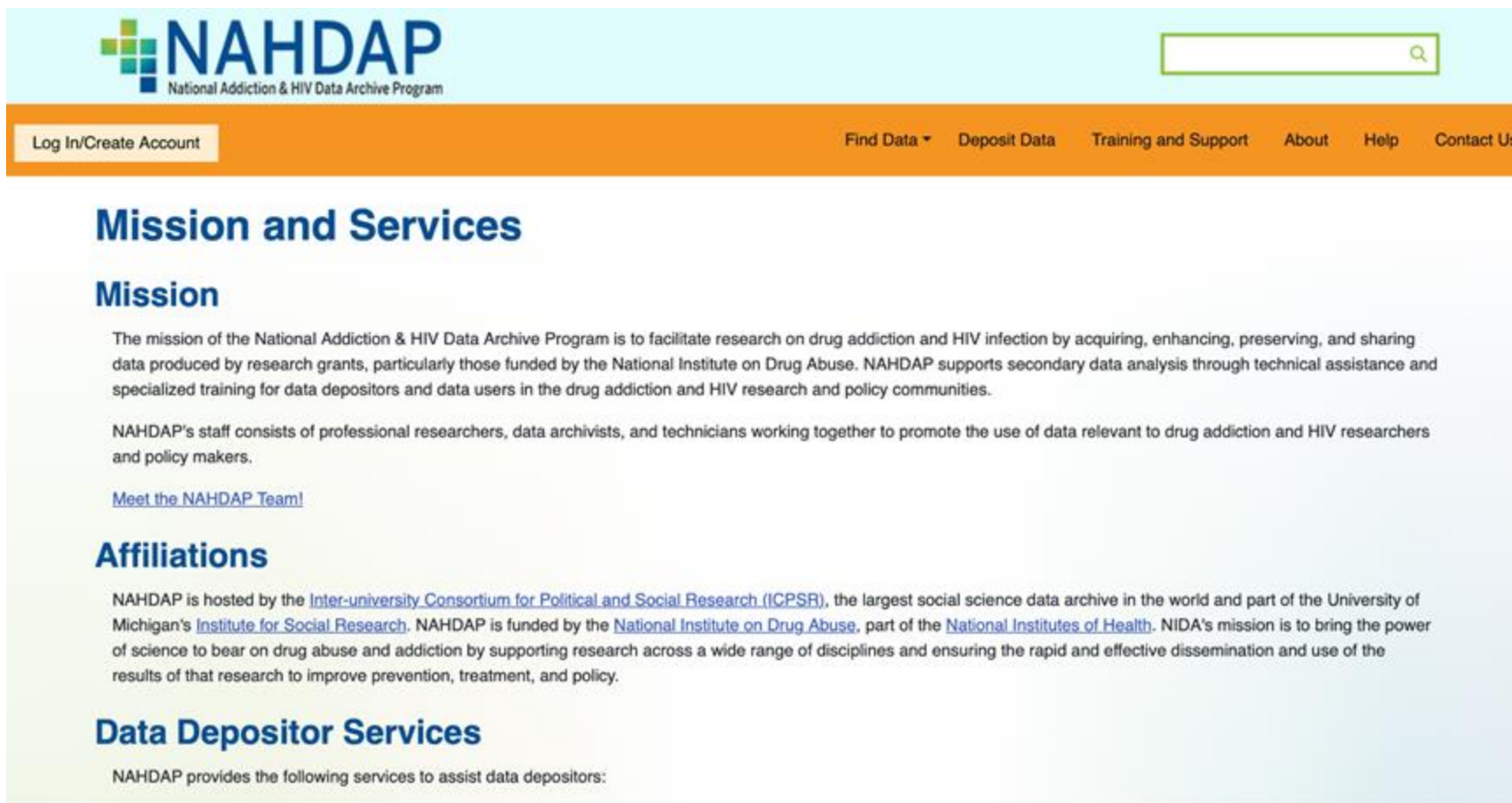
Desirable Characteristics for All Data Repositories



Where to Share

- **National repositories** like the [National Center for Biotechnology Information](#) or one from the NIH [list of approved sharing repositories](#)
- **Community-sponsored websites** like the [Neuroscience Information Framework](#)
- An **generalist repository**, such as UW's [Dryad](#) or [QDR](#)

Example repository: [NAHDAP](#)



The image shows the header and mission statement of the NAHDAP website. The header features the NAHDAP logo on the left, a search bar on the right, and a navigation bar with links for 'Log In/Create Account', 'Find Data', 'Deposit Data', 'Training and Support', 'About', 'Help', and 'Contact Us'. Below the header, the 'Mission and Services' section is displayed, starting with a 'Mission' sub-header. The mission text describes the program's goal to facilitate research on drug addiction and HIV infection by acquiring, enhancing, preserving, and sharing data. It also mentions the staff's role in promoting the use of data relevant to drug addiction and HIV researchers and policy makers. A link to 'Meet the NAHDAP Team!' is provided. The 'Affiliations' section follows, stating that NAHDAP is hosted by the Inter-university Consortium for Political and Social Research (ICPSR) and funded by the National Institute on Drug Abuse, part of the National Institutes of Health. The 'Data Depositor Services' section is partially visible at the bottom.

NAHDAP
National Addiction & HIV Data Archive Program

Log In/Create Account

Find Data ▾ Deposit Data Training and Support About Help Contact Us

Mission and Services

Mission

The mission of the National Addiction & HIV Data Archive Program is to facilitate research on drug addiction and HIV infection by acquiring, enhancing, preserving, and sharing data produced by research grants, particularly those funded by the National Institute on Drug Abuse. NAHDAP supports secondary data analysis through technical assistance and specialized training for data depositors and data users in the drug addiction and HIV research and policy communities.

NAHDAP's staff consists of professional researchers, data archivists, and technicians working together to promote the use of data relevant to drug addiction and HIV researchers and policy makers.

[Meet the NAHDAP Team!](#)

Affiliations

NAHDAP is hosted by the [Inter-university Consortium for Political and Social Research \(ICPSR\)](#), the largest social science data archive in the world and part of the University of Michigan's [Institute for Social Research](#). NAHDAP is funded by the [National Institute on Drug Abuse](#), part of the [National Institutes of Health](#). NIDA's mission is to bring the power of science to bear on drug abuse and addiction by supporting research across a wide range of disciplines and ensuring the rapid and effective dissemination and use of the results of that research to improve prevention, treatment, and policy.

Data Depositor Services

NAHDAP provides the following services to assist data depositors:



Deposit Data

DEPOSIT DATA

GOT DATA? We Want To Hear From You

WHY SHARE DATA WITH NAHDAP?

PREPARING DATA FOR DEPOSIT?

IRBS AND DATA SHARING

NAHDAP's mission is to preserve deposited files in perpetuity. Our online deposit form provides a secure upload of files and establishes the terms and conditions of the data transfer.

Data Deposit Form



To complete a data deposit, click on the "Data Deposit Form" button above. Please make sure that "NAHDAP" appears in the "Archive" drop-down menu so that your data is archived with us. Most fields are self-explanatory but contact us at nahdap@icpsr.umich.edu if you have any questions.

NAHDAP follows ICPSR's standard model for preparing data. Information about [how ICPSR prepares data for public release](#) is

Dryad

[Who we are](#)[What we do](#)[Join us](#)[Help ▾](#)[Login](#)

Trusted, simple, community driven

Explore data

[Search](#)

Photo by [USGS](#) on [Unsplash](#)

Over 15 years serving the community with easy,
dependable data sharing solutions.
[Learn about membership](#)

[Submit now](#)



UNIVERSITY LIBRARIES
UNIVERSITY of WASHINGTON

Dryad

1. Login with your ORCID ID

2. Describe your Data

3. Data Types

Dryad accepts many different types of data:

- If submitting human subjects data, makes sure your data is anonymized and follows both legal and ethical guidelines.
- All data deposited with Dryad must be complete, and open to the public.
- All data must be compatible with the **Creative Commons Zero license**. Find out more about Creative Commons licenses [here](#).

Dryad

4. Data Curation

Dryad curates (reviews) and then publishes your data. Select the *Private for Peer Review* check box if you want to suspend data publication until after the peer review process is complete.

5. Data Publication

Dryad notifies you once your data is published and provides you with a permanent DOI (digital object identifier) you can use to cite your data. You can update and re-version your data set at any time.

What to Share

- Raw data
- Analyzed data
- PII/human subjects
- Anonymization

Data preparation

- [Protecting Participant Privacy When Sharing Scientific Data](#) (NIH)
- [Data Enhancement](#) (ICPSR)
 - Dealing with [confidential information](#)
- [ICPSR Guide to Social Science Data Preparation and Archiving](#)
 - long (texty) page with 11 steps to data prep

Extra Credit

- [UW's OA policy](#)
- [Open Access @ UW](#)
- [How do I make my articles open?](#)

Related: [UW Reproducibility and Open Research group](#)

Thank you

Jenny Muilenburg

jmuil@uw.edu

[Link to slides](#)

[Support for Authors](#)

