



Research Data Services  
UNIVERSITY OF WISCONSIN-MADISON

We **love** your data too.

```
01000      10101
0001000100 0011110001
10001010101010010100101010
0000010000100100001011001010
0010101001010010100101010010
01001010101010101010101010
01101010100101010100100100
0110101010100100100100100
100110101001
010110
10
```



data management plans

personal consultations

education & training

[researchdata.wisc.edu](https://researchdata.wisc.edu) & @UWMadRsSchSvc

defining terms.



**Data management** → a set of practices to handle information collected and created during research; the compilation of many small practices that make data easier to find, easier to understand, less likely to be lost, and more likely to be usable in the future

**Data sharing** → broad umbrella term for the scholarly sharing of data between researchers, both:

- collaborative sharing between researchers within a research project
- broadly refers to sharing data through a repository (to make open or public!)

**Open data** → research data that is available for download via the internet without restriction or cost and licensed in such a way that allows for re-use and re-distribution by anyone for any purpose.

- Openness has both technical and legal requirements to it

**Public access to data** → federal funding agency mandates that require making research data underlying publications from funded projects available to anyone for free ~~or at nominal cost~~ through a repository

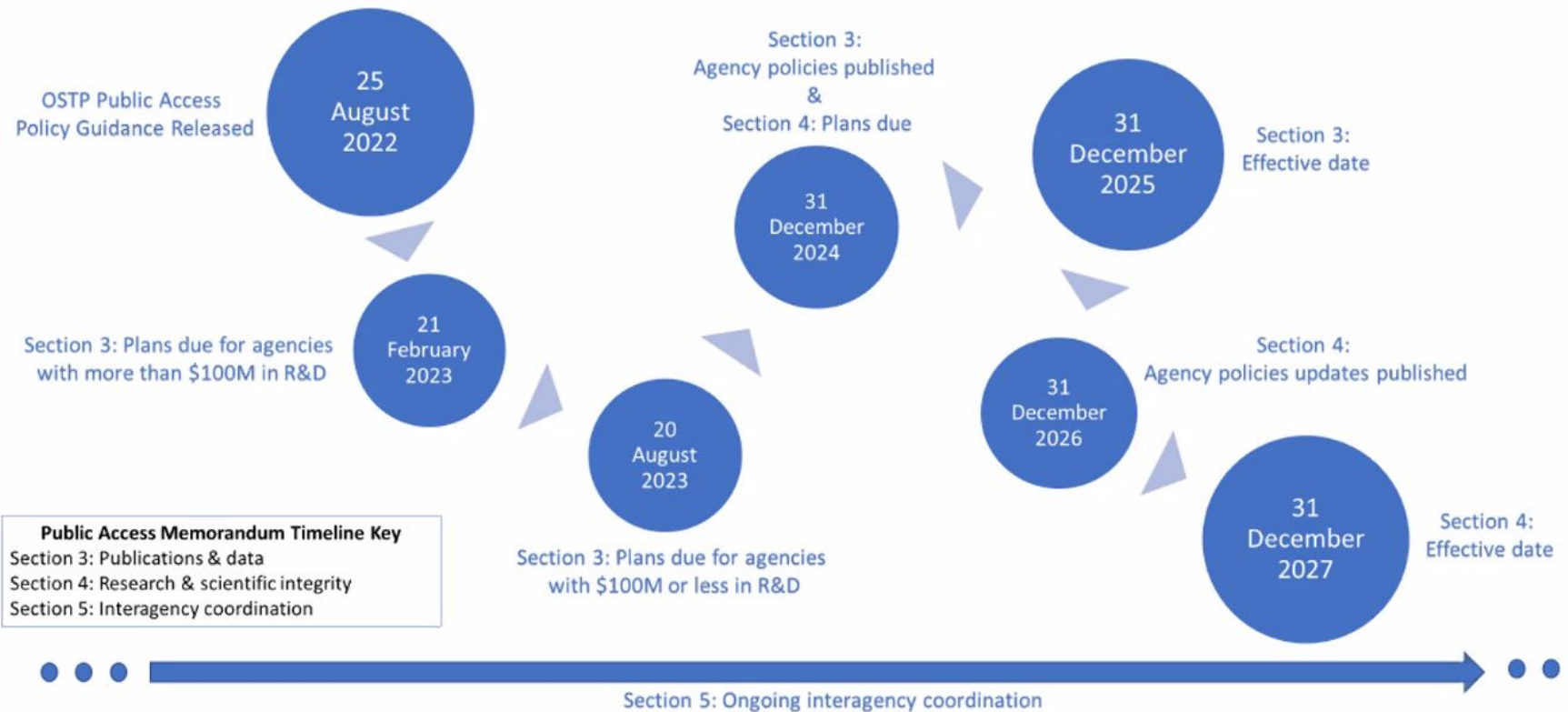
- **Public access  $\neq$  open data**
- Open or moderated, restricted access data can be compliant
- Data should be shared as ethically appropriate for the dataset.
  - 'as open as possible, as closed as necessary'.

**Data repository** → Managed third party platform/service with policies for collection, retention, etc. for research data that assumes stewardship of your deposited data, makes sure it is available, and conforms to best practices

- Frequently called a data archive, but is not archival storage solutions in a common IT sense
  - Think repository = making your data available to others
  - Archiving = storing it away long-term, little intention to access it
- Agencies are signaling that chosen repositories should be existing repositories that comply with the [Desired Characteristics of Data Repositories for Federally Funded Research](#)

# new federal policy + past policy

## Timeline for 2022 OSTP Public Access Policy Guidance



# NIH rationale?

- Equitable delivery of federally funded research results and data
- Foster public trust in research
  - reproducibility & transparency in research
- Maximize research participants' contributions
- Speed research/discovery, especially during crises like COVID

# NIH Data Management & Sharing Policy Overview

## Required for competing applications submitted after January 25, 2023

- Submit a 2-page plan outlining how scientific data from their research will be managed and shared
  - Scientific data: data of sufficient quality to validate and replicate research findings
- **Data should be shared as soon as possible, and no later than the time of an associated publication or end of performance period (whichever comes first)**
  - Data must be shared regardless of whether the data are used to support scholarly publications
- This plan represents the minimum requirements
  - NIH ICOs may expect more specificity in their plans
- Complete list of [NIH activity codes subject to the DMS Policy](#)
  - Review your FOA to determine if the DMS Policy applies to your application



# Elements to include in a DMP/DMSP

1. **Data Types:** Data to be preserved and shared (file types and sizes)
2. **Related Tools, Software, Code:** Tools and software needed to access & manipulate data, (open and/or proprietary)
3. **Common Data Standards:** Standards applied to data/metadata (protocols, data dictionaries, common data elements)
4. **Data Preservation, Access, Timelines:** Repository to be used, persistent unique identifiers used by repo, and when the data will be made available (when the work is published or the award/support period ends)
5. **Access, Distribution, Reuse Conditions:** Legal, technical or ethical considerations that may affect the extent of data sharing
6. **Oversight of Data Management:** Name and ORCID of Lead PI, a statement on broad compliance oversight from UW-Madison

# Should my data be shared?

Under element 5, a compelling rationale for limited sharing of some or all of the data produced can be explained. NIH examples include:

- informed consent will not permit or will limit the scope or extent of sharing and future research use
- existing consent prohibits sharing or limits the scope or extent of sharing and future research use
- privacy or safety of research participants would be compromised or place them at greater risk of re-identification or suffering harm, and protective measures such as de-identification and Certificates of Confidentiality would be insufficient
- explicit federal, state, local, or Tribal law, regulation, or policy prohibits disclosure
- restrictions imposed by existing or anticipated agreements

# Submission & Review of DMS Plans

- Submit DMS Plans and budget requests as part of the funding application or proposal.
- *Peer Review will not see or review DMS Plans, but will consider any related budget items.*
- NIH program staff will review the DMS Plan for acceptability and may request modifications prior to award as appropriate.
- Provide updates on data management and sharing activities in annual progress reports.
- If plans change over the course of the project, work proactively with NIH Program Officer to obtain review and approval of modifications.

# Budgeting for Data Management & Sharing

## What to include in the budget justification summary regarding DMS costs

- include the requested dollar amount and a brief justification of the proposed activities that will incur costs (no more than half a page)
  - curating data and developing supporting documentation (\*data activities)
  - local data management considerations
  - preserving and sharing data through established repositories

## Personnel effort included and budgeted as part of the DMSP

- salary and fringe benefits corresponding to the time it takes personnel to undertake data activities \*
- must be included in the single line item on the R&R Budget Form in section F. Other Direct Costs

**All allowable costs must be incurred during the performance period**

# generalist repository ecosystem initiative (GREI)

- Office of Data Science Strategy aims for GREI
  - “...further enhance the biomedical data ecosystem and help researchers find and share data from NIH-funded studies in generalist repositories.”
  - The 6 repositories will work together to “establish consistent metadata, develop use cases for data sharing, train and educate researchers on FAIR data and the importance of data sharing,”
  - Increase discoverability of data, improve reproducibility and reuse of data
- NIH is working with 6 data repositories
  - Dryad (UW-Madison is a member!)
  - Dataverse
  - Figshare
  - Mendeley Data
  - Open Science Framework (OSF)
  - Vivli



# DMPTool

Build your Data Management Plan



## BuckyBackup Archive

Plan

## ResearchDrive

Closeout

Collect & store data



Publish research artifacts

Analyze & visualize data



**\* What research project are you planning?****\* Select the primary research organization**

Research organization

**\* Select the primary funding organization**

Funder

[Project Details](#)[Collaborators](#)[Write Plan](#)[Research outputs](#)[Request feedback](#)[Finalize](#)[Download](#)

This plan is based on the "NIH-GEN DMSP (2023) " template provided by National Institutes of Health (nih.gov) - (ver: 8, pub: 2023-01-24).

[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 \(0 / 1\)](#)

+

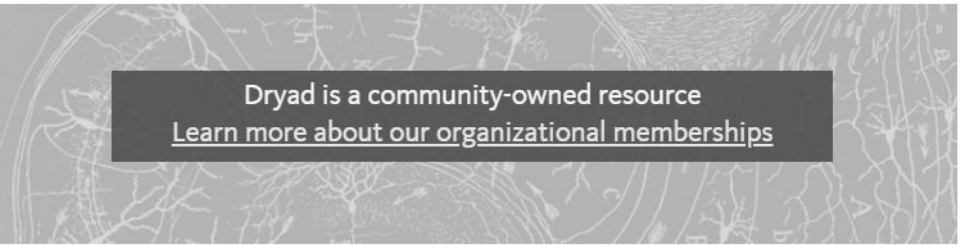
# Dryad



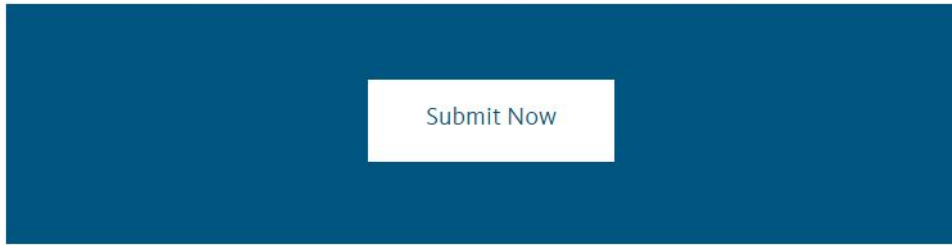
[Admin](#) | [Explore Data](#) | [About](#) ▼ | [Help](#) ▼ | [My Datasets](#) | [Logout](#)



*for your research data*



Dryad is a community-owned resource  
[Learn more about our organizational memberships](#)



Submit Now

<https://datadryad.org/>



## Simulated GOES ABI brightness temperatures from the 2019 OU-CAPS Hazardous Weather Testbed FV3-LAM simulations

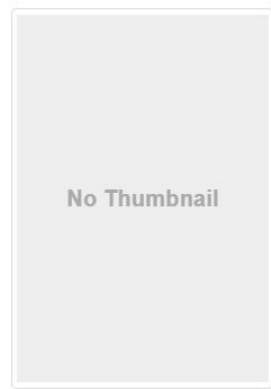
To access the dataset, follow the permanent link below or copy and past the following into your browser: [https://app.globus.org/file-manager?origin\\_id=86401693-5974-4013-b498-eb4484e08eb4&origin\\_path=%2FSSEC%2FCIMSS%2F](https://app.globus.org/file-manager?origin_id=86401693-5974-4013-b498-eb4484e08eb4&origin_path=%2FSSEC%2FCIMSS%2F)

Subject  
Brightness temperature  
Remote sensing

Related Material/Data  
[https://app.globus.org/file-manager?origin\\_id=86401693-5974-4013-b498-eb4484e08eb4&origin\\_path=%2FSSEC%2FCIMSS%2F](https://app.globus.org/file-manager?origin_id=86401693-5974-4013-b498-eb4484e08eb4&origin_path=%2FSSEC%2FCIMSS%2F)

DOI  
<https://doi.org/10.21231/6s0z-zt47>

Citation  
Griffin, Sarah, Jason Otkin, and Timothy Supinie. "Simulated GOES ABI Brightness Temperatures from the 2019 OU-CAPS Hazardous Weather Testbed FV3-LAM Simulations." Cooperative Institute for Meteorological Satellite Studies, University of Wisconsin-Madison, 2021.



File(s)  
Dataset README (1.816Kb)

Date  
2021

[minds.wisconsin.edu](https://minds.wisconsin.edu)

Search

Search MINDS@UW  
 This Collection

- BR
- FILE MANAGER
- BOOKMARKS
- ACTIVITY
- ENDPOINTS
- GROUPS
- Q A

### File Manager

Collection

Path

select all   up one folder   refresh list

NAME ▾

- Simulated-GOES-ABI-brightness-temperatures
- Simulated-NCEP-GFS-hyperspectral-infrared-sounder



## Find Data

 [Search](#)[view all](#)[search tips](#)

### Browse

[Topics / Series / Thematic data collections](#)

- [List studies for which online analysis is available](#)
- [List self-published data, including replication datasets](#)
- [List studies with learning guides](#)

#### New/Updated Data Releases:

- [In the last week](#)

### Statistics



16,575 studies

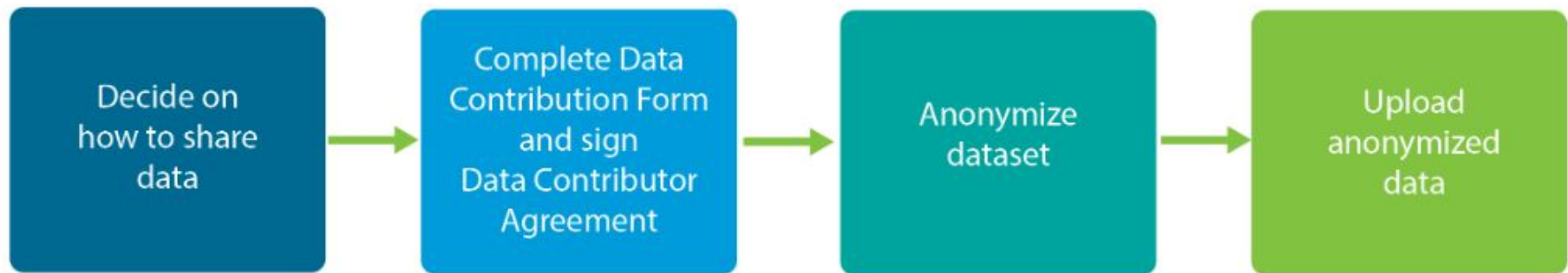


5,888,749 variables

### Most Popular Search Terms



## Steps to sharing your clinical research data on Vivli



- Managed access repository
  - Users request access to the data for a specific research proposal
- Data collected
  - anonymized individual participant-level data (IPD)
  - raw data collected during a clinical trial
- Vivli Data Use Agreement, user agrees to:
  - maintain data security
  - refrain from any attempts to reidentify research participants
  - submit a valid scientific question, including a statistical analysis plan
- Vivli agrees to store data for 10 years

# ResearchDrive?

- File storage service for research data - 5 TB per PI
- Share data with UW colleagues or external collaborators
- [Approved tool for exchanging HIPAA data](#)
- Cross platform - Linux, Mac OS, Windows
- High speed connectivity on campus or remotely accessible via VPN and the web.

uw / app / researchdrive / pis / mplayde / rdrive-mplayde-lab

## rdrive-mplayde-lab Group details

Path: uw:app:researchdrive:pis:mplayde:rdrive-mplayde-lab

Total Members: 7

Publish to Campus Active Directory?: Yes

Contacts Members Privileges Invitations Services

Members

+ Add member(s) Remove selected member(s)

mplayde (\research.drive.wisc.edu) x +

This PC > mplayde (\research.drive.wisc.edu) (R) >

Name	Date modified	Type	Size
500GB	2022-07-14 12:11	File folder	
logs	2021-04-14 15:28	File folder	
pc	2022-11-19 15:51	File folder	
project1	2022-03-09 14:42	File folder	
transfer test	2022-05-12 10:39	File folder	
transfertest	2022-05-17 12:31	File folder	
web	2020-12-10 15:26	File folder	
1G.dat	2021-02-10 14:41	DAT File	976,563 KB
10M.dat	2020-12-08 13:22	DAT File	9,766 KB
100M.dat	2021-02-10 14:41	DAT File	97,657 KB

22 items

# Gather data

## Research Cores

- Data from instruments
- [Biotech Center](#) and [UW-CCC](#)

## Data Protection

- Automated data protection
- Snapshots and offsite mirroring of data
- [Ransomware protection](#) including offline backups

### Core Facilities



[Bioinformatics](#)



[Next Gen Seq](#)



[Genome Editing](#)



[Animal Models](#)



[Gene Expression](#)



[Mass Spec](#)



**UWCCC RESEARCH**

Resource for Researchers

[INNOVATION](#) [RESOURCES](#) [CLINICAL TRIALS](#) [FUNDING OPPORTUNITIES](#) [RESEARCH](#)

[HOME](#) / [SAIRF](#) / [NEED TO ACCESS AND ANALYZE SAIRF DATA?](#)

## NEED TO ACCESS AND ANALYZE SAIRF DATA?

Posted on [December 14, 2021](#)

SAIRF data is automatically backed up to UW ResearchDrive and can be accessed and analyzed via:

# Analyze & visualize data

## Endpoint Computers

- [UW Endpoint Management and Security standards](#)
- Can easily share data with other departments



## Computing Centers

- Easy to move data in/out of computing centers
- [Platform X](#) and [SSCC](#)



Public



Internal



Sensitive



Restricted

# Share data with collaborators

## UW-Madison NetID

- Share data with anyone with a NetID

## External Collaborators

- Send invites to create a NetID with VPN and Duo-MFA

## Remote Access

- Collaborators use UW VPN and Duo-MFA

uw / app / researchdrive / pis / mplayde / rdrive-mplayde-external

### rdrive-mplayde-external Group details

Path: uw:app:researchdrive:pis:mplayde:rdrive-mplayde-external

Total Members: 2

[Contacts](#) [Members](#) [Privileges](#) [Invitations](#) [Services](#)

#### Invitations

Permission to invite external users (without NetIDs)

#### Send invitation(s)

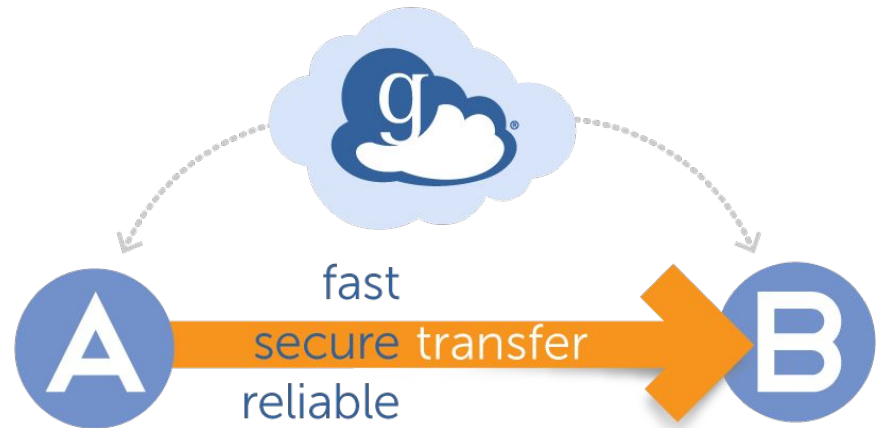
Enter a comma-separated list of e-mail addresses

**Invitations to send**

No invitations

# Globus?

- Web-based research collaboration platform that supports secure data sharing and transfers
- Connects disparate storage systems together through Globus endpoints



<https://it.wisc.edu/services/globus>



# Data transfer

## Research Cores

- [Cryo-EM](#) and [Biotech Center](#)

## Computing Centers

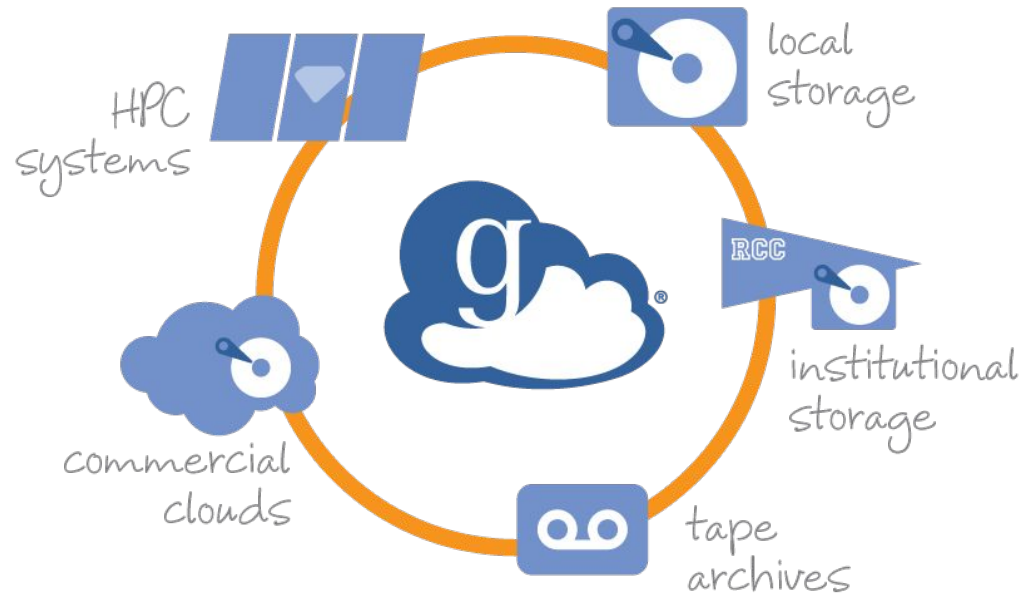
- [SSCC](#) and [Platform X](#)

## Campus Storage

- [ResearchDrive](#), [Shared Drive](#), and [S3](#)

## Public Clouds

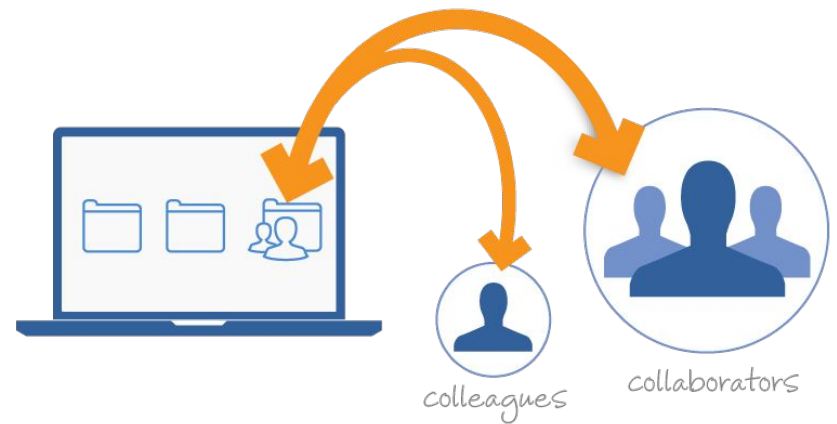
- [AWS](#), [Google Cloud](#), and [Microsoft Azure](#)



# Data sharing

Securely exchange data with external collaborators

- Allows sharing between institutions using existing identities
- [Approved tool for exchanging HIPAA data](#)
- Supports sharing files with web links



The screenshot shows the 'File Manager' interface. The collection is 'wisc-drive-mplayde' and the path is '/'. The 'Shareable links to:' section shows two options: 'access this Collection and Path in File Manager' with a 'view URL' button, and 'access the selected file directly' with a 'view URL' button and a text input field containing the URL <https://g-b8c688.dtn.globus.wisc.edu/1G.dat>. Below this is a table of files:

NAME	LAST MODIFIED	SIZE
<input type="checkbox"/> 100M.dat	2/10/2021, 02:41 ...	100 MB
<input type="checkbox"/> 10M.dat	12/8/2020, 01:22 ...	10 MB
<input checked="" type="checkbox"/> 1G.dat	2/10/2021, 02:41 ...	1 GB
<input type="checkbox"/> 2021-10-11_11-40-05.png	10/11/2021, 11:40 ...	360.33 KB
<input type="checkbox"/> 24096__acclivity__henry5.mp3	9/17/2021, 11:35 AM	6.72 MB
<input type="checkbox"/> 500GB	7/14/2022, 12:11 PM	—

The right sidebar shows various actions: Permissions, Transfer or Sync to..., New Folder, Rename, Delete Selected, Download, Open, Upload, Get Link, Show Hidden Items, and Manage Consent.

# Publish research artifacts

Publish open access or moderated data sets through the web

- Deposit large data sets in [Minds@UW](https://minds.uw.edu/) with Globus
- [Globus platform](https://www.globus.org/) integration for web applications and data portals



# **NIH DMSP Drop-ins & DMS Plan Office Hours**

## **Drop-in Information Sessions**

Health Sciences Learning Center 3330

- Every Wednesday, January 18 - March 15, 2023 (Noon-1PM)

## **DMS Plan Open Office Hours**

Health Sciences Learning Center 2320

- Every Tuesday, January 10 - March 28 (11:30AM-1:30PM)
- Every 3rd Tuesday, January – March (4:30PM-6:30PM)

**QUESTIONS?**



TRISHA ADAMUS

adamus@wisc.edu

CAMERON COOK

cameron.cook@wisc.edu

MICHAEL LAYDE

michael.layde@wisc.edu

researchdata.wisc.edu

THANK YOU.



# Researcher Toolkit

Your Guide to Researcher Resources at UW-Madison

Lifecycle ▾

Resource List

Acronyms

Feedback

About

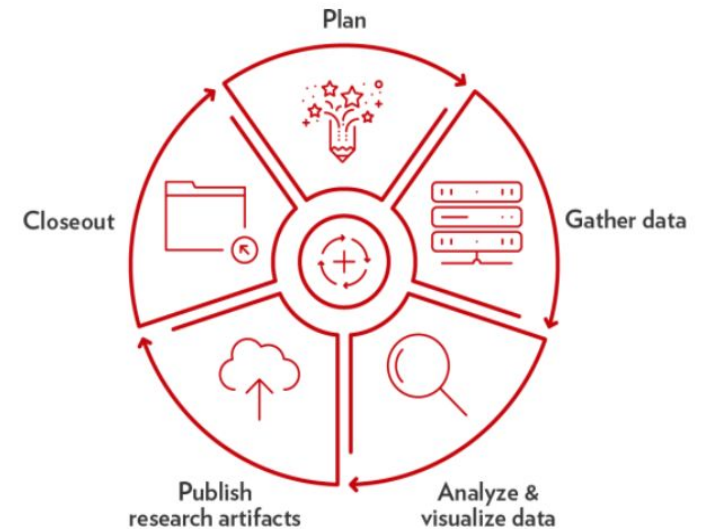
## What is the research lifecycle?

### The research lifecycle

The research lifecycle includes planning, collecting & storing data, analyzing & visualizing data, publishing research artifacts, and closing out the findings.

*Plus* outreach, compliance, reporting, and security happening throughout the process.

About Research Lifecycle Components



# UW-Madison resources.

[Research Data Services](#) - data management help, writing DMPs, data sharing/archiving

[Data Science Hub](#) - Coding Meetup, training, office hours

- Data & Software Carpentry - [sign up for the Data Science Hub newsletter](#)

[CALS Statistical Consulting](#) - study design, data analyses, software use

[Social Science Computing Cooperative](#) - software training, SILO

[UW-Madison Libraries](#) - finding data, copyright, publishing questions, etc.

[IRB](#), [IRB Guidance](#)

[MINDS@UW](#) - data sharing via UW institutional repository

[Dryad](#) - data repository available via UW-Madison membership

[Researcher Toolkit](#)

[Data Storage Finder](#)