The New NIH Policy for
Data Management and Sharing
What you need to know

Notice Number: NOT-OD-21-013
Effective Date: January 25, 2023

NIH DMS Plan | Requirements

The DMSP policy has **two main requirements:**

1. **Submission of a two-page data management and sharing plan:** Research proposals without a Plan will not be considered for funding.

2. **Compliance with the approved plan.** Failure to provide updates in grant reporting may result in enforcement actions, including the addition of special terms and conditions or award termination. Failure to deposit data after the end of the funding period may negatively influence future opportunities.

Effective Date: Applies to **competing grant applications** that are submitted to NIH for **January 25, 2023**, and subsequent receipt dates and **all proposals for contracts** that are submitted to NIH on or after January 25, 2023. Applies only to projects that result in the generation of "scientific data."

NIH DMSP Policy | Applicability and Timing

Applicability: Applies to all research, funded or conducted in whole or in part by NIH, that results in the generation of **scientific data**, regardless of funding level.
Does not include projects that do not generate scientific data, e.g., Training (T), Fellowships (Fs), Construction (C06), Conference Grants (R13), Resource (Gs), and Research-Related Infrastructure Programs (e.g., S06)

Timelines:

Proposal
Submit 2-page DMS Plan, describe in budget justification

JIT
Feedback from program official

Life of Award
Follow/modify throughout life of award

End of Award/Publication
Share data no later than publication or end of award, including first no-cost-extension (for unpublished data)

How long to share/maintain data
At a minimum, 3 years following closeout of a grant or contract agreement. Contracts may specify different time periods.

NIH DMSP Policy | Reason for Policy

Final NIH Policy for Data Management (NOT-OD-21-013):

"Data sharing enables researchers to rigorously test the validity of research findings, strengthen analyses through combined datasets, reuse hard-to-generate data, and explore new frontiers of discovery.

In addition, NIH emphasizes the importance of good data management practices, which provide the foundation for effective data sharing and improve the reproducibility and reliability of research findings.

NIH encourages data management and data sharing practices consistent with the [Findability, Accessibility, Interoperability, and Reuse (FAIR)] data principles."

NIH DMS Plan | Assessment of DMS

Assessment by NIH Program Staff

- Peer review will not see or review DMS Plans but will consider any related budget items
- DMS Plans are **NOT** part of scored peer review criteria unless specifically noted in the Funding Opportunity Announcement
- NIH Program Staff will review the DMS Plan for acceptability and may request modifications prior to award as appropriate
- Plans must be approved by the funding institute prior to award

Just-in-Time Review

If DMS Plan provided in the application cannot be approved based on the information provided, applicants will be notified that additional information is needed through the Just-in-Time (JIT) process.

NIH DMSP Policy | Scientific Data

Definition of "Scientific Data" that must be shared

- Recorded factual material commonly accepted in the scientific community as necessary to validate and replicate research findings, regardless of whether the data are used to support scholarly publications

- Does not include laboratory notebooks, preliminary analyses, completed case report forms, drafts of scientific papers, plans for future research, peer reviews, communications with colleagues, or physical objects, such as laboratory specimens

DMS Plan | NIH Required Elements

1. Data type

- Identifying estimated type and amount of data to be generated (i.e., modality, level of aggregation, and degree of data processing)
- Which data to be preserved and shared
- Accompanying metadata, other relevant data, and associated documentation to be made available

2. Related tools, software, code

- Tools and software needed to access and manipulate data

3. Standards

- Standards, if any, that will be applied to the scientific data and associated metadata (i.e., data formats, data dictionaries, data identifiers, definitions, unique identifiers, and other data documentation)

DMS Plan | NIH Recommended Elements (Cont'd)

4. Data preservation, access, and associated timelines

- Proposed repository to be used consistent with Supplemental Information on Repository Selection
- How data will be findable and accessible (e.g., persistent unique identifier)
- When data will be made available and for how long

5. Access, Distribution, and Reuse Considerations

- Description of factors potentially affecting data access, distribution, or reuse related to informed consent or privacy and confidentiality protections
- Whether access to human data will be controlled

6. Oversight of data management

- Describe how compliance with the DMS Plan will be monitored and managed

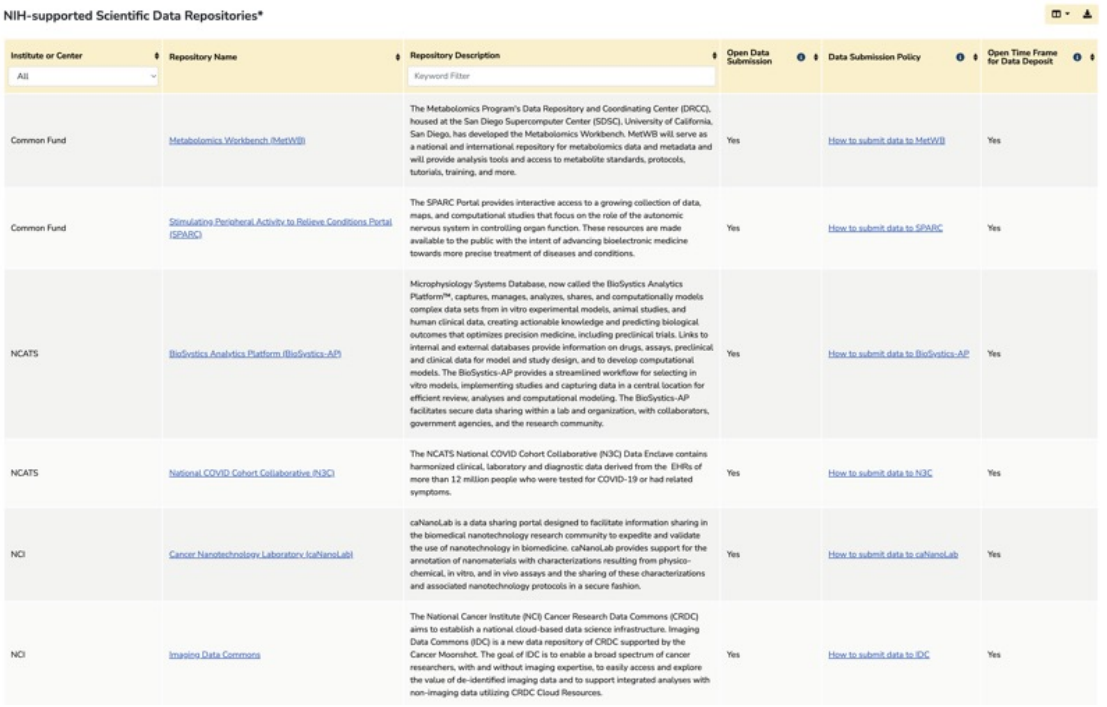
Note(s):

Data management responsibilities extend beyond just the PI or researcher(s) who create or collect the data. Various parties involved in the research process (including collaborators) must play a role in ensuring good quality data stewardship throughout the life of the project. It will be essential that roles and responsibilities of data management be clearly defined and assigned, rather than assumed.

Data Preservation & Access

Selecting a Repository for Data Resulting from NIH-Supported Research

- The DMS Policy does not currently mandate any particular repository for data, but some individual NIH institutes have specified required repositories (e.g., NIMH).
- NIH encourages researchers to select the repository that is most appropriate for their data type and discipline. A Generalist repository can also be used if researchers are unable to find a repository that suits their data.



NIH-supported Scientific Data Repositories*

Institute or Center	Repository Name	Repository Description	Open Data Submission	Data Submission Policy	Open Time Frame for Data Deposit
All		Keyword Filter			
Common Fund	Metabolomics Workbench (MetWB)	The Metabolomics Program's Data Repository and Coordinating Center (DRCC), housed at the San Diego Supercomputer Center (SDSC), University of California, San Diego, has developed the Metabolomics Workbench. MetWB will serve as a national and international repository for metabolomics data and metadata and will provide analysis tools and access to metabolite standards, protocols, tutorials, training, and more.	Yes	How to submit data to MetWB	Yes
Common Fund	Stimulating Peripheral Activity to Relieve Conditions Portal (SPARC)	The SPARC Portal provides interactive access to a growing collection of data, maps, and computational studies that focus on the role of the autonomic nervous system in controlling organ function. These resources are made available to the public with the intent of advancing bioelectronic medicine towards more precise treatment of diseases and conditions.	Yes	How to submit data to SPARC	Yes
NCATS	BioSystics Analytics Platform (BioSystics-AP)	Microphysiology Systems Database, now called the BioSystics Analytics Platform™, captures, manages, analyzes, shares, and computationally models complex data sets from in vitro experimental models, animal studies, and human clinical data, creating actionable knowledge and predicting biological outcomes that optimizes precision medicine, including preclinical trials. Links to internal and external databases provide information on drugs, assays, preclinical and clinical data for model and study design, and to develop computational models. The BioSystics-AP provides a streamlined workflow for selecting in vitro models, implementing studies and capturing data in a central location for efficient review, analyses and computational modeling. The BioSystics-AP facilitates secure data sharing within a lab and organization, with collaborators, government agencies, and the research community.	Yes	How to submit data to BioSystics-AP	Yes
NCATS	National COVID Cohort Collaborative (N3C)	The NCATS National COVID Cohort Collaborative (N3C) Data Enclave contains harmonized clinical, laboratory and diagnostic data derived from the EHRs of more than 12 million people who were tested for COVID-19 or had related symptoms.	Yes	How to submit data to N3C	Yes
NCI	Cancer Nanotechnology Laboratory (caNanoLab)	caNanoLab is a data sharing portal designed to facilitate information sharing in the biomedical nanotechnology research community to expedite and validate the use of nanotechnology in biomedicine. caNanoLab provides support for the annotation of nanomaterials with characterizations resulting from physico-chemical, in vitro, and in vivo assays and the sharing of these characterizations and associated nanotechnology protocols in a secure fashion.	Yes	How to submit data to caNanoLab	Yes
NCI	Imaging Data Commons	The National Cancer Institute (NCI) Cancer Research Data Commons (CRDC) aims to establish a national cloud-based data science infrastructure. Imaging Data Commons (IDC) is a new data repository of CRDC supported by the Cancer Moonshot. The goal of IDC is to enable a broad spectrum of cancer researchers, with and without imaging expertise, to easily access and explore the value of de-identified imaging data and to support integrated analyses with non-imaging data utilizing CRDC Cloud Resources.	Yes	How to submit data to IDC	Yes

https://sharing.nih.gov/data-management-and-sharing-policy/sharing-scientific-data/repositories-for-sharing-scientific-data

Budgeting for Data Management & Sharing

Investigators may request funds towards data management and sharing in the budget justification section of their application.

Allowable Costs

- Curating data
- Developing supporting documentation
- Formatting data according to accepted community standards, or for transmission to and storage at a selected repository for long-term preservation and access
- De-identifying data
- Preparing metadata to foster discoverability, interpretation, and reuse
- Local data management considerations, such as unique and specialized information infrastructure necessary to provide local management and preservation (for example, before deposit into an established repository).
- Preserving and sharing data through established repositories, such as data deposit fees (If the Data Management & Sharing (DMS) plan proposes deposition to multiple repositories, costs associated with each proposed repository may be included).

Note(s):

- All allowable costs submitted in budget requests must be incurred during the performance period, even for scientific data and metadata preserved and shared beyond the award period.

Budgeting for Data Management & Sharing (Cont'd)

Unallowable Costs

Budget requests must **NOT include**:

- Infrastructure costs that are included in institutional overhead (for instance, Facilities and Administrative costs)
- Costs associated with the routine conduct of research, including costs associated with collecting or gaining access to research data.
- Costs that are double charged or inconsistently charged as both direct and indirect costs

More guidance to come on budgeting –
November 18 Town Hall will focus on this

Budget Justification – Hot off the press!

The Data Management and Sharing justification must be clearly labeled as "Data Management and Sharing Justification" in the within the budget justification attachment.

Provide a brief summary of **type and amount of scientific data to be preserved and shared** and the name of the established repository(ies) where they will be preserved and shared.

Indicate general cost categories such as curating data and developing supporting documentation, local data management activities, preserving and sharing data through established repositories, etc., **including an amount for each category and a brief explanation**.

Specify in the justification if no costs will be incurred for Data Management and Sharing, if applicable. The recommended length of the justification should be no more than half a page.

Post-award Monitoring and Oversight

The NIH-approved plan becomes a **Term and Condition** of the Notice of Award.

- Principal investigators must 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.
- During the funding period, compliance with the Plan will be determined by the NIH ICO. Compliance with the Plan, including any Plan updates, may be reviewed during regular reporting intervals.
- Note that while data management responsibilities will require a team effort, **PIs are ultimately responsible for ensuring that the DMS plan is properly executed in a timely and accurate manner**.

Flexibility

"The final DMS Policy does not create a uniform requirement to share all scientific data. ….[A]ppropriate data sharing is likely to be varied and contextual….The final DMS Policy retains the Draft Policy's factors (i.e., ethical, legal, or technical) that may necessitate variations in the extent of scientific data preservation and sharing, and researchers should convey such factors in their Plans. The final DMS Policy has also been modified to clarify these factors are not limited to data derived from human research participants. **We believe this will provide the necessary flexibility for researchers to accommodate the substantial variety in research fields, projects, and data types that this expectation will encompass.**"

https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-013.html

NIH FAQs | Secondary data

Q. If researchers are reusing existing, shared data to generate new datasets, are they expected to reshare the primary data they incorporated into their new analysis? Are the derived data generated considered scientific data and expected to be shared?

The DMS Policy applies to research that results in the generation of scientific data. Scientific data can result from secondary research, but researchers are not expected to share the existing, shared primary data used to conduct the secondary research. Researchers are, however, expected to maximize appropriate sharing of any new, derived data generated as a result of their research. Note that use of data obtained from repositories or other sources and derived data may be subject to limitations on sharing as a condition of access, which is a justifiable reason for limiting sharing under the DMS Policy.

https://osp.od.nih.gov/scientific-sharing/nih-policy-for-data-management-and-sharing-faq/

NIH FAQs | Limiting data sharing

Q. What are justifiable reasons for limiting sharing of data

NIH expects that researchers will take steps to maximize data sharing but recognizes that certain factors (i.e., ethical, legal, or technical) may necessitate limiting sharing. Anticipated limitations should be described in the DMSP. Potential factors include but are not limited to:

- Informed consent will not permit or will limit the scope or extent of sharing and future research use
- Existing consent (e.g., for previously collected biospecimens) 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 (e.g., with third party funders, with partners, with repositories, with Health Insurance Portability and Accountability Act (HIPAA) covered entities that provide Protected Health Information under a data use agreement, through licensing limitations attached to materials needed to conduct the research)
- Datasets cannot practically be digitized with reasonable efforts

https://osp.od.nih.gov/scientific-sharing/nih-policy-for-data-management-and-sharing-faq/

NIH FAQs | Timing of sharing data

Q. Does the timeline for sharing scientific data change in the event of a no-cost extension?

Scientific data should be made accessible as soon as possible, and no later than the time of an associated publication or the end of the performance period of the extramural award that generated the data. If a no cost extension is granted for an extramural award, scientific data should be made accessible no later than the time of an associated publication, or the **end of the no cost extension**, whichever comes first.

Scientific data underlying findings not disseminated through peer-reviewed journal articles should be shared by the end of the performance period unless the grant enters into a no-cost extension…. These scientific data may underlie unpublished key findings, developments, and conclusions; or findings documented within preprints, conference proceedings, or book chapters. For example, **scientific data underlying null and negative findings are important to share even though these key findings are not always published**.

https://osp.od.nih.gov/scientific-sharing/nih-policy-for-data-management-and-sharing-faq/

NIH FAQs | Data derived from human participants

Q. Which steps does the DMS policy take to ensure institutions and researchers protect research participants?

Award recipients must comply with any applicable laws, regulations, statutes, guidance, or institutional policies related to research with human participants and that protect participants' privacy. The DMS Policy encourages respect for participants by encouraging researchers and award recipients to:

- Address data management and sharing plans during the informed consent process to ensure prospective participants understand how their data will be managed and shared
- Outline steps they will take for protecting the privacy, rights, and confidentiality of prospective participants (i.e., through de-identification, Certificates of Confidentiality, and other protective measures)
- Assess limitations on subsequent use of data and communicate these limitations to the individuals or entities (e.g., repositories) preserving and sharing the data
- Consider whether access to shared scientific data derived from humans should be controlled, even if de-identified and lacking explicit limitations on subsequent use. Sharing via controlled access may be specified by certain funding opportunity announcements (FOAs) or the funding NIH ICO(s).

https://osp.od.nih.gov/scientific-sharing/nih-policy-for-data-management-and-sharing-faq/

NIH FAQs | Compliance and enforcement

Q. How will noncompliance with the NIH DMS Policy be handled?

NIH will monitor compliance with Plans over the course of the funding period during regular reporting intervals (e.g., at the time of annual Research Performance Progress Reports (RPPRs)).

Noncompliance with Plans may result in:

- NIH ICO adding special Terms and Conditions of Award
- Termination of the award

If award recipients are not compliant with Plans at the end of the award, noncompliance may be factored into future funding decisions.

Note(s):

- To avoid potential issues when reporting progress, ensure that the submitted DMS plan contains sufficient details for the program officer to be able to evaluate compliance.

https://osp.od.nih.gov/scientific-sharing/nih-policy-for-data-management-and-sharing-faq/

Getting Help

NIH Resources

- NIH Data Management & Sharing Policy
- Supplemental Information to the NIH Policy for Data Management and Sharing: Elements of an NIH Data Management and Sharing Plan
- Supplemental Information to the NIH Policy for Data Management and Sharing: Allowable Costs for Data Management and Sharing
- 2023 DMSP FAQs
- Preview/Draft DMS Plan format page (PDF)
- Selecting a Repository for Data Resulting from NIH-Supported Research
- Repositories for Sharing Scientific Data
- Open Domain-Specific Data Sharing Repositories (September 2020)
- Generalist Repositories (October 2021)

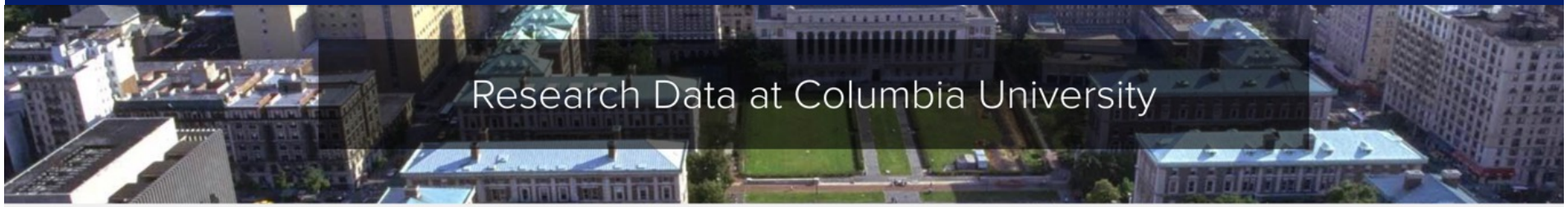
Visit **sharing.nih.gov** for more information!

Getting Help

Other Resources

- DMPTool | Create Data Management Plans that meet requirements and promote your research
- DMPTool NIH DMSP Template (PDF)
- FASEB DataWorks! Data Management Plan (DMP) Challenge Evaluation Rubric
- Visualize key recommendations in the PLOS Computational Biology article "Ten simple rules for maximizing the recommendations of the NIH data management and sharing plan"
- "Forecasting Costs for Preserving, Archiving, and Promoting Access to Biomedical Data"
- re3data (Registry of Research Data Repositories): Offers researchers an overview of existing repositories for research data
- Briney KA, Coates H, Goben A (2020) Foundational Practices of Research Data Management. Research Ideas and Outcomes 6: e56508
- Wilkinson, M., Dumontier, M. et al, The FAIR Guiding Principles for Scientific Data Management and Stewardship (March 2016)

Columbia Resources: data.research.columbia.edu



Getting Started with Research Data Management

Writing a Data Management and Sharing Plan

Sharing Data and Finding the Right Repository

Data Security

Ownership of Data

Data Retention

Funding Agency Data Policies

NIH Data Management and Sharing Plan (2023)

FEATURED RESOURCES

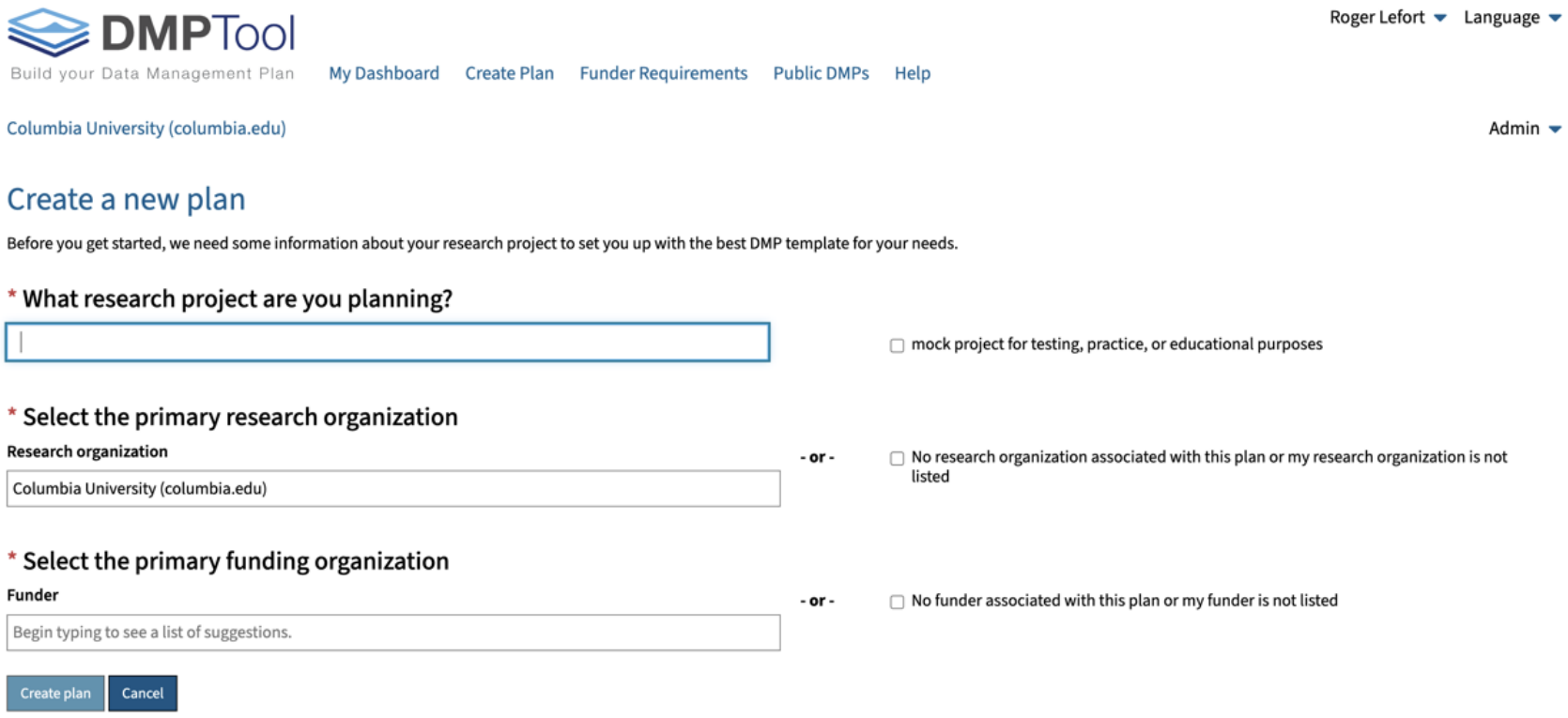
Columbia Resources | Data storage options

UNIT	SERVICE	FREE	UNLIMITED	MUTABLE	BACKUPS	VERSION CONTROL	PII/RHI SECURE	PHI CERTIFIED SECURE	ACCESS CONTROL	SHARABLE	ARCHIVAL PRESERVATION	LONG-TERM STORAGE	PUBLIC ACCESS
CUIT	LionMail Drive	Yes	No	Yes	Yes	No	No	No	Yes	Yes	No	No	No
	Amazon Web Services (AWS)	Free Tier	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
	Google Cloud Platform (GCP)	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	No	Yes	Yes
	Secure Data Enclave (SDE)	No	No	Yes	No	No	Yes	Yes	Yes	No	No	No	No
	Microsoft Azure	Free Tier	Yes	Yes	Yes	Yes	No	No	Yes	Yes	No	No	No
	Globus (data transfer)	Yes	N/A	N/A	No	N/A	No	No	Yes	Yes	No	N/A	No
	Columbia Data Platform	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
	Box	Free Tier	Unlimited Tier	Yes	Yes	Yes	Yes	No	Yes	Yes	No	Yes	No
CUIMCIT	CUIMC IT Storage	No	Unlimited Tier	Yes	Yes	Yes*	Yes	Yes	Yes	Yes	No	Yes*	Yes*
	CUIMC IT FTP Server	No	No	Yes	No	No	Yes*	Yes*	Yes	Yes	No	No	No
	SharePoint Online – Office 365	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
	Virtual Servers	No	No	Yes	Yes	N/A	Yes*	Yes*	Yes	Yes	No	No*	No
Libraries	Academic Commons	Yes	No	No	Yes	Yes*	No	No	No	Yes	Yes	Yes	Yes
	Dryad	Yes	Yes	No	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes
CUIT/Libraries	LabArchives	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No*	No
Open Source Other	Open Science Framework	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes	No	Yes	Yes
	PC/Mac Server (w/ options)*	No	Unlimited Tier	Yes	Yes	Yes*	No	No	Yes*	Yes	No	Yes	No

Research Data Storage Options

DMS Plan | DMPTool

DMPTool is a free tool to help researchers create data management plans that comply with funder requirements. It walks users through the components required by the funder they're applying to, and allows them to create a completed DMP, which can be exported as a Word document or PDF.



Where to start | First steps

1. **Determine your personal timeline**: If you currently have an active NIH award up for renewal or are planning to submit a proposal, then developing a DMS plan should be a high priority, especially if you will be working with collaborators, as it may take time to set up appropriate data procedures and agreements.

2. **Carefully read and monitor the CU webpage** to stay up to date with any changed and updates to the policy, including supplements.

3. **Familiarize yourself with the FAIR principles** (Wilkinson et. al, 2016). The FAIR (Findable, Accessible, Interoperable, Reusable) data principles are the guiding principles the NIH has used in creating the new policy.

4. **Assess** your own project needs and current data management practices against to new policy, and document and update existing practices. Start developing new ones, if needed, to address the increased emphasis on data sharing and oversight.

5. **Review** campus data services and resources (computing, storage, consulting services, etc.) and determine if they meet your needs. Also consider any additional costs for labor related to data curation and documentation.

Prepare your data for sharing

1. **Gather your data and related information**
   - raw data
   - processed and transformed data
   - documentation (README files, data dictionaries)
   - data tables and statistical summaries
   - figures and visualizations
   - methods and protocols
   - code: data processing and analysis scripts

2. **Verify that your files can be shared publicly**
   - remove copyrighted or user-restricted content
   - remove sensitive or personally identifiable information
   - anonymize human subject data
   - mask collection sites or locations of endangered or vulnerable species

3. **Organize data and documents**
   - use descriptive, meaningful and consistent file names
   - organize directories/folders logically (e.g., Scripts, Raw Data, Processed Data, Results)

4. **Choose open file formats over proprietary**
   - for example, save in text (.txt, csv) or rather than Microsoft Word formats (.docx, .xlsx)
   - see the Library of Congress recommendations for various file formats

5. **Upload data files to your chosen data repository**
   - get permanent identifier (DOI) for your dataset

Adapted from https://datadryad.org/stash/best_practices

Town Halls and other Outreach

Fall 2022 Town Halls scheduled:

- Session 1: Thursday, October 27, 2022; 12:00pm – 1:00pm EST
- Session 2: Friday, November 18, 2022; 2:00pm – 3:00pm EST
- Session 3: Thursday, December 1, 2022; 12:30pm – 1:30pm EST

Also planning **faculty-focused smaller sessions**:

- Faculty Sharing Effective Data Management Practices
- DMS Plan Drafting Workshops

What **other resources** or **outreach** would be helpful?

Questions?

If you have questions about NIH's new data management and sharing policy, you may email:

data-management@columbia.edu

We will route as appropriate to SPA, Research Compliance, Libraries, or others with appropriate expertise.

Questions?

For questions regarding…	Please contact…
• Proposal submission • Budgeting and allowable costs • NIH requirement for DMSP • Data sharing and use agreement	SPA Project Officer
• Creating a DMS plan • Selecting a data repository • Managing data during research projects • Publishing and preserving data research	Research Data Services ORCT
• Informed consent and data privacy related to DMSP • Documenting and obtaining IRB approval for DMSP	Human Research Protection Office and IRBs
• Resources available to meeting research data needs • Consultation and workshops • Computation and data storage	Research Data Services ORCT CUIT CUIMCIT
• Data security and secure information technology systems • Identification of architectural requirements • Best practices in data security	CUIT CUIMCIT