NIH Data Management and Sharing Policy

Meredith C. Hermosura, Ph.D.

Pre-Award Office

02/02/2023

Talk Outline

- DMS Policy - Details

- Budget, Justification, and Submission

- The DMS Plan – Elements 1 to 6

Useful Overview Page:

https://sharing.nih.gov/data-management-and-sharing-policy/planning-and-budgeting-for-data-management-and-sharing/writing-a-data-management-and-sharing-plan#after

Applications for Receipt Dates
BEFORE Jan 25 2023

Applications for Receipt Dates
ON/AFTER Jan 25 2023

Applications for Receipt Dates ON/AFTER Jan 25 2023

ON THIS PAGE:

- Writing a Data Management and Sharing Plan
- Submitting Data Management and Sharing Plans
- Data Management and Sharing Plan Format
- Elements to Include in a Data Management and Sharing Plan
- Sample Plans
- Assessment of Data Management and Sharing Plans
- Revising Data Management and Sharing Plans
- Additional Considerations

From the NIH Grants Conference (Feb 1-2, 2023) – The NIH Policy for Data Management and Sharing (DMS) in Effect: Planning for Success

Goals of Data Sharing

- **Advance rigorous and reproducible research**
  - Enable validation of research results
  - Make high-value datasets accessible
  - Accelerate future research directions
  - Increase opportunities for citation and collaboration

- **Promote public trust in research**
  - Foster transparency and accountability
  - Demonstrate stewardship over taxpayer funds
  - Maximize research participants' contributions
  - Support appropriate protections of research participants' data

NIH National Institutes of Health
Office of Extramural Research

From the NIH Grants Conference (Feb 1-2, 2023) – The NIH Policy for Data Management and Sharing (DMS) in Effect: Planning for Success

Policy Details

- **Requirements (effective January 25, 2023):**
  1. **Submission** of Data Management & Sharing Plan
  2. **Compliance** with ICO-approved Plan (may affect future funding)

- **Scope:** All NIH-supported research generating _scientific data_

  *Whatever is used to make tables or figures that can be **validated** and **replicated***

  - **What's in:** "Recorded factual material… of <u>sufficient quality to validate and replicate research findings</u>, regardless of whether the data are used to support scholarly publications"
  - **What's out:** Lab notebooks, preliminary analyses, peer reviews, physical objects

- **Timelines:**
  - **When to share data:** No later than <u>publication</u> or <u>end of award</u> (for unpublished data), whichever comes first
  - **How long to share data:** Consider other relevant requirements and expectations (e.g., journal policies, repository policies)

NIH National Institutes of Health
Office of Extramural Research

'Which brings us to my next point.'

More on Submission and Compliance

- **Submission**: If a DMS Plan is <u>required</u> for that activity code and not included in the submission, 'it will likely be returned without review' (eRA validation)

- **Compliance** will be monitored by Program Officers during the *annual progress report* reviews (RPPR); noncompliance may result in (1) termination of award, (2) addition of special Terms and Conditions of Award and (3) factored into future funding decisions

Is your proposed research subject to new DMS policy?

✓ **All** NIH research that results in the generation of **scientific data**

Subject to DMS Policy (Examples)	**Not Subject to DMS Policy**
▪ Research Projects (Rs)	▪ Training (Ts)
▪ Some Career Development Awards (Ks)	▪ Fellowships (Fs)
▪ Small Business SBIR/STTR	▪ Some conference grants (R13)
▪ Research Centers (Ps)	▪ Construction
▪ Cooperative Agreements (Us)	▪ Resource Centers (Gs)
	▪ Research-Related Infrastructure Programs (e.g., S06, S10)

NIH Activity Codes subject to DMS Policy

https://sharing.nih.gov/sites/default/files/flmngr/List-of-Activity-Codes-Applicable-to-DMS-Policy.pdf

*This is a partial screenshot of page 1 of the **List of Activity Codes** pdf*

*There are 19 pages in this document, listing Program Codes from **D** to **U***

NIH Data Management and Sharing Policy: Applicable Activity Codes

Revised on December 16, 2022

A comprehensive listing of all NIH activity codes that generally require applicants to submit a Data Management and Sharing Plan.

Program Code	Activity Code	Category	Title	Description	DMS Plan Requirement
D	DP1	Institutional Training and Director Program Projects	NIH Director's Pioneer Award (NDPA)	To support individuals who have the potential to make extraordinary contributions to medical research. The NIH Director's Pioneer Award is not renewable.	Y
D	DP2	Institutional Training and Director Program Projects	NIH Director's New Innovator Awards	To support highly innovative research projects by new investigators in all areas of biomedical and behavioral research.	Y
D	DP3	Institutional Training and Director Program Projects	Type 1 Diabetes Targeted Research Award	To support research tackling major challenges in type 1 diabetes and promoting new approaches to these challenges by scientific teams.	Y
D	DP4	Institutional Training and Director Program Projects	NIH Director's Pathfinder Award - Multi-Yr Funding	To support multi-year funded research with unique, high impact ideas for addressing biomedical research including assuring a balanced and effective workforce. This research grant program will encourage exceptionally creative scientists to develop potentially transforming approaches for supported research. The proposed research must reflect ideas that are substantially different from those already being pursued or they must apply existing research designs in new and innovative ways. This is a multi-year, funded companion activity code to the existing Pioneer Award (DP1); thus ICs need OER prior approval to use the DP4.	Y
D	DP5	Institutional Training and Director Program Projects	Early Independence Award	To support the independent research project of a recent doctoral degree recipient. This research grant program will encourage exceptionally creative scientists to bypass the typical post-doc research training period in order to move rapidly to research independence. It will encourage institutions to develop independent career tracks for recent graduates in order to demonstrate the benefits of early transition to independence both in terms of career productivity for the candidate and research capability for the institution.	Y
E	E11	Health Professions Programs	Grants for Public Health Special Projects	To provide grants to public or nonprofit organizations for planning, development, demonstration, research, training, public information projects for preventive medicine, health promotion and disease prevention to improve the health of targeted populations.	Y
K	K00	Research Career Programs	Post-doctoral Transition Award	To support the second phase of a Pre-Doctoral to Post-Doctoral Transition award program that provides 3-4 years of career support. Note: The K00 Post-doctoral Transition Award is anticipated to only be used in conjunction with the F99 Pre-Doctoral Award.	Y
K	K01	Research Career Programs	Research Scientist Development Award - Research & Training	For support of a scientist, committed to research, in need of both advanced research training and additional experience.	Y

Policy for No Cost Extensions & Competitive Renewals

- <u>No cost extension</u>: share scientific data *no later* than the time of an associated **publication**, or the **end** *of the no cost extension*, whichever comes first

- <u>Competitive renewals</u>: a DMS Plan should be submitted at the **time** of the **renewal funding application** for research generating scientific data; if the previous award included a Plan, an updated Plan consistent with the proposed competitive renewal application should be submitted

Policy for Genomic Data

- On or after January 25, 2023, NIH will **no longer** collect separate **Genomic Data Sharing** (GDS) Plans

- **Submit a single Plan**, addressing genomic data sharing considerations (e.g., where and when genomic data will be shared) within DMS Plan Elements

  For example: describe genomic data such as sequencing reads and variant call files in Element 1 (Data Types) of the DMS Plan

When should scientific data should be shared?

As soon as possible, whichever is earlier of these **2 timepoints**

- No later than the than the date on which an associated publication is first made available in print or electronic format, or

- End of the performance period - at this time, all scientific data including null and negative findings, other unpublished findings documented in conference proceedings, book chapters, preprints*, etc.

*note that some preprint servers and data repositories may require that data be shared upon preprint posting

DMS Plan Submission

New **FORM-H** Package (effective 01/25/2023)

- new form field **#11** called 'Other Plans' on the **PHS 398 Research Plan** Form

DMS Plan Submission

New **FORM-H** Package (effective 01/25/2023)

- form field **#17** 'Other Plans' on the **Career Development Aw**ard forms

How will DMS Plans be reviewed?

✓ by **Program staff** at the proposed NIH Institute, Center, Office (ICO)

- Program staff will assess whether the DMS Plan has adequately addressed **6 recommended elements**

- Peer reviewers will **not** be asked to comment on the DMS Plan

- The Plan will **not** factor into the Overall Impact score (*unless* data sharing is *integral* to the project design and specified in the FOA; in this case, it may be factored into the score)

- Applications selected for funding will **only** be funded if the DMS Plan is complete and acceptable

DMS Plan revisions at the Pre-Award Stage

➢ proposals with <u>fundable scores but without an approved</u> DMS plan will be asked for *more information* or submit a *revised Plan*

- applicants will be notified through the **Just-in-Time** (JIT) process that the DMS Plan cannot be approved as submitted and more information is needed

- applicants should communicate with their NIH ICO Program Officer and/or Grants Management Specialist to resolve issues (a *revised Plan* may be required)

DMS Plan revisions at the Post-Award Stage

➢ the Plan should be *updated* or *revised* when changes occur during the support period that affect how data is managed or shared

- The NIH ICO-approved DMS Plan becomes **part** of the **Terms and Conditions of Award**

- Potential **changes** should be discussed with the NIH ICO Program Officer

- The updated Plan should be **approved** by the funding NIH ICO

DMS Plan – BUDGET

A budget is **not** required, unless DMS-related costs are requested

Data Management and Sharing Costs

ALLOWABLE COSTS:
- Curating data/developing supporting documentation
- Preserving/sharing data through repositories
- Local data management considerations
- **IMPORTANT:** Must be incurred during the performance period

UNALLOWABLE COSTS:
- Infrastructure costs typically included in indirect costs
- Costs associated with the routine conduct of research (e.g., costs of gaining access to research data)

From the NIH Grants Conference (Feb 1-2, 2023) – The NIH Policy for Data Management and Sharing (DMS) in Effect: Planning for Success

DMS Plan – BUDGET

Requesting Costs - R&R Detailed Budget Form

Label costs as '**Data Management and Sharing Costs**' line item, under **F. Other Direct Costs** (anywhere between line items 8-17, as applicable)

IMPORTANT: Type in **$0** if **not** requesting costs (or ASSIST will return 'error' message

Justifying Costs

- Recommended to be ½ a page
- Include brief summary of type and amount of scientific data to be preserved, shared, name of repository(ies) to be used, and general cost categories.

F. Other Direct Costs — Funds Requested ($)

1. Materials and Supplies
2. Publication Costs
3. Consultant Services
4. ADP/Computer Services
5. Subawards/Consortium/Contractual Costs
6. Equipment or Facility Rental/User Fees
7. Alterations and Renovations
8.
9.
10.
11.
12.
13.
14.
15.
16.
17.

Up to 10 additional Other Direct Costs line items can be added. Examples of possible uses: Tuition Remission, Technical Assistance, and Patient Care Costs.

FORMS-H: If a Data Management and Sharing (DMS) plan is included, you must include a "Data Management and Sharing Costs" line item covering DMS costs, including personnel costs (e.g., personnel who will be curating data for the project). If no cost incurred, enter 0. Type the string as requested (without quotation marks) and do not combine the line item with any "Other" costs.

Total Other Direct Costs

G. Direct Costs — Funds Requested ($)
Total Direct Costs (A thru F)

H. Indirect Costs

Indirect Cost Type	Indirect Cost Rate (%)	Indirect Cost Base ($)	Funds Requested ($)

Total Indirect Costs

Cognizant Federal Agency (Agency Name, POC Name, and POC Phone Number)

I. Total Direct and Indirect Costs — Funds Requested ($)
Total Direct and Indirect Institutional Costs (G + H)

J. Fee — Funds Requested ($)

K. Total Costs and Fee — Funds Requested ($)
Total Costs and Fee (I + J)

L. Budget Justification
(Only attach one file.)   Add Attachment   Delete Attachment   View Attachment

Budget Justification is required and must cover all budget periods.

FORMS-H: If a Data Management and Sharing (DMS) plan is included, you must include a section titled "Data Management and Sharing Justification" that provides a brief brief summary of DMS activities and justification for their costs.

DMS Plan – BUDGET

Requesting and Justifying Costs – PHS 398 Modular Budget Form

Use the **Additional Narrative Justification** attachment of the PHS 398 Modular Budget Form

✓ This attachment should have information on the requested costs and justification

Preparing the DMS Plan

- **Structure**: length and content

- **DMS Plan Template** - https://grants.nih.gov/grants/forms/all-forms-and-formats/data-management-and-sharing-plan-format-page

- **Sample DMS Plans** - https://sharing.nih.gov/data-management-and-sharing-policy/planning-and-budgeting-for-data-management-and-sharing/writing-a-data-management-and-sharing-plan#sample-plans

Showing Again: Overview Page has the relevant links

https://sharing.nih.gov/data-management-and-sharing-policy/planning-and-budgeting-for-data-management-and-sharing/writing-a-data-management-and-sharing-plan#after

Preparing the DMS Plan - Structure

- <u>Length</u>: recommended to be **2 pages** or less (except for large projects or those that generate many data types)

- <u>Content</u>: address **6 elements**
  1) Data Type
  2) Related Tools, Software and/or Code
  3) Standards
  4) Data Preservation, Access, and Associated Timelines
  5) Access, Distribution, or Reuse Considerations
  6) Oversight of Data Management and Sharing
  - **<u>No hyperlinks</u> (no URLs**): use **name** of the data repository, dataset identifier, or relevant resource or NIH may *withdraw* your application from consideration

From the NIH Grants Conference (Feb 1-2, 2023) – The NIH Policy for Data Management and Sharing (DMS) in Effect: Planning for Success

Elements of a Data Management and Sharing Plan

- Data type
  - Identifying data to be preserved and shared
- Related tools, software, code
  - Tools and software needed to access and manipulate data
- Standards
  - Standards to be applied to scientific data and metadata
- Data preservation, access, timelines
  - Repository to be used, persistent unique identifier, and when/ how long data will be available
- Access, distribution, reuse considerations
  - Description of factors for data access, distribution, or reuse
- Oversight of data management
  - Plan compliance will be monitored/ managed and by whom

The NIH Sharing Plan Format (Template)

DATA MANAGEMENT AND SHARING PLAN

If any of the proposed research in the application involves the generation of scientific data, this application is subject to the NIH Policy for Data Management and Sharing and requires submission of a Data Management and Sharing Plan. If the proposed research in the application will generate large-scale genomic data, the Genomic Data Sharing Policy also applies and should be addressed in this Plan. Refer to the detailed instructions in the application guide for developing this plan as well as to additional guidance on sharing.nih.gov. The Plan is recommended not to exceed two pages. Text in italics should be deleted. There is no "form page" for the Data Management and Sharing Plan. The DMS Plan may be provided in the *format* shown below.

Public reporting burden for this collection of information is estimated to average 2 hours per response, including the time for reviewing instructions, searching existing data sources, gathering, and maintaining the data needed, and completing and reviewing the collection of information. An agency may not conduct or sponsor, and a person is not required to respond to, a collection of information unless it displays a currently valid OMB control number. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to: NIH, Project Clearance Branch, 6705 Rockledge Drive, MSC 7974, Bethesda, MD 20892-7974, ATTN: PRA (0925-0001 and 0925-0002). Do not return the completed form to this address.

Element 1: Data Type
A. **Types and amount of scientific data expected to be generated in the project:**
   Summarize the types and estimated amount of scientific data expected to be generated in the project,

B. **Scientific data that will be preserved and shared, and the rationale for doing so:**
   Describe which scientific data from the project will be preserved and shared and provide the rationale for this decision.

C. **Metadata, other relevant data, and associated documentation:**
   Briefly list the metadata, other relevant data, and any associated documentation (e.g., study protocols and data collection instruments) that will be made accessible to facilitate interpretation of the scientific data.

Element 2: Related Tools, Software and/or Code:
State whether specialized tools, software, and/or code are needed to access or manipulate shared scientific

---

data, and if so, provide the name(s) of the needed tool(s) and software and specify how they can be accessed.

Element 3: Standards:
State what common data standards will be applied to the scientific data and associated metadata to enable interoperability of datasets and resources, and provide the name(s) of the data standards that will be applied and describe how these data standards will be applied to the scientific data generated by the research proposed in this project. If applicable, indicate that no consensus standards exist.

Element 4: Data Preservation, Access, and Associated Timelines
A. **Repository where scientific data and metadata will be archived:**
   Provide the name of the repository(ies) where scientific data and metadata arising from the project will be archived; see Selecting a Data Repository).

B. **How scientific data will be findable and identifiable:**
   Describe how the scientific data will be findable and identifiable, i.e., via a persistent unique identifier or other standard indexing tools.

C. **When and how long the scientific data will be made available:**
   Describe when the scientific data will be made available to other users (i.e., no later than time of an associated publication or end of the performance period, whichever comes first) and for how long data will be available.

Element 5: Access, Distribution, or Reuse Considerations
A. **Factors affecting subsequent access, distribution, or reuse of scientific data:**
   NIH expects that in drafting Plans, researchers maximize

---

the appropriate sharing of scientific data. Describe and justify any applicable factors or data use limitations affecting subsequent access, distribution, or reuse of scientific data related to informed consent, privacy and confidentiality protections, and any other considerations that may limit the extent of data sharing. See Frequently Asked Questions for examples of justifiable reasons for limiting sharing of data.

B. **Whether access to scientific data will be controlled:**
   State whether access to the scientific data will be controlled (i.e., made available by a data repository only after approval).)

C. **Protections for privacy, rights, and confidentiality of human research participants:**
   If generating scientific data derived from humans, describe how the privacy, rights, and confidentiality of human research participants will be protected (e.g., through de-identification, Certificates of Confidentiality, and other protective measures).

Element 6: Oversight of Data Management and Sharing:
Describe how compliance with this Plan will be monitored and managed, frequency of oversight, and by whom at your institution (e.g., titles, roles).

Sample Plans

9 DMS Plans are currently available at NIH, for *educational purposes only.* NIH notes that these documents are not intended to be used as templates; their use does **not guarantee** NIH approval

- Sample Plan A: **Clinical and/or MRI** data from **human** research participants
- Sample Plan B: **Genomic** data from **human** research participants
- Sample Plan C: **Genomic** data from a **non-human** source
- Sample Plan D: **Secondary** data analysis
- Sample Plan E: **Human genomic** data
- Sample Plan F: **Technology** development
- Sample Plan G: **Human clinical** and **genomics** data
- Sample Plan H: **Gene expression** analysis data from **non-human model** organism (zebrafish)
- Sample Plan I: **Human survey** data

DMS Plan – Element 1: Data Type

A. **Type and amount of scientific data expected to be generated in the project:**
   In general terms, address the ***type and amount/size*** of scientific data expected to be collected and used in the project; indicate the ***data modality*** (e.g., imaging, genomic, mobile, survey), ***level of aggregation*** (e.g., individual, aggregated, summarized), and/or the ***degree of data processing*** (i.e., how raw or processed the data will be)

B. **Scientific data that will be preserved and shared, and the rationale**
   NIH does not anticipate that all data will be preserved and shared. Decide which to preserve and share based on ***ethical, legal, and technical*** factors that may affect the extent to which data are preserved and shared. Provide the rationale for the decision

C. **Metadata and other relevant data, and associated documentation**
   Briefly list the ***metadata***, other relevant data, and any ***associated documentation*** (e.g., study protocols and data collection instruments) that will be made accessible to ***facilitate interpretation*** of the scientific data

DMS Plan – Element 2: Related Tools, Software and/or Code

- Indicate whether **specialized tools** are needed to access or manipulate shared scientific data to support replication or reuse, and **name(s)** of the needed **tool(s)** and **software**.

  - For example (1) *data will be processed and analyzed with STATA and shared in many widely accessible formats, including SAS, STATA, SPSS, dBase, Excel, and ASCII*

  - (2) *All newly developed software and code for processing and analyzing data will be distributed as version controlled, open-source code written in R or Python via GitHub, with detailed user documentation*

- If applicable, specify how **needed tools** can be **accessed** (e.g., *open source* and *freely available* or generally *available for a fee* in the marketplace, available *only from the research team*). If known, whether such tools are likely to remain available for as long as the scientific data remain available.

DMS Plan – Element 3: Standards

- Many scientific fields have developed and adopted *common data standards* - indicate what standards will be applied (i.e., data formats, data dictionaries, data identifiers, definitions, unique identifiers, and other data documentation)

    Example: RNA-Seq data: Data will be QC'd and analyzed according to ENCODE Bulk RNA-seq Data Standards. FASTQs, BAM alignment files, and TSV transcript quantifications will be shared.

- In fields were there are no common standards, indicate that *no consensus data standards exist* for the scientific data and metadata to be generated, preserved, and shared.

DMS Plan – Element 4: Data Preservation, Access, and Associated Timelines

A. Repository where scientific data and metadata will be archived:
    Will discuss in next few slides

B. How scientific data will be findable and identifiable
    For example: '*The dataset will be registered in dbGaP and assigned an ID. Data will be findable and identifiable using standard data indexing tools. We will reference the accession number(s) for our dataset(s) in all relevant future publications*

C. When and how long the scientific data will be made available
    For example: '*We will meet the data submission and release timeframes specified by the DMS Policy as described on NIH's data sharing website*';
    *The data will be available for **as long as the repository preserves** the dataset*

Data Repositories - Element 4 of DMS Plan

NIH Guidelines on Repositories

- Use <u>FOA-</u> or <u>ICO-designated</u> data repositories

- If no specific data repository is noted, select one that is <u>appropriate for the data</u> generated from the research project, as long as repository meets list of *desirable characteristics*

- Preference given to data repositories that are <u>discipline</u> or <u>data-type specific</u> to support effective data discovery and reuse

Selecting a data repository

https://sharing.nih.gov/data-management-and-sharing-policy/sharing-scientific-data/selecting-a-data-repository

NIH-Supported Repositories:
https://www.nlm.nih.gov/NIHbmic/domain_specific_repositories.html

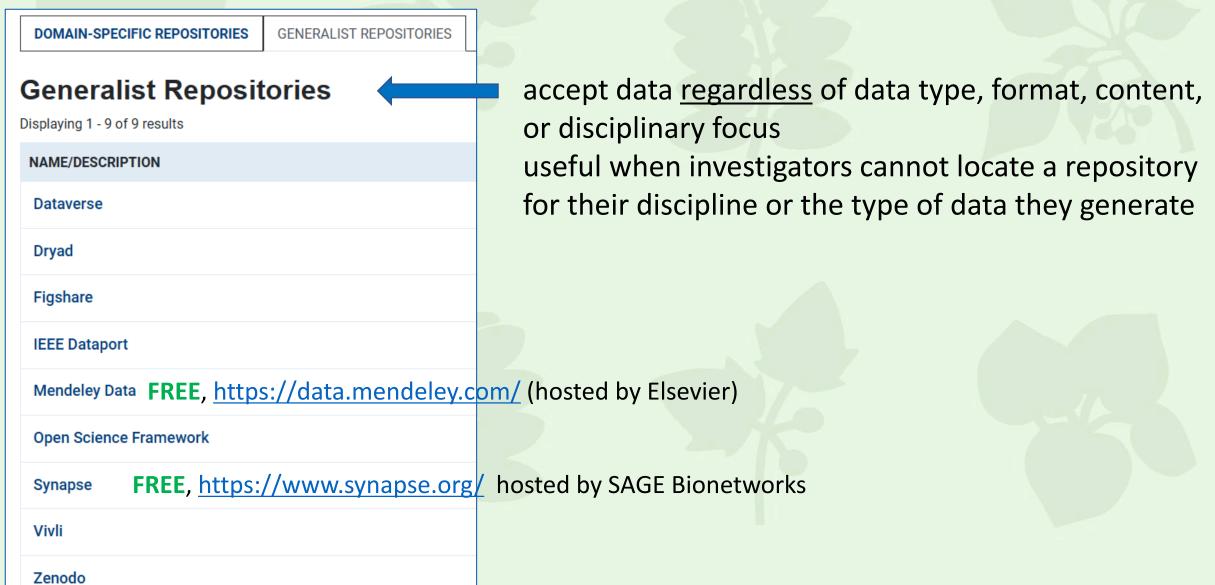# Domain-Specific Repository Spreadsheet (from previous slide)

Has information on: **URL**, **Subject Area, Access Type** *(open, registration required, controlled),* **Model System,** and **Open Data Submission**

Open data submission: accepts data from a broad set of investigators; any investigator funded by a particular NIH ICO. There are **75** repositories that allow open data submission in this list

ICO	Repository_Name	Repository_Url	Description	Data_Sub	Data_Access_Url	Subject_Area	Access_Type	Model_System	NIH_Fund	Open_Data
Common	Metabolomics Workbench	https://www.metabolomic	The NIH Common Fund's National Metabolomics Data Reposi	https://w	https://www.metabolor	Clinical research, Comput	open	human, non-huma	TRUE	TRUE
Common	Stimulating Peripheral Activity to	https://sparc.science	The SPARC Portal provides interactive access to a growing col	https://dc	https://docs.sparc.scien	Clinical research, Comput	open	human, non-huma	TRUE	TRUE
Common	Open Data Commons for Spinal Cc	https://odc-sci.org/	The Open Data Commons for Spinal Cord Injury (ODC-SCI.org	https://oc	https://odc-sci.org/	Computational biology, N	controlled, registe	non-human	TRUE	TRUE
Common	SenNet Consortium	https://data.sennetconsort	The goal of the Cellular Senescence Network (SenNet) is to i	undefined	undefined	Computational biology, Ir	controlled, registe	human, non-huma	TRUE	TRUE
Common	4D Nucleome Data Portal	https://data.4dnucleome.o	The 4D Nucleome Data Portal is an open repository for genor	https://da	https://data.4dnucleome	Imaging, Sequence biolog	registered, open	human, non-huma	TRUE	TRUE
DPCPSI	Monarch Initiative	https://monarchinitiative.o	The Monarch Initiative is an integrative data and analytic plat	undefined	https://archive.monarchi	Clinical research, Imaging	open	human, non-huma	TRUE	TRUE
DPCPSI, N	The Universal Protein Resource (U	https://www.uniprot.org/	The Universal Protein Resource (UniProt) is a comprehensive	http://ww	https://www.uniprot.org	Sequence biology	open	human, non-huma	TRUE	TRUE
FIC, NCAT	INCLUDE Data Hub	https://portal.includedcc.o	The INCLUDE (INvestigation of Co-occurring conditions acros	https://w	https://portal.includedcc	Behavioral and social scie	controlled, registe	human, non-huma	TRUE	TRUE
NCATS	BioSystics Analytics Platform (Bio§	https://mps.csb.pitt.edu/	The BioSystics Analytics Platform accesses, manages, analyze	https://m	https://mps.csb.pitt.edu	Clinical research, Comput	registered, open	human, non-huma	TRUE	TRUE
NCATS, NI	National COVID Cohort Collaborat	https://covid.cd2h.org/	The N3C Data Enclave is a secure platform through which the	https://nc	https://covid.cd2h.org/a	Clinical research	controlled, registe	human	TRUE	TRUE
NCCIH	Natural Products Magnetic Resona	https://np-mrd.org/	The Natural Products Magnetic Resonance Database is a free	https://np	https://np-mrd.org/dow	Other	open	non-human	TRUE	TRUE
NCCIH, NE	OpenNeuro	https://openneuro.org/	**A free and open platform** for validating and sharing BIDS-con	undefined	https://openneuro.org/s	Imaging	open	human, non-huma	TRUE	TRUE
NCCIH, NE	The DANDI Archive	https://dandiarchive.org/	The BRAIN Initiative archive for publishing and sharing neurc	undefined	https://dandiarchive.org	Neuroscience	open	human, non-huma	TRUE	TRUE
NCCIH, NE	The Neuroscience Multi-omic Arch	https://nemoarchive.org/	The Neuroscience Multi-omic Archive (NeMO Archive) is a da	undefined	https://nemoarchive.org	Sequence biology	controlled, open	human, non-huma	TRUE	TRUE
NCCIH, NE	Brain Image Library (BIL)	https://www.brainimagelib	The Brain Image Library (BIL) is a national public resource ena	https://w	https://submit.brainimag	Imaging, Neuroscience	open	human, non-huma	TRUE	TRUE
NCCIH, NE	Brain Observatory Storage Service	https://bossdb.org/	BossDB is a volumetric database for 3D and 4D neuroscience	undefined	https://bossdb.org/proje	Imaging, Neuroscience	open	human, non-huma	TRUE	TRUE
NCI	Cancer Nanotechnology Laborator	https://cananolab.cancer.g	caNanoLab is a data sharing portal designed to facilitate infor	https://w	https://wiki.nci.nih.gov/	Other	registered, open	human	TRUE	TRUE
NCI	The Cancer Imaging Archive (TCIA)	https://www.cancerimaging	TCIA is a service which de-identifies and publishes medical in	https://w	https://www.cancerimag	Imaging	controlled, registe	human, non-huma	TRUE	TRUE
NCI	The Network Data Exchange (NDE)	https://www.ndexbio.org	NDEx is an online commons where scientists can upload, sha	https://hc	https://home.ndexbio.c	Computational biology, O	open	human, non-huma	FALSE	TRUE
NCI, NIAID	Protein Data Bank	https://www.rcsb.org/	The mission of the RCSB Protein Data Bank (PDB) is to sustain	https://w	https://www.rcsb.org/	Computational biology, Ir	open	human, non-huma	TRUE	TRUE
NHGRI	The NHGRI Genomic Data Science .	https://anvil.terra.bio	The NHGRI Genomic Data Science Analysis, Visualization, and	https://an	https://anvilproject.org/	Clinical research, Comput	controlled, registe	human	TRUE	TRUE
NHGRI	Zebrafish Information Network (Z	https://zfin.org/	The Zebrafish Information Network (ZFIN) is the database of	https://zfi	https://zfin.org/	Sequence biology	open	non-human	TRUE	TRUE
NHGRI	WormBase	https://wormbase.org/#012	WormBase is an international consortium of biologists and cc	https://w	http://www.informatics.	Behavioral and social scie	open	non-human	TRUE	TRUE

NIH-Supported Repositories:
https://www.nlm.nih.gov/NIHbmic/domain_specific_repositories.html

DOMAIN-SPECIFIC REPOSITORIES	GENERALIST REPOSITORIES

Generalist Repositories

Displaying 1 - 9 of 9 results

NAME/DESCRIPTION

Dataverse

Dryad

Figshare

IEEE Dataport

Mendeley Data

Open Science Framework

Synapse

Vivli

Zenodo

accept data <u>regardless</u> of data type, format, content, or disciplinary focus

useful when investigators cannot locate a repository for their discipline or the type of data they generate

Mendeley Data **FREE**, https://data.mendeley.com/ (hosted by Elsevier)

Synapse **FREE**, https://www.synapse.org/ hosted by SAGE Bionetworks

Other good resources for finding suitable data repositories

- https://www.nature.com/sdata/policies/repositories#general

View data repositories
- **Biological sciences:** Nucleic acid sequence; Protein sequence; Molecular & supramolecular structure; Neuroscience; Omics; Taxonomy & species diversity; Mathematical & modelling resources; Cytometry and Immunology; Imaging; Organism-focused resources
- **Health sciences**

- https://journals.plos.org/plosone/s/recommended-repositories

Repositories by type

Biochemistry	Neuroscience	Social Sciences
Biomedical Sciences	Omics	Structural Databases
Marine Sciences	Physical Sciences	Taxonomic & Species Diversity
Model Organisms	Sequencing	Unstructured and/or Large Data

- https://data-repository-finder.ll.mit.edu/

- https://github.com – there is a **free** option (basic service)

Data Repositories - More Options

- Small datasets (up to 2 GB in size): included as supplementary material to accompany articles submitted to **PubMed** Central

- University of Hawaii Cyberinfrastructure and the UH Library System: we are exploring these options (*should meet the accessibility requirement of the NIH DMS Policy*)

DMS Plan – Element 5: Access, Distribution, or Reuse Considerations (largely applicable to human subject research)

Discuss and justify applicable ***factors*** or ***data use limitations*** affecting subsequent access, distribution, or reuse of scientific data related to *informed consent, privacy, proprietary issues, confidentiality protections*, and any other considerations that may *limit* the extent of data sharing

- <u>Informed Consent</u>: discuss disease-specific limitations, particular communities' concerns

- <u>Protections</u> for privacy, rights, and confidentiality of human research participants (i.e., de-identification, Certificates of Confidentiality, and other protective measures)

  > For example: *'only genomic data will be shared; we are not obtaining demographic or phenotypic information from BIOBANK X'*; OR *'upon receipt of an NIH Award, the data for this study will be protected by a Certificate of Confidentiality'*

DMS Plan – Element 5 *continued*

- Discuss whether **access to scientific data** derived from **humans** will be **controlled** (i.e., made available by a data repository only after approval).

- Discuss any **restrictions** imposed by federal, Tribal, or state laws, regulations, or policies, or existing or anticipated agreements (e.g., with third party funders, with partners, 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).

  Please see sample DMS Plans with human subjects and clinical data

DMS Plan – Element 6: Oversight of Data Management and Sharing

Indicate how compliance with the Plan will be monitored and managed, frequency of oversight, and by whom (e.g., titles, roles).

For example:

The **study PI** will oversee the execution of this Data Management and Sharing Plan.

Progress on data sharing will be reported in the Research Performance Progress Report. If changes occur during the project that might affect the proposed data management and sharing plans, we will provide an update and/or revision.

A Few More Notes -

- Some NIH ICOs have **ICO-specific** DMS Plan requirements

  NIMH, for example requires use of their data repository, the NIMH Data Archive (or NDA). They also require a Validation Schedule in the Plan and submission of a Data Use Agreement signed by the PI and Institutional Official

- Some **FOAs** also have **specific** requirements

- If you need assistance, please go to the Pre-Award website and submit a ticket: https://jabsom.hawaii.edu/admin/dir-rge/support-for-research/

We request those who will receive NIH feedback on their submitted DMS Plans (sometime in June-July) to please share the feedback with us so that we can develop templates and examples for future submissions from JABSOM faculty

Thank you!!

Questions?