

Research Data Management Resources at Northwestern

Date: August 20, 2025

Tobin Magle

Lead Data
Management
Specialist

Research Computing
& Data Services

Matt Carson

Head of Data
Management &
Technology

Galter Health
Science Library

Sara Gonzales

Senior
Data
Librarian

Galter Health
Sciences Library

Dieyun Song

Research Data
Management
Librarian

University
Libraries

Today we'll cover...

- Why is data management a hot topic?
- When should I be thinking about data management?
- What resources does Northwestern provide?
- How can I get help?

The Nelson Memo



EXECUTIVE OFFICE OF THE PRESIDENT
OFFICE OF SCIENCE AND TECHNOLOGY POLICY
WASHINGTON, D.C. 20502

August 25, 2022

MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES

FROM:

Dr. Alondra Nelson

A handwritten signature in blue ink that reads "Alondra Nelson".

Deputy Assistant to the President and Deputy Director for Science and Society
Performing the Duties of Director
Office of Science and Technology Policy (OSTP)

SUBJECT:

Ensuring Free, Immediate, and Equitable Access to Federally Funded Research

New requirements for researchers

Goal: Make federal research findings “as open as possible”

- Make all federally funded, peer-reviewed **publications freely available and publicly accessible** without embargo after publication
- Require **data management and sharing plans** with funding applications
- Make **scientific data underlying peer-reviewed scholarly publications freely available and publicly accessible** by default at the time of publication

A photograph of the Northwestern University archway, a large ornate metal structure with a crest at the top, set against a background of green trees. The archway is partially obscured by a semi-transparent grey overlay containing text.

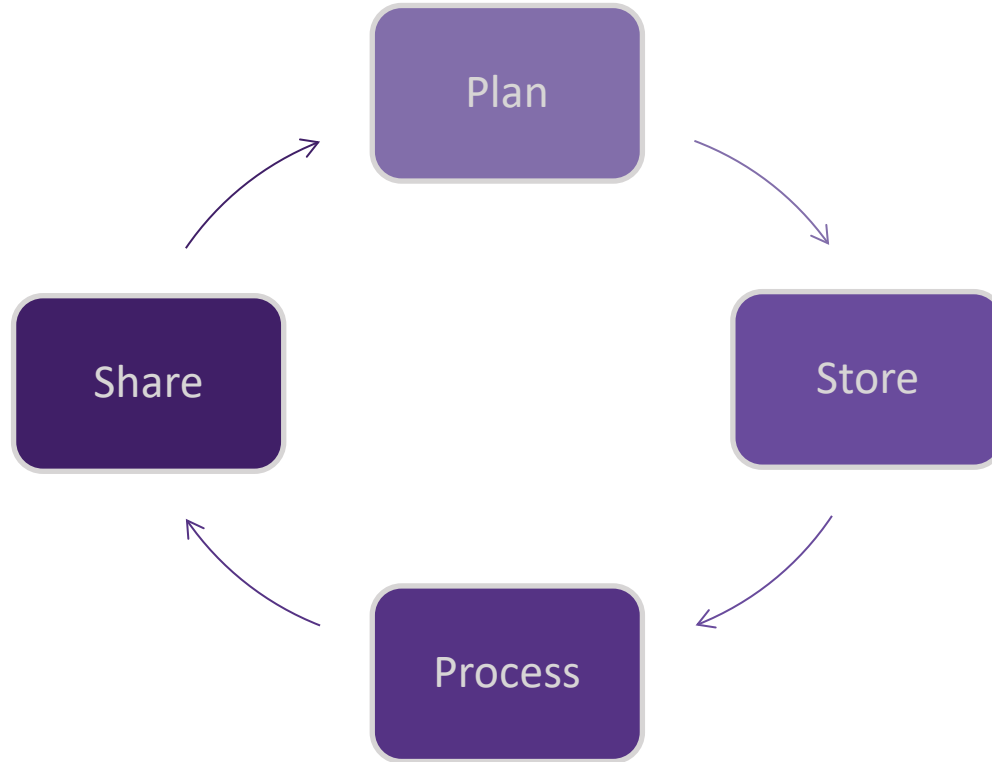
Research data management support at Northwestern University

A collaboration among Research Computing and Data Services, Galter Health Sciences
Library and University Libraries

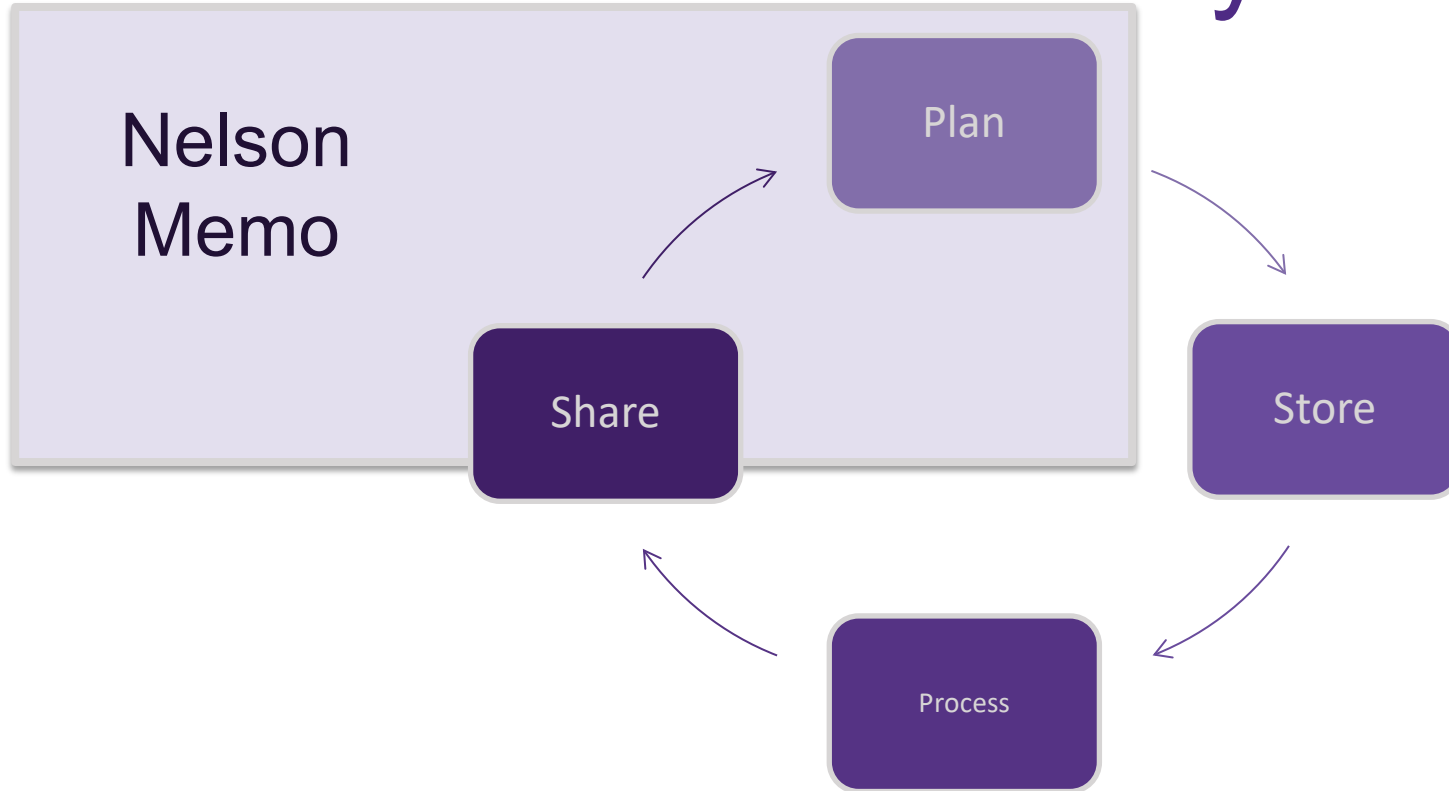
web: [https://www.it.northwestern.edu/departments/it-services-
support/research/data-storage/](https://www.it.northwestern.edu/departments/it-services-support/research/data-storage/)

contact: researchdata@northwestern.edu

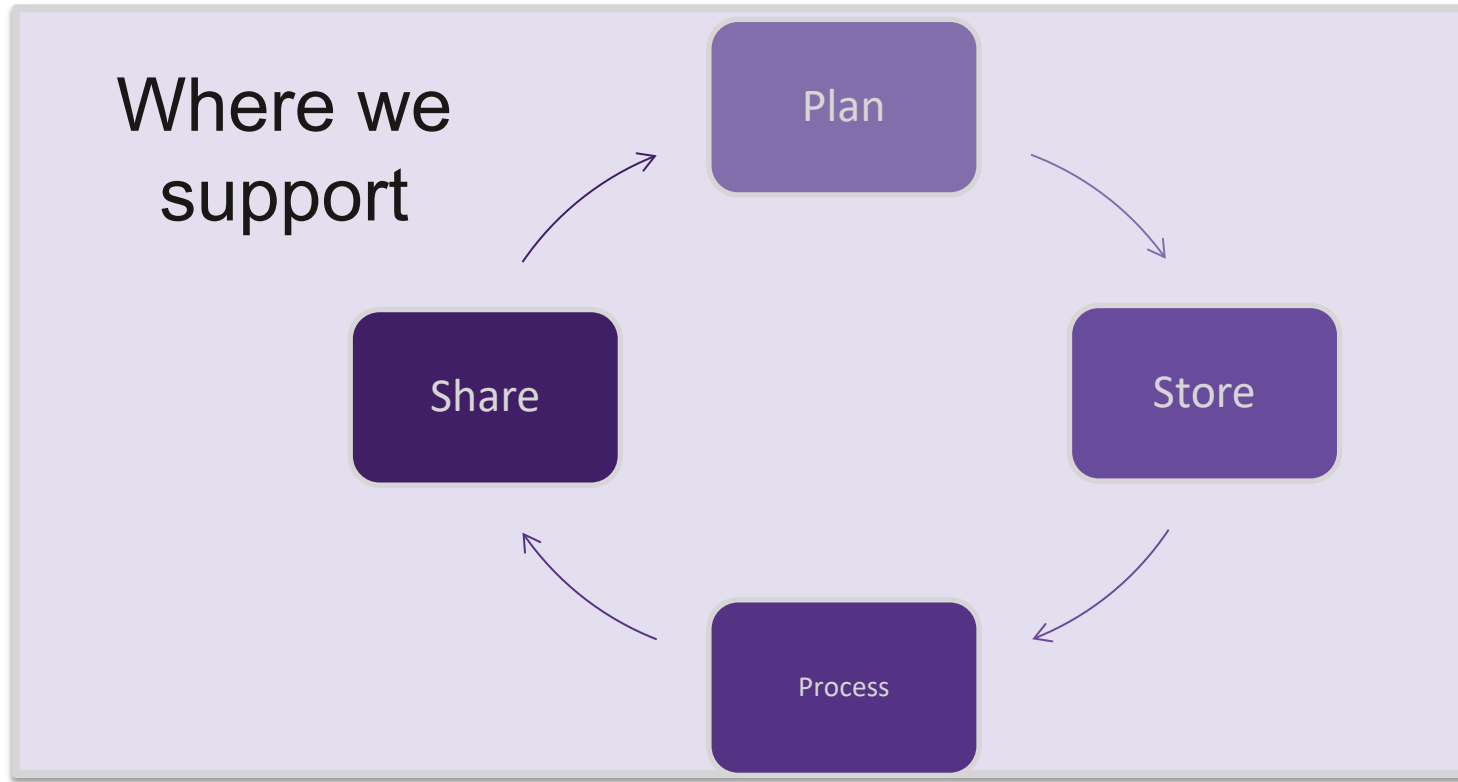
The Data Lifecycle



The Data Lifecycle



The Data Lifecycle



Highlighted Resources

Plan

- DMPTool

Store

- Storage Options
- Security

Process

- Globus
- Starfish

Share

- Arch
- Prism
- GREI

Plan:

Writing and Getting Help with Data
Management and Sharing Plans

What is a Data Management and Sharing Plan?

A description of how to describe,
preserve and share research data

Elements of a DMSP

1. Inventory of data to be produced
2. Description of what it is and how to use it
3. A plan for preserving and providing access
4. How will you make sure you follow your plan?

Funders that require DMSPs

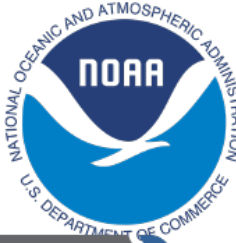
GORDON AND BETTY
MOORE
FOUNDATION



National
Science
Foundation



U.S. Department of Defense



INSTITUTE of
Museum and Library
SERVICES



NATIONAL
ENDOWMENT
for the **ARTS**
arts.gov

NIST



National Institutes
of Health

 **USGS**
science for a changing world



ALFRED P. SLOAN
FOUNDATION

Each Agency is different

What might vary?

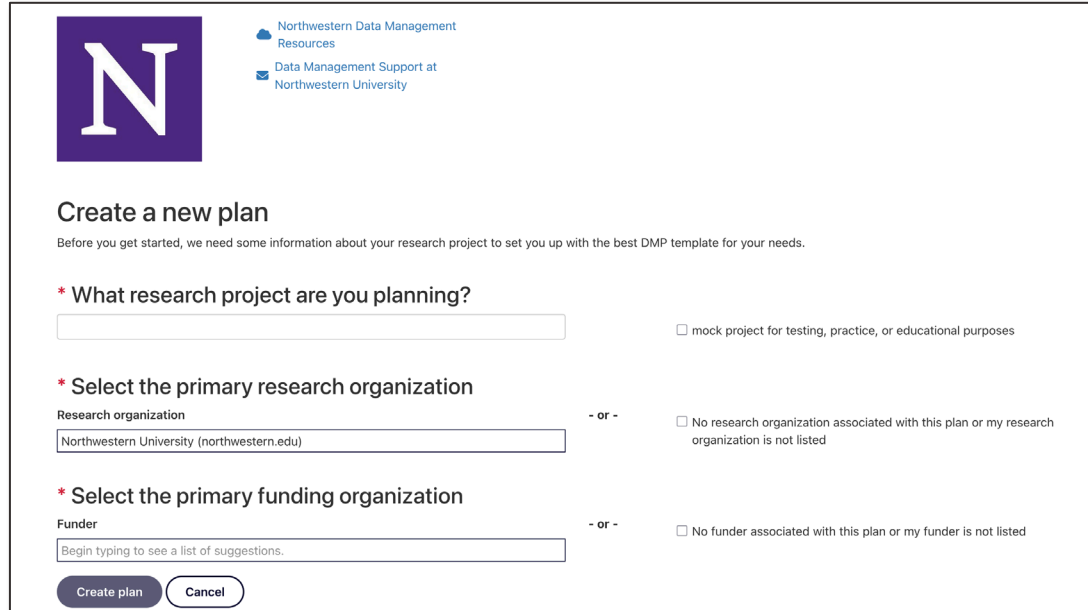
- Data Management plan format
- Rules for budgeting data management expenses
- Requirements for data sharing
- Progress reporting requirements
- Enforcement of data sharing policies

DMPTool.org: A DMSP Resource

Sign-in Steps

- Go to DMPTool.org
- Enter your Northwestern email address in the "Sign in / Sign up" box and click "Continue"
- After the prompt "Your address is associated with: Northwestern University (northwestern.edu)," click "Sign in with Institution to Continue"
- Enter your Northwestern credentials and click "Sign In"

dmptool.org/plans/new



The screenshot shows the 'Create a new plan' page on DMPTool.org. At the top left is the Northwestern University logo (a purple square with a white 'N'). To its right are two links: 'Northwestern Data Management Resources' and 'Data Management Support at Northwestern University'. The main heading is 'Create a new plan', followed by a subtext: 'Before you get started, we need some information about your research project to set you up with the best DMP template for your needs.' The form contains three required sections, each marked with an asterisk:

- * What research project are you planning?**: A text input field. To its right is a checkbox labeled 'mock project for testing, practice, or educational purposes'.
- * Select the primary research organization**: A dropdown menu labeled 'Research organization' with 'Northwestern University (northwestern.edu)' selected. To its right is a checkbox labeled 'No research organization associated with this plan or my research organization is not listed'.
- * Select the primary funding organization**: A dropdown menu labeled 'Funder' with 'Begin typing to see a list of suggestions.' as the selected option. To its right is a checkbox labeled 'No funder associated with this plan or my funder is not listed'.

At the bottom of the form are two buttons: 'Create plan' and 'Cancel'.

Writing a DMSP with DMPTool.org

Test

Project Details

Collaborators

Write Plan

Research outputs

Request feedback

This plan is based on the "NIH-Default DMSP" template provided by National Institutes of Health (nih.

expand all | collapse all

0/12

+ 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)

- After selecting organization, funder, and template type, fill in Project Details (only a Title is required) and add Collaborators (optional)
- The template type will show various section of the Plan to open and fill out. Utilize the funder- and institution-specific guidance included
- Finalize and Download the plan, then attach the PDF to your application. Feedback is available on plans from us

Sample plans available from sharing.nih.gov

Sample Plans

NIH has provided sample DMS Plans as examples of how a DMS Plan could be completed in different contexts, conforming to the elements described above. These sample DMS Plans are provided for educational purposes to assist applicants with developing Plans but are not intended to be used as templates and their use does not guarantee approval by NIH.

Note that the sample DMS Plans provided below may reflect [additional expectations](#) established by NIH or specific NIH Institutes, Centers, or Offices that go beyond the DMS Policy. Applicants will need to ensure that their Plan reflects any additional, applicable expectations (including from NIH policies and any ICO- or program-specific expectations as stated in the funding opportunity).



Sample Plan Description	NIH Institute or Center	Tags Filter
Clinical and/or MRI data from human research participants	NIMH	human_subjects controlled_access imaging video multimedia clinical_research phenotypic
Genomic data from human research participants	NIMH	genomics human_subjects controlled_access phenotypic
Genomic data from a non-human source	NIMH	genomics animal_models
Secondary data analysis	NIMH	human_subjects controlled_access secondary_data_analysis

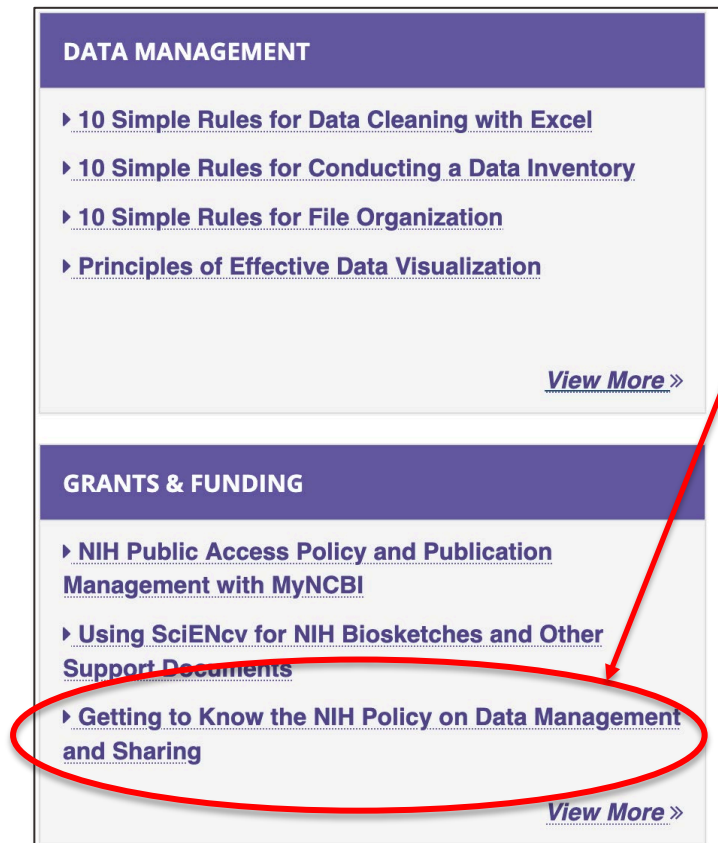
Trainings on writing DMSPs, plus other data management topics

NIH DMSP training:

- Offered on a regular basis through Galter Library
- Repository selection advice; real-world scenarios

Add'l Galter classes on RDM topics:

- Basic guidance:
 - [Best Practices in Research Data Management and Sharing](#)
 - [10 Simple Rules for File Organization](#)
 - [10 Simple Rules for a Conducting a Data Inventory](#)
- Skills classes:
 - [10 Simple Rules for Data Cleaning with Excel](#)
 - [Data Projects: Setup and Key Tips for Excel](#)
 - [Cleaning Spreadsheet Data with OpenRefine](#)
 - [Introduction to SQL for Clinical Researchers](#)
 - [Principles of Effective Data Visualization](#)



DATA MANAGEMENT

- ▶ [10 Simple Rules for Data Cleaning with Excel](#)
- ▶ [10 Simple Rules for Conducting a Data Inventory](#)
- ▶ [10 Simple Rules for File Organization](#)
- ▶ [Principles of Effective Data Visualization](#)

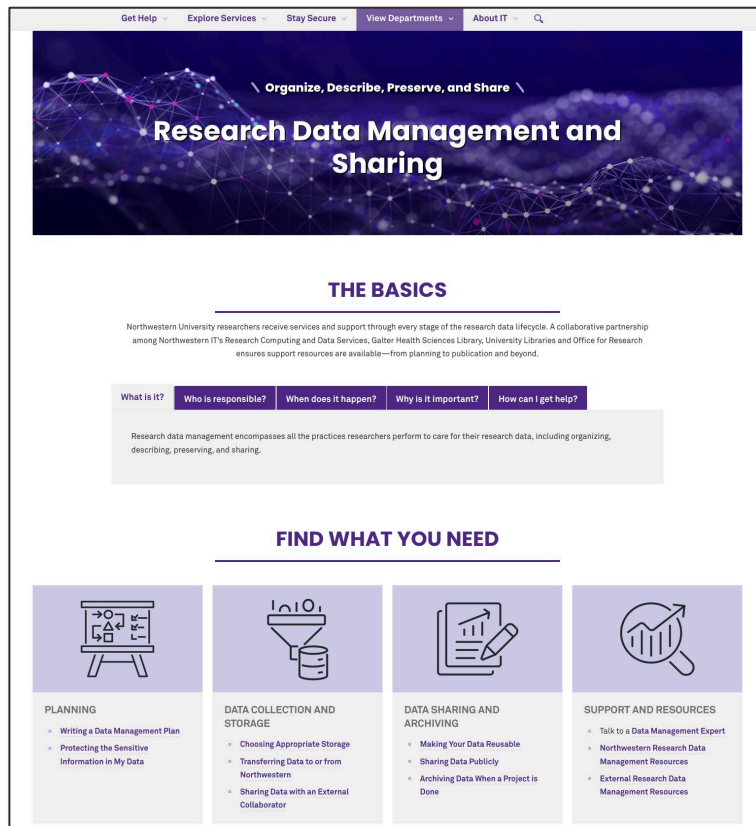
[View More »](#)

GRANTS & FUNDING

- ▶ [NIH Public Access Policy and Publication Management with MyNCBI](#)
- ▶ [Using SciENcv for NIH Biosketches and Other Support Documents](#)
- ▶ [Getting to Know the NIH Policy on Data Management and Sharing](#)

[View More »](#)

Get help with writing a data management and sharing plan



All campuses:

- Research Data Management and Sharing website
- ResearchData@northwestern.edu
- Use DMPTool Request Feedback feature

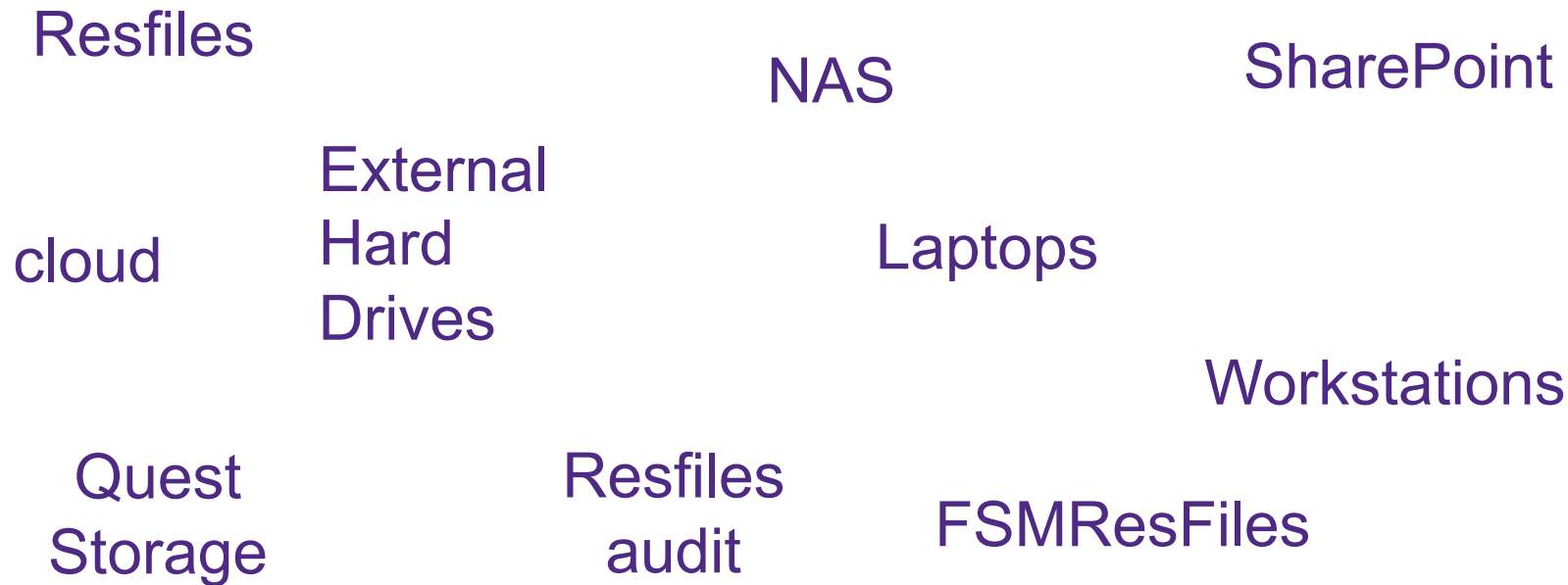
Feinberg School of Medicine:

- DataLab@northwestern.edu
- sara.gonzales2@northwestern.edu - NIH DMSPs and general RDM support
- NIH DMSP GalterGuide: libguides.galter.northwestern.edu/NIH-Data-Management-Plan

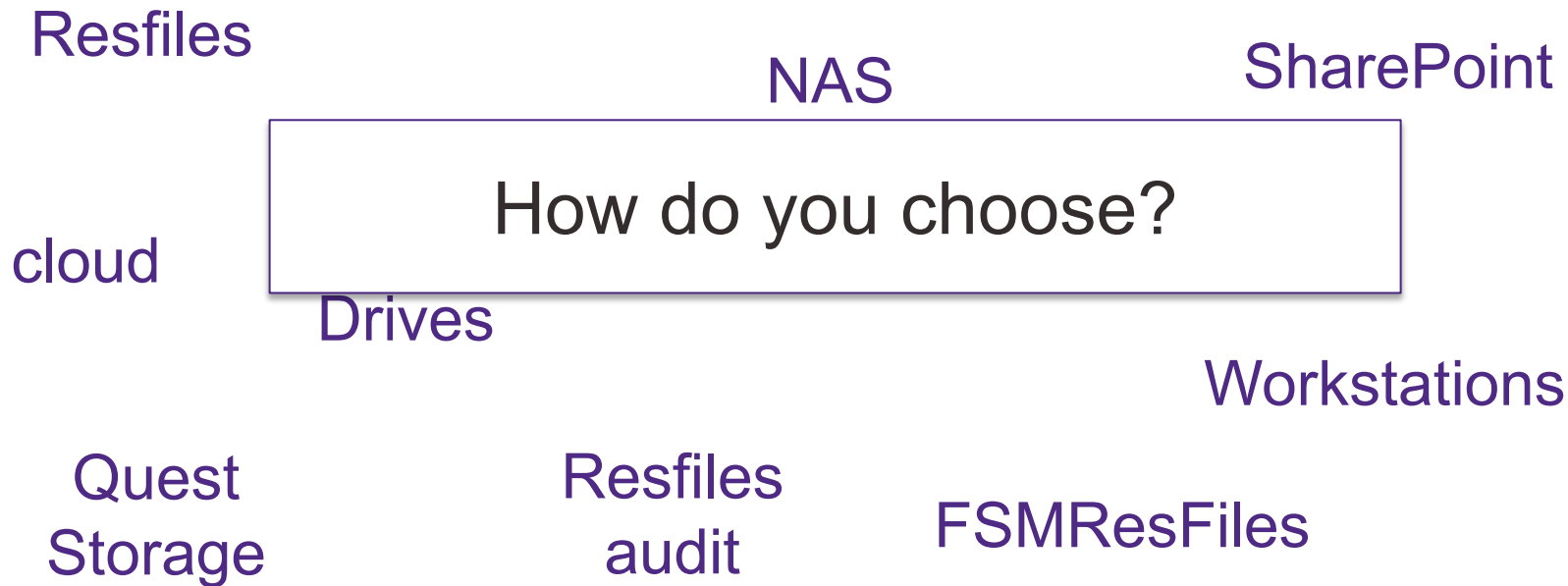
Store:

Choosing appropriate
research data storage

Storage Options



Storage Options



Things to consider

- **How much?** (data and cost)
- **Technical specs:**
 - Access control
 - File sizes
 - Transfer speeds
- **Security and compliance:**
 - **Regulations** that apply to your data
 - **Security controls** required
- **Data workflows:** moving your data efficiently between systems

Things to consider

- **How much?** (data and cost)
- **Technical specs:**
 - Access control
 - File sizes
 - Transfer speeds
- **Security and compliance:**
 - **Regulations** that apply to your data
 - **Security controls** required
- **Data workflows:** moving your data efficiently between systems

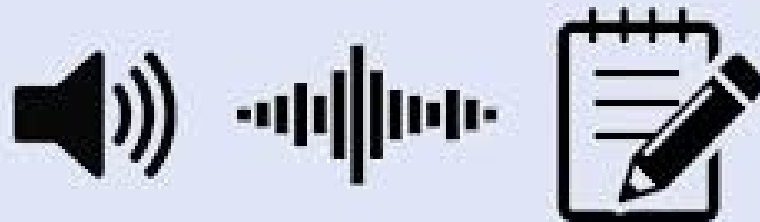
Security and compliance

Common requirements

- **Basic safeguarding:** for all research data – using Northwestern-provided services and managed devices satisfies many requirements
- **Complying with state and federal law**
 - **Common regulations:** HIPAA, FERPA, BIPA, NIST 180-171
 - All have distinct, overlapping requirements
- **Data use agreements (DUAs):** Agreements signed by Northwestern for researchers to receive data
 - must go through sponsored research
 - Have distinct requirements

Transcription services

- No approved vendors currently
- Options: Use auto transcription
 - on Zoom - record locally
 - on Teams - record to OneDrive
- Use transcription software local to your computer
 - Adobe Premiere - via Creative Suite license
 - Whisper – Open AI model



Help with sensitive data

- [Protecting the sensitive information in my data](#) (RDM website)
- [Choosing appropriate data storage](#) (RDM website)
- Ask for help:
 - Feinberg School Of Medicine: email fsm-help@northwestern.edu
 - All others: Email researchdata@northwestern.edu for help

Process:

Creating efficient data workflows

Data workflows



Globus data transfer tool

Move data between Northwestern storage platforms

- High speed transfers
- Easy to use web interface
- Resumes interrupted data transfers
- Verifies file integrity using checksums
- Automate data management tasks



Starfish

Manage data stored on RDSS, FSMResFiles, and Quest

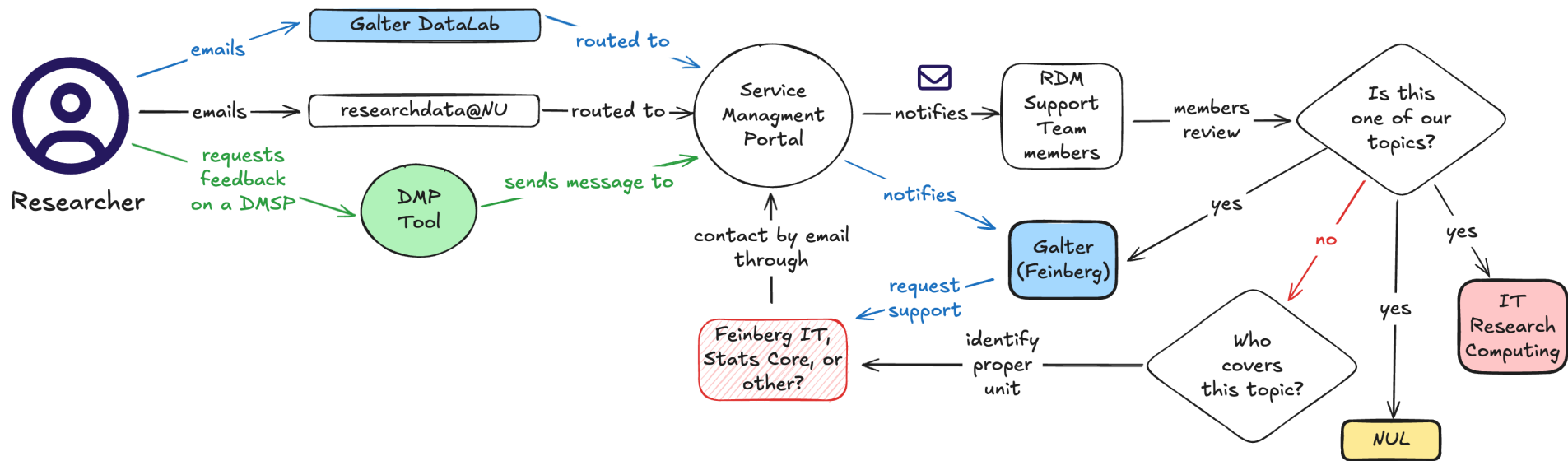
- Identify files that can be moved to cold storage
- Identify duplicate files that could be deleted
- Automate data workflows
- Generate custom reports



Share:

Complying with data
sharing requirements

Research Data Services Support Workflow



Galter DataLab Training & Support

Data Management

[Back To All Classes](#)

Available Classes

10 Simple Rules for Conducting a Data Inventory

In this "Coffee & Learn" session, enjoy some downstairs!) and join us for a 45-minute discussion.

10 Simple Rules for Data Cleaning with Excel

Join us for a 1 hour crash course in data cleaning.

Through this workshop, participants will learn:

- Simple and advanced filtering
- Creating histograms through the Data Analysis Toolpak
- Splitting multi-valued cells and trimming
- Data validation and custom lists

***Disclaimer:** This class is not a general Education Sciences Library & Learning Center does not

10 Simple Rules for File Organization

Join us for a 45-minute discussion on best practices for file organization.

Through this workshop, participants will:

- Learn best practices for file naming
- Learn best practices for digital folder organization
- Gain tips for adding searchable keywords

Best Practices in Research Data Management and Data Sharing

An introduction to basic concepts in research management plan requirements, data documentation, and data sharing.

Upon completion of this one-hour workshop, participants will:

- Understand and be able to apply best practices for data management
- Be familiar with basic tidy data best practices
- Be familiar with metadata best practices
- Understand and be able to locate online resources
- Be familiar with publication and data sharing

Grants & Funding

[Back To All Classes](#)

Available Classes

Getting to Know the NIH Policy on Data Management and Sharing

Mastering SciENcy for NSF Grant Applications

NIH Biosketch: Presenting Your Research in Your Grant Application

NIH, My NCBI and PMCIDs - Virtual Q & A "Office Hours"

NIH Public Access Policy and Publication Management with MyNCBI

Using SciENcy for NIH Biosketches and Other Support Documents

These classes provide guidance on common forms and compliance issues surrounding grant applications, with particular emphasis on NIH grants.

The screenshot shows the GitHub profile for Galter DataLab. The profile includes a bio: "Data-related consultation, training, and collaborative services for researchers in Northwestern University's Feinberg School of Medicine". It lists 4 followers and provides the website URL <https://galter.northwestern.edu/datalab> and email datalab@northwestern.edu. The pinned repositories are:

- galterdatalab.github.io**: Public repository for bridging the skills gap for translational researchers through education & training, consultation, and innovation. 1 star.
- crdm-training**: Public repository for a clinical data retrieval and management training program for researchers.
- dmp-tool-api-queries**: Public repository for scripts accessing the DMP Tool API. 1 star.
- zenodo-api-queries**: Public repository for Python scripts accessing the Zenodo's API. 1 star.
- intro-sql-clinical-researchers**: Public repository for an introduction to databases and SQL for clinical researchers. 1 star.
- foundations-sequence-similarity**: Public repository.
- principles-of-effective-data-viz**: Public repository for an introduction to concepts and best practices for data visualization. 1 star.
- intro-sql-clinical-researchers**: Public repository for an introduction to databases and SQL for clinical researchers. 1 star.
- dmp-tool-api-queries**: Public repository for scripts accessing the DMP Tool API.

Top languages: Jupyter Notebook, R, HTML, SCSS, TSQ.

Most used topics: r, rstudio, clinical-data, clinical-data-warehouse, data-management.

https://galter.northwestern.edu/subject_info/8
https://galter.northwestern.edu/subject_info/6
<https://github.com/galterdatalab>

Federal Policy Compliance Consultation

- New policy mandates and accelerated timelines have increased the need for support
- 'NIH Policies & Procedures Sessions: A Roadmap for Compliance' seminar
- NIH Public Access Policy and Publication Management with MyNCBI Galter class offered monthly and available by request for individuals or groups

Galter Health Sciences Library & Learning Center

[Galter Health Sciences Library & Learning Center](#) / [Galter Guides](#) / [NIH Announcements and Updates](#) / [Policy Change Notices](#)

NIH Announcements and Updates

This guide serves as a listing service for updates and announcements from the National Institutes of Health (NIH)

Home

Policy Change Notices

NIH News Releases

Executive Orders

NIH Public Access Policy

Describes the NIH Public Access Policy and compliance, including the process for obtaining PMCIDs

[What is the NIH Public Access Policy?](#)

[Details on the Policy and Potential Issues for Authors](#)

[Enforcement on Progress Reports](#)

[Outing PMCIDs for your Papers](#)

[Deposit Papers to NIHMS](#)


[Advice to Authors](#)

[NIH Policy FAQs](#)

[Resources](#)

[My NCBI My Bibliography for Grant + Publication Management](#)

Bioinformatics & Bioinformatics Librarian,
NIH Public Access Compliance Reporter



Pamela Shaw
she/her

[Email Me](#)

Contact:
Digital Systems Dept.
Galter Health Sciences Library
320 E Superior St.
Chicago, IL 60605

[NEW FAQs on Limit to NIH Proposals per Calendar Year](#)

The NIH has added new FAQs to its FAQ page on Use of Generative AI in Peer Review.

These new FAQs address the new policy on limiting the number of proposals per NIH applicant to six (6) per calendar year.

- The policy begins with applications submitted to the September 25, 2025 receipt date and beyond. Investigators who have submissions prior to September 25 that may put them over the limit of six proposals for calendar year 2026 will not have these proposals withdrawn.
- Applications where the investigator is the PI/PO or PI count toward the investigator's yearly limit.
- The policy applies to all "new, renewal, resubmission, or revision submissions. Administrative supplements, non-competing renewals, and Type 7 applications are not counted toward the maximum. The Policy applies to all activity codes except T activity codes and R13 Conference Grant Applications. Applications that are withdrawn after peer review count toward the maximum."

Please see NOT-OD-25-132 Supporting Fairness and Originality in NIH Research Applications for the details of the policy.

[NEW! July 30, 2025 - NIH Issues RFI for Public Input on APC Limits](#)

The NIH Office of Science Policy has just released a request for information (RFI - NOT-OD-25-138) on article publishing cost limits for NIH-funded manuscripts.

This notice is a follow-up to the announcement (directly below this section of this guide) to limit publishing costs for NIH-funded manuscripts.

Responses to the RFI will shape the limits for article publishing costs that are allowable to be charged to NIH awards, and whether they will be capped per manuscript, per life of the award, or some combination of these limits (or NO allowable publishing costs at all). The NIH is proposing five options for respondents to consider.

Please take some time to visit the proposed APC limit notice and consider submitting a response. The link to the response document is in the notice.

[NEW \(July 8, 2025\) The NIH Announces a Crack-Down on Article Publishing Charges](#)

In a July 8, 2025 NIH news item, the NIH announced they will address the issue of high article publishing charges (APCs) charged by journal publishers:

... As part of its ongoing commitment to scientific transparency and responsible stewardship of taxpayer dollars, the National Institutes of Health (NIH) today announced plans to implement a new policy that will cap how much publishers can charge NIH-supported scientists to make their research findings publicly accessible.

...

To address this imbalance, NIH will introduce a cap on allowable publication costs starting in Fiscal Year (FY) 2026, ensuring that publication fees remain reasonable across the research ecosystem. The policy aims to curb excessive APCs and ensure the broad dissemination of research findings without unnecessary financial barriers.

There is no new policy yet linked to this announcement.

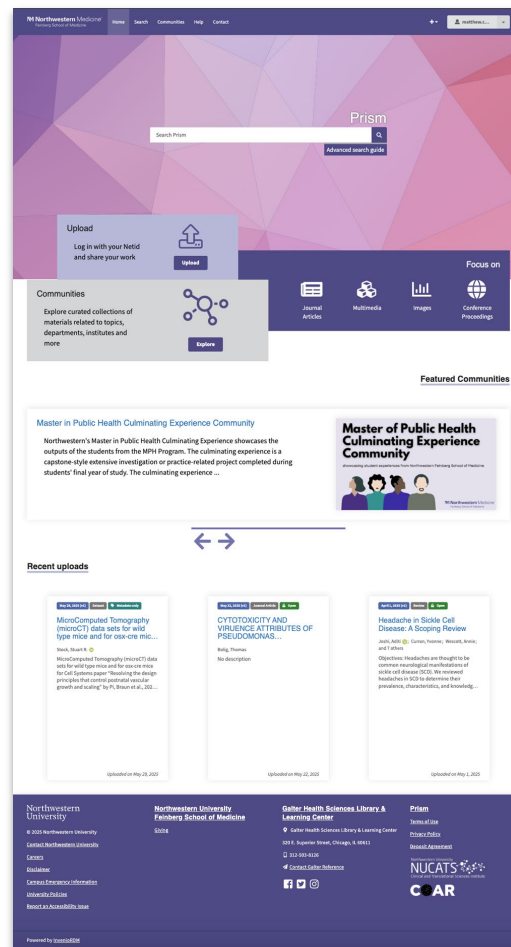


National Institutes
of Health



Prism: Feinberg's IR

- Launched in April 2023
- Maintained, managed, and operated by Galter Health Sciences Library & Learning Center
- User base: Faculty, staff, and students of Feinberg School of Medicine



Communities

M Northwestern Medicine
Feinberg School of Medicine

Home Search Communities Help Contact Search records...

COVID-19 COMMUNITY
Topic

Records Requests Members Settings Curation policy About

74 results found

Versions

☐ View all versions

Access status

☐ Open 71
☐ Restricted 3

Resource types

☐ Articles 74

File type

October 30, 2020 (v1.0.0) Review Open

Coronavirus Disease-19: An Interim Evidence Synthesis of the World Association of Immunological Disorders (Waidid)

Abu-Raya, Bahaa; Migliori, Giovanni Battista; O'Ryan, Miguel; and 27 others

No description

Part of COVID-19 Community

Uploaded on March 30, 2023

November 11, 2020 (v1.0.0) Review Open

Coping With Stress and Burnout Associated With Telecommunication and On

Mheidly, Nour; Fares, Mohamad Y.; Fares, Jawad Youssef

No description

Part of COVID-19 Community

Uploaded on March 30, 2023



Generalist Repository Ecosystem Initiative (GREI)

The mission of GREI is to establish a common set of capabilities, services, metrics, and social infrastructure; raise general awareness and facilitate researchers to adopt FAIR principles to better share and reuse data.

This initiative will further enhance the biomedical data ecosystem and help researchers find and share data from NIH-funded studies in generalist repositories.

Goals of the Generalist Repository Ecosystem Initiative



GREI is Supporting Transparency and Compliance Through Coopetition

Objectives



- Catalog Use Cases Supported
- Implement Open Metrics
- Prepare Training Materials
- Conduct Outreach
- Commit to 'Coopetition'
- Implement Best Practices for Data Repositories
- Support Discovery of NIH Funded Data
- Adopt Consistent Metadata Models
- Facilitate QA/QC
- Connect Digital Objects

Participating Repositories



Contact

GREI@nih.gov • <https://bit.ly/3en3rRy>

Repository Selection Flowchart:

<https://doi.org/10.5281/zenodo.11105430>

Generalist Repository Comparison Chart:

<https://doi.org/10.5281/zenodo.3946719>

Data Management Plan Guidance:

<https://doi.org/10.5281/zenodo.14278957>

Data Submission Checklist:

<https://doi.org/10.5281/zenodo.14278906>

Data Library Project

Data Library / Research Data Management

Develop a comprehensive Data Library platform framework that includes technical infrastructure, operational partnership, and policy compliance. Develop a research data management service model, including repositories for research data and curated datasets in compliance with federal data-sharing policies, that is scalable using technologies like machine learning and generative AI that both highlights the Libraries' expertise and reduces redundancies in services offered across campus.

How is it different?

1. Focused on use not preservation

Data Library is different than Arch, Dryad & Prism (preservation & compliance)

Examples – [WRDS](#), [Dataverse Project](#) & [KRS](#)

2. Created with faculty & student input

3. Curated

Ready to download and use – no cleaning necessary

4. Flexible

Data

Data + models

Data + models + code

1. Partnership

Collaborative team – NUL (AI, IT, AP) & NUIT RCDS

Working with Kellogg Research Support

Interviews with faculty and graduate students

2. Curated datasets

Identify already purchased datasets & data from other research groups (KRS & NUIT RCDS)

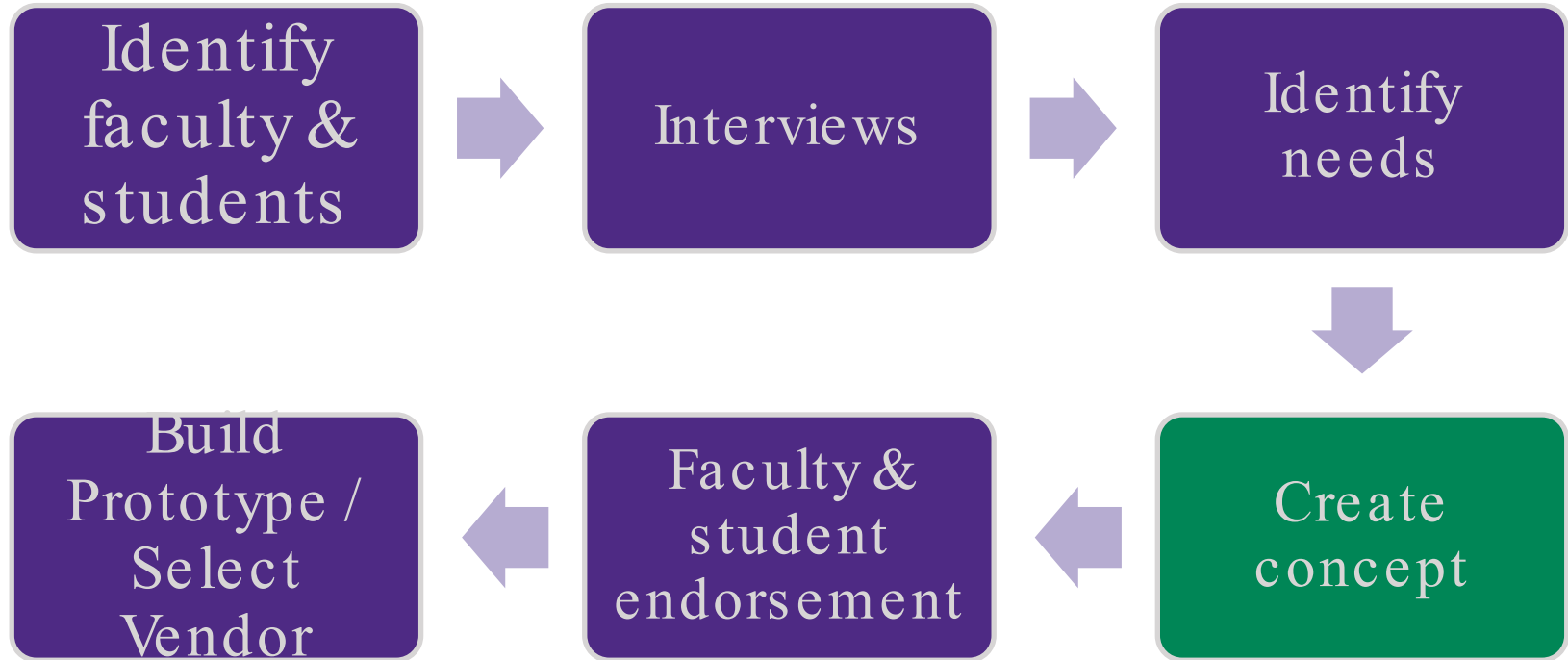
Datasets used by DSRDM – Reddit, Semantic Scholar, YouTube Comments, JStor

3. Reduce redundancy (see partnership)

4. Technical infrastructure

Scalable

Process // Phase 1



A stylized illustration of a radio tower with several signal waves emanating from its top, set against a purple gradient background.

Share Your Scholarship

Research and Data Repository

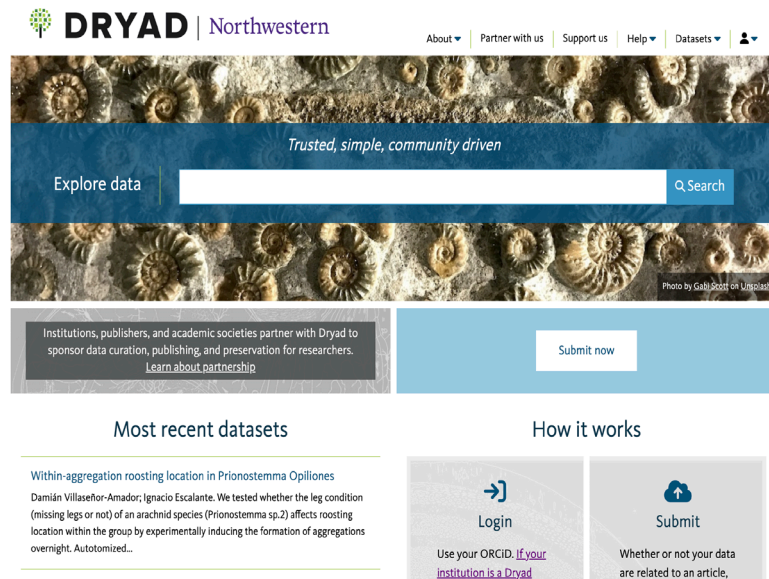
GET STARTED

Arch: Northwestern Institutional Repository

- ⑩ A generalist repository
- ⑩ Created and managed by Northwestern University Libraries
- ⑩ For all Northwestern faculty, staff, and students
- ⑩ Supports individual deposits up to 3GB
- ⑩ Contact Research Data Management Librarian for individual deposits between 3-5GB for technical support

Dryad: A Generalist Repository

- ⑩ For datasets larger than 10GB by all Northwestern faculty, staff, and students
- ⑩ Dryad supports and approves each individual publication
- ⑩ Paid and supported by Northwestern University Libraries
- ⑩ Starting from January 1, 2026, contact Research Data Management Librarian dieyun.song@northwestern.edu for publication funding support



How can I get help?

- Scheduled Trainings
- Custom Trainings
- Individual consultation

Scheduled Trainings

RCDS and University Libraries

Data Management, Data Science and Compute

- **Data Storage 101:** Understanding your options
- **Data Storage 102:** Time-Saving Data Organization Strategies
- Quest Orientation
- Genomics compute cluster
- R/Python
- Statistics
- AI for Researchers

Galter Health Science Library

Data management

- 10 simple rules for ...
- Best Practices in research data management and sharing
- Introduction to SQL Clinical Researchers
- Principles of Effective Data Visualization
- Prism: collect, share, and preserve the full spectrum of your work

Other Topics

- Bioinformatics & Biosciences
- Grants and Funding
- Research Dissemination and Impact

Custom Trainings

Email researchdata@northwestern.edu to request customized training for your research group

Need individualized help?

In person (Evanston)

- [Research Computing and Data Drop-in support](#)

Email

- researchdata@northwestern.edu
- datalab@northwestern.edu (Feinberg Only)

NU / FSM Policies and Support Links

- Northwestern IT: [Choosing Appropriate Data Storage](#)
- [FSM Data Storage Policy](#) (update coming soon)
- Includes Data De-identification guidance:
 - [FSM General Information Security Policy](#)
 - [FSM Information Security and Access](#)
 - [Research Privacy and Confidentiality](#)
- Research Use of NMEDW Data Policy:
 - [Research Use of Electronic Medical Record Data](#)
 - [Feinberg IT Policy Compliance](#)
- [Data Sharing & Storage Support](#)
 - Includes Sharepoint, Teams, OneDrive, Delve, Research Data Storage Service (RDSS)