The Lifecycle of an Expenditure: NIH DMSP Cost Budgeting and Monitoring

NU-RES Boston Workshop 2023

June 12th, 2023

2:00pm –2:45pm

Presenters:
Juan Carlos Hincapie, Director - Pre-Award Administration, College of Engineering
Fiorella Chavez, Director - Research Finance, NURES
Jen Ferguson, Head - Library Research Data Services

Agenda

1. Overview of NIH Data Management and Sharing Policy (DMSP): Scope, Applicability, Requirements, Expectations.
2. Getting started: NIH's six mandatory plan elements, Budgeting DMS costs in proposal, Considering human subjects protections.
3. Institutional tools and services for data sharing at NEU
4. NEU Implementation Efforts
5. Post-Award: Compliance and enforcement, Overall considerations
6. Q&A

NIH Data Management & Sharing Policy

The NIH has issued a Data Management and Sharing (DMS) policy, effective January 25, 2023, to promote the sharing of scientific data. There are multiple benefits to sharing scientific data, and ultimately this will facilitate the development of treatments and products that improve human health.

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

Under the DMS policy, NIH intramural investigators will:

- **Prospectively plan for the managing and sharing of scientific data**
- **Submit a DMS plan**
- **Comply with the approved plan**

Scope

- Scope
  - Applies to **ALL** research, funded or conducted in whole or in part by NIH, that results in generation of ***Scientific Data.***
  - *Not limited to awards greater than $500K*
  - Replaces 2003 policy

  Note: other Federal Agencies are considering a similar or integrated approach. All investigators and administrators should become familiar with this requirement.

- Scientific Data
  - ”[T]he recorded factual material commonly accepted in the scientific community as of sufficient quality to validate and replicate research findings, *regardless of whether data are used to support scholarly publications.”*

Scope Continued

- Scientific Data **DO NOT** include
  - Data not necessary for or of sufficient quality to validate and replicate findings
  - Lab notebooks
  - Preliminary analyses
  - Completed case report forms
  - Plans for future research
  - Peer reviews
  - Communications with colleagues
  - Lab specimens

Applicability

- **All research generating scientific data including**
  - Research Projects
  - Certain Career Development Awards (K)
  - Small Business SBIR/STTR
  - Research Centers

- Policy **DOES NOT** apply to projects not generating scientific data or non-research projects
  - Training grants (T)
  - Fellowships (F)
  - Certain non-research Career Awards (e.g., KM1)
  - Construction grants (C06)
  - Conference Grants (R13)
  - Resources (Gs)
  - Research Related Infrastructure Programs (e.g., 506)

Basic Requirements

- Include two-page written Data Management/Sharing Plan (DMSP) with application/proposal in “Other Plans” field on Form H

- Treat DMSP as a living document that should be updated throughout the award period and changes reviewed/approved by NIH

- Costs, associated with data curation and storage, **incurred during performance period** may be direct charged to grant if NOT included in IDC

Sharing Parameters

- DMSP should maximize sharing
  - As soon as possible but not later than time of publication of findings in peer-reviewed journal **OR** end of award, whichever comes first

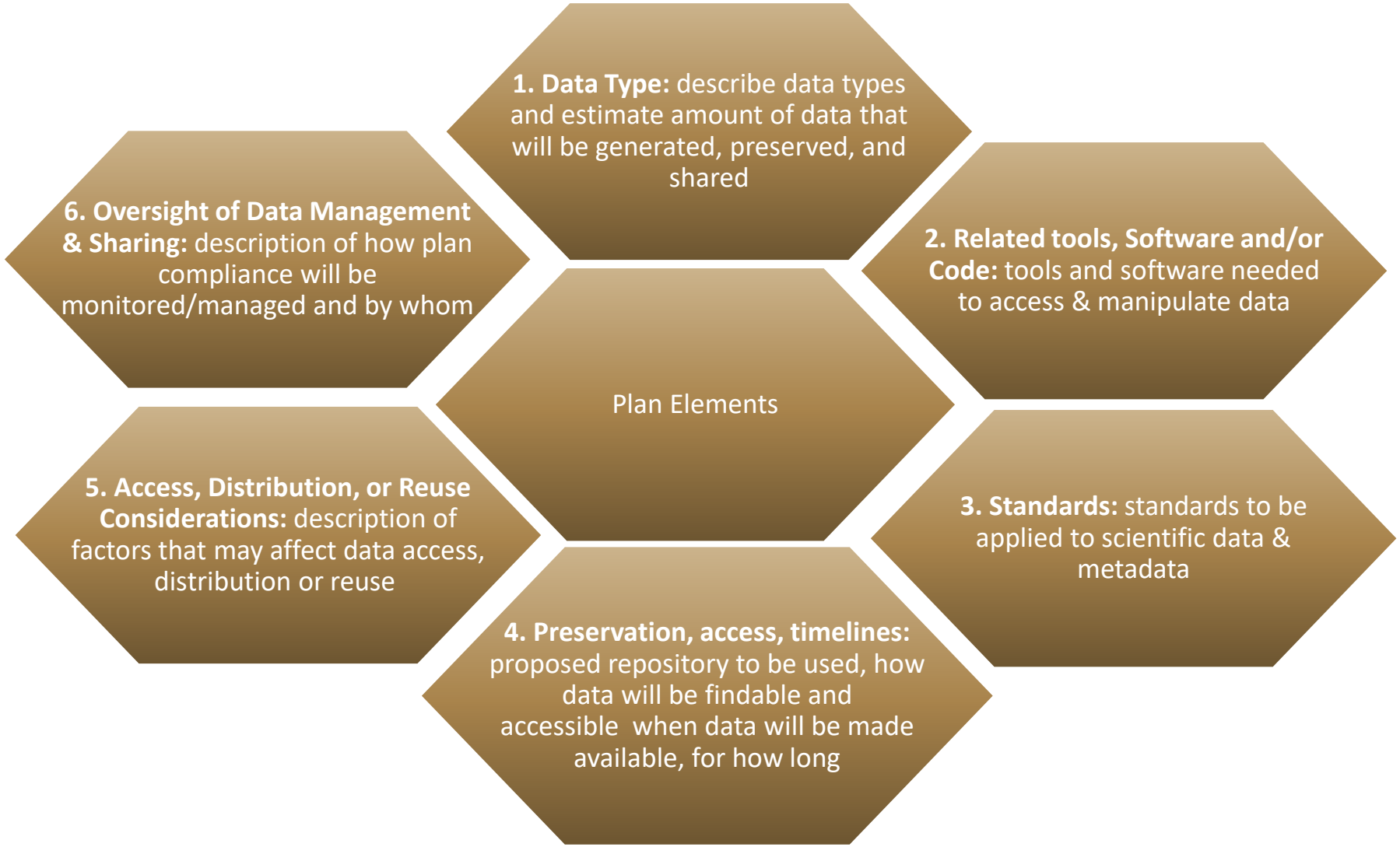
- NIH has identified reasons when acceptable to limit sharing
  - Not permitted by informed consent, government or agreement restrictions
  - Participant privacy or safety would be compromised

- NIH has identified reasons **NOT** justifiable for limiting sharing
  - Data too small
  - Anticipated data will not be widely used
  - No suitable repository

NIH PI DMSP Expectations

- PIs will plan and budget for management and sharing of scientific data (as defined by NIH)
- All competing applications/contract proposals will include a DMSP
- If the project is also subject to the NIH Genomic Data Sharing Policy, PIs will consider these requirements when drafting DMSP
  - See **GDS Policy Overview | Data Sharing (nih.gov)**
- PIs will implement DMSP as written and approved by NIH Center/Institute making changes and securing approval as data needs evolve throughout the award
- PIs are expected to maximize data sharing on a timely basis but no later than end of award
- **Compliance is the joint responsibility of PI and NU-RES**

Getting Started: What should be included in DMSP?
NIH has identified 6 mandatory plan elements



Getting Started (Continued): Choosing a Repository

- Selecting a Data Repository | Data Sharing (nih.gov)
  - Using established data repositories is highly promoted by NIH
  - For some programs and types of data, NIH will identify specific data repositories (or sets of repositories) for data preservation and sharing
  - For data generated from research for which no data repository is specified by NIH, researchers are encouraged to select a data repository that is appropriate for the data generated from the research project. See the website listed below
    - Repositories for Sharing Scientific Data | Data Sharing (nih.gov)
      - There are 70+ NIH-Supported Repositories that can be filtered through
      - Other repositories are available if you can't find one in the list above
  - PI should talk to the library or Research Computer resources about suitable public repositories

Data Sharing and Human Subjects Protections

- While the data should be made as widely and freely available as possible, the PI, IRB, and the Institution have **responsibility to protect the rights of research participants and the confidentiality of the data**.

- Researchers who are planning clinical trials and intend to share the resulting data should think carefully about the study design, **the informed consent documents, and the structure of the resulting dataset prior to the initiation of the study.**

- If research participants are promised that their de-identified data will not be shared with other researchers, the new protocol application should explain the reasons for such promises.
  - **For NIH funded research, those statements in ICFs should not be made routinely and without adequate justification**.

Data Collected Without Prior Consent for Future Use

- This refers to data previously collected from prior research studies, clinical data collected with a consent waiver, or de-identified clinical data that does not require informed consent.

  For non-US subject data, PI must ensure that data meets the definition of 'de-identified' as defined by GDPR, PIPL, etc. "The General Data Protection Regulation (GDPR) standardizes data protection law across all 28 European Union (EU) countries and imposes strict new rules on controlling and processing of personal information. It will come into effect as of May 25, 2018": GDPR_Application in Research Settings.pdf (jhsph.edu)

- Sharing options include
  - Anonymized data (including the removal of indirect identifiers)
  - De-identified data with restricted access to the repository and/or a written data-sharing agreement

Reasonable costs allowed in budget requests

- Curating Data and Documentation Costs
  - **Formatting** data according to accepted community standards
  - **De-identifying** data
  - **Preparing metadata** to foster discoverability, interpretation, and reuse
  - **Formatting data** for transmission to and storage at a selected repository for long-term preservation and access

- Local Data Management Costs
  - **Unique and specialized information infrastructure**
  - **Necessary** to provide local management and preservation (e.g., before deposit into an established repository)
  - Not included in the indirect cost rate

Reasonable costs allowed in budget requests - continued

- Preserving and Sharing Data Costs
  - **Preserving and sharing data** through established repositories, such as **data deposit fees** necessary for making data available and accessible
  - **Important:** costs must be incurred and charged against the award during the performance period

Budget Tool: NIH NDA Data Submission Cost Estimation

- Tool to estimate how much staff time and effort will be required to deposit data
  - Allows users to replace sample answers to questions listed below with research project specifics
  - Cost estimate is for the entire project not per year.

Questions

1. # of subjects?
2. # of sites collecting patient data?
3. # times data from this project submitted?
4. # data structures will be submitted?
5. # unique experiments (e.g., omics, EEG, eye tracking, fMRI) will be conducted in the study?
6. # publications analyzing human subjects data are expected for this project?
7. $ hourly rate charged for the Principal Investigator?
8. $ hourly rate charged for the Data Manager?

Tool: https://nda.nih.gov/ndapublicweb/Documents/NDA_Data_Submission_Cost_Estimation_Tool.xlsx

Where to Include DMS Budget in Application

- Identify direct costs as "**Data Management and Sharing Costs**" on **R&R Budget Form:** line item in section F. Other Direct Costs

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. Data Management and Sharing Costs	
9.	
10.	

PHS 398 Modular Budget Form: <u>within</u> Additional Narrative Justification

2. Budget Justifications

Personnel Justification		Add Attachment	Delete Attachment	View Attachment
Consortium Justification		Add Attachment	Delete Attachment	View Attachment
Additional Narrative Justification		Add Attachment	Delete Attachment	View Attachment

LVX VERITAS VIRTVS

Where to Include Budget Justification in Application

- Label "**Data Management and Sharing Justification**" within the overall budget justification
    - No more than half page (see Application Instructions for details)

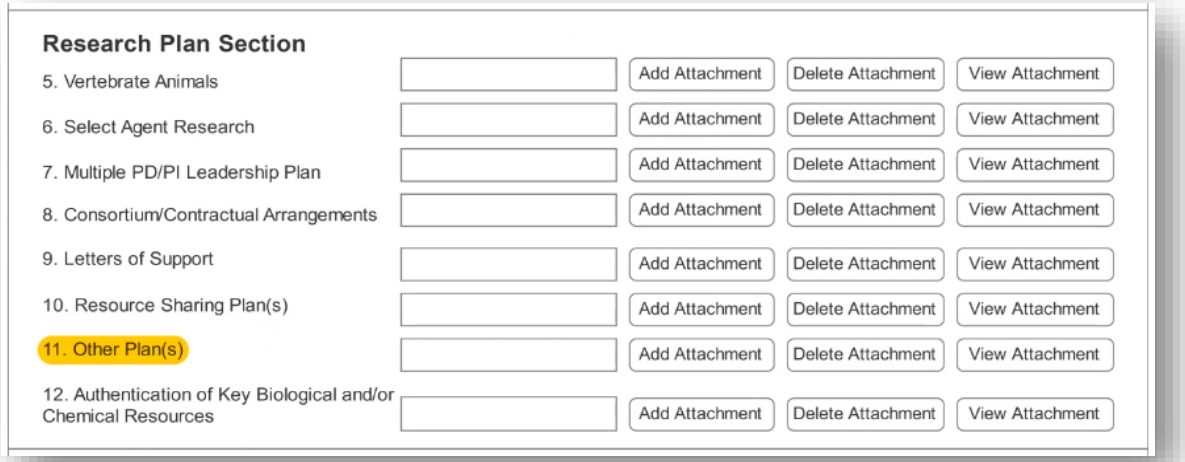
- Include activities in the DMS Plan that will incur costs
    - A **brief summary** of type and amount of scientific data to be preserved and shared
    - Name of the established **repository(ies)** and
    - General cost categories

See Budgeting for Data Management & Sharing for details

LVX
VERITAS
VIRTVS

Where to Submit DMS Plan in Application

- A new "**Other Plan(s)**" field added to the PHS 398 form to collect a single PDF attachment
- **Data Sharing Plans will no longer be submitted through the "Resource Sharing Plan(s)" field**



See Budgeting for Data Management & Sharing for details

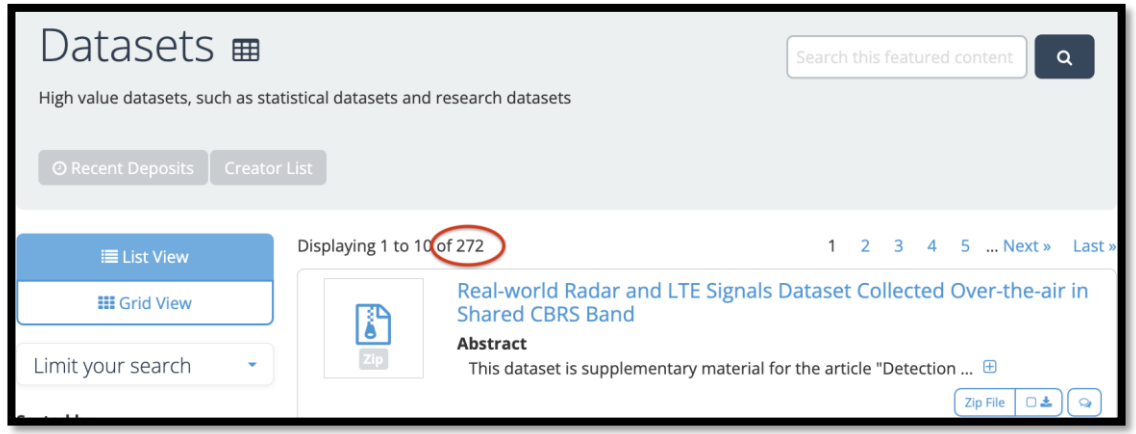
Tools for Data Sharing at Northeastern (and beyond)

NIH encourages the use of **domain-specific** repositories, but notes other potential options such as:

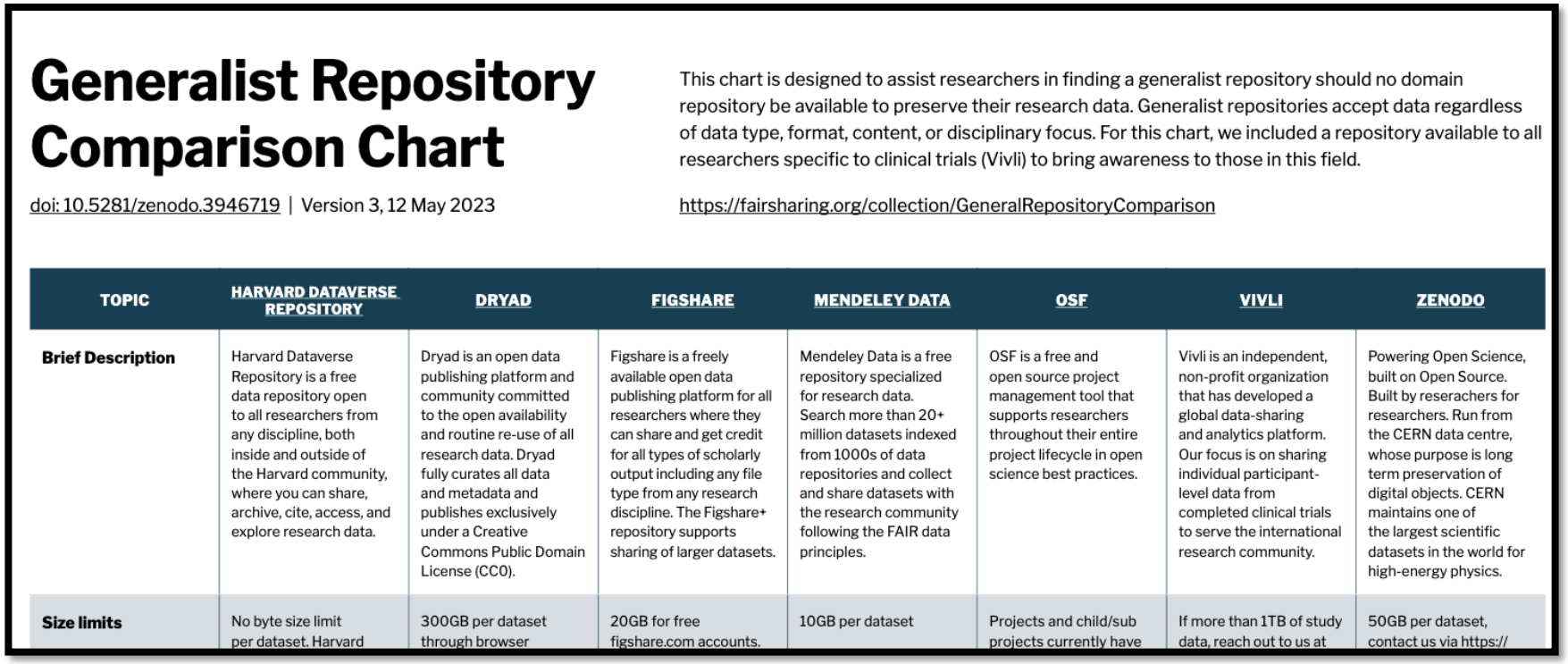
- Institutional repositories like NU's Digital Repository Service
- PubMed Central
- Generalist repositories

Digital Repository Service (DRS)

- Secure, permanent, and free storage of scholarship
- Easy deposit and discovery
- DOIs and embargos
- Usage reports to demonstrate impact
- Suggested text for DMS Plans
- repository.library.northeastern.edu



Compare generalist repository options



Generalist Repository Comparison Chart

This chart is designed to assist researchers in finding a generalist repository should no domain repository be available to preserve their research data. Generalist repositories accept data regardless of data type, format, content, or disciplinary focus. For this chart, we included a repository available to all researchers specific to clinical trials (Vivli) to bring awareness to those in this field.

doi: 10.5281/zenodo.3946719 | Version 3, 12 May 2023

https://fairsharing.org/collection/GeneralRepositoryComparison

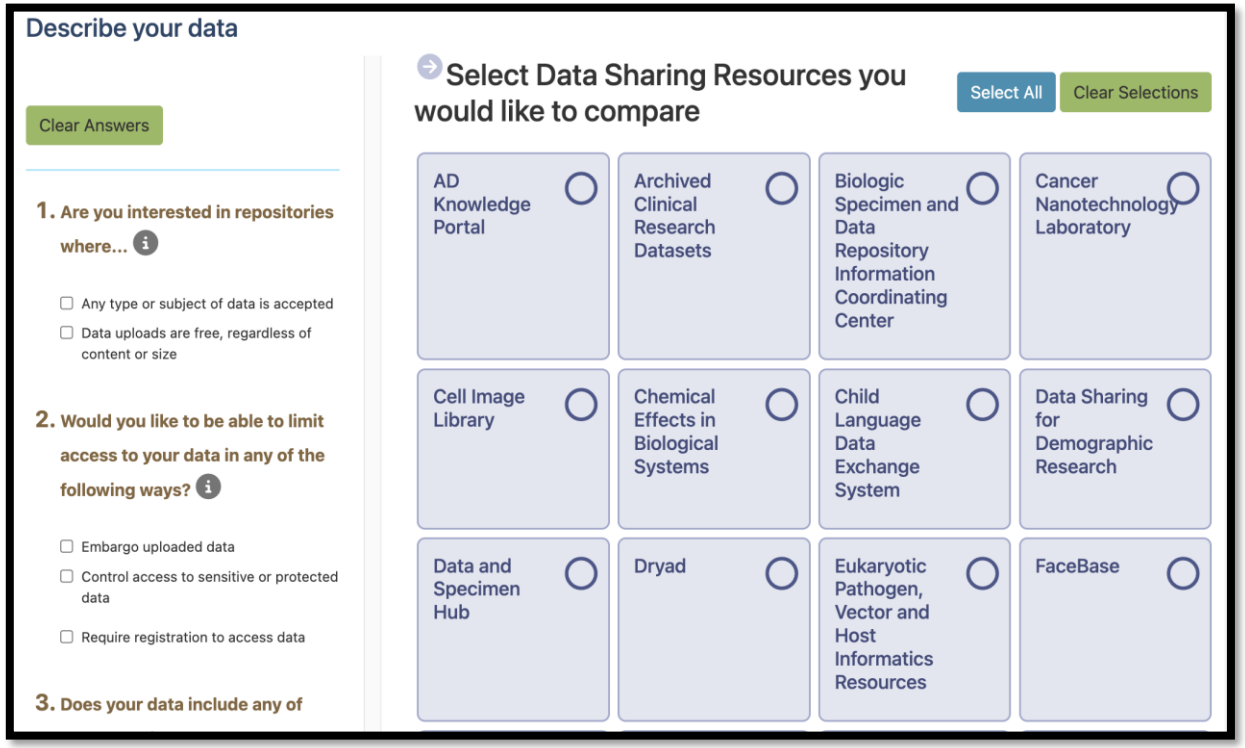
TOPIC	HARVARD DATAVERSE REPOSITORY	DRYAD	FIGSHARE	MENDELEY DATA	OSF	VIVLI	ZENODO
Brief Description	Harvard Dataverse Repository is a free data repository open to all researchers from any discipline, both inside and outside of the Harvard community, where you can share, archive, cite, access, and explore research data.	Dryad is an open data publishing platform and community committed to the open availability and routine re-use of all research data. Dryad fully curates all data and metadata and publishes exclusively under a Creative Commons Public Domain License (CC0).	Figshare is a freely available open data publishing platform for all researchers where they can share and get credit for all types of scholarly output including any file type from any research discipline. The Figshare+ repository supports sharing of larger datasets.	Mendeley Data is a free repository specialized for research data. Search more than 20+ million datasets indexed from 1000s of data repositories and collect and share datasets with the research community following the FAIR data principles.	OSF is a free and open source project management tool that supports researchers throughout their entire project lifecycle in open science best practices.	Vivli is an independent, non-profit organization that has developed a global data-sharing and analytics platform. Our focus is on sharing individual participant-level data from completed clinical trials to serve the international research community.	Powering Open Science, built on Open Source. Built by reserachers for researchers. Run from the CERN data centre, whose purpose is long term preservation of digital objects. CERN maintains one of the largest scientific datasets in the world for high-energy physics.
Size limits	No byte size limit per dataset. Harvard	300GB per dataset through browser	20GB for free figshare.com accounts.	10GB per dataset	Projects and child/sub projects currently have	If more than 1TB of study data, reach out to us at	50GB per dataset, contact us via https://

- https://zenodo.org/record/7946938

NNLM Repository Finder

Interactive tool to identify the best NIH-supported repository for your data

- https://www.nnlm.gov/finder



NEU Implementation Efforts

- Work In Progress!
- DMS pilot by a national level organization in collaboration with NIH[1]
  - Aims to generate greater consistency in requirements and mitigate administrative burden. There are two pilot phases:
  - Phase 1 (Mar23): NIH Data Management and Sharing Plan Template Pilot – review effectiveness and usability of DMSP templates, gather feedback from participating institutions, refine templates based on pilot data.
  - Phase 2 (Dec23): Cost Policies – establish common cost principles, identify types of costs required, determine how to identify additional/unforeseen costs that may be required.
- Discussions within NEU are ongoing. Updates will be shared as more information is available.

1. *Source: https://thefdp.org/default/fdp-nih-data-management-and-sharing-pilot/*

LVX VERITAS VIRTVS

Post-award: NIH Compliance and Enforcement

Are there any additional post-award financial reporting requirements associated with the Data Management and Sharing Policy?

- At this time there are no additional financial reporting requirements related to Data Management and Sharing. Reporting must be done as outlined in the Federal Financial Report (FFR) instructions. Internally, recipients should track expenditures in accordance with their institutional policies.

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 the NIH ICO adding special Terms and Conditions of Award or terminating the award. If award recipients are not compliant with Plans at the end of the award, noncompliance may be factored into future funding decisions.

- For contracts, noncompliance with the DMS Plan will be handled in accordance with the terms and conditions of the contract and applicable Federal Acquisition Regulation (FAR).

Source: https://sharing.nih.gov/faqs#/data-management-and-sharing-policy.htm?anchor=56553 (Updated 2/2/23)

Post-Award: Overall Considerations

- Final proposal submissions (budget forms and budget justification) will be critical to identifying allowability of DMS costs during the life of the award.
- All costs for data management and sharing activities, including personnel costs, must be included in the single line item on the R&R Budget Form in section F. Other Direct Costs. Supporting details, including a breakdown of any personnel effort, must be included in the budget justification.
- If there are no anticipated data management and sharing costs, this should be noted in the budget and justification.
- All allowable costs submitted in budget requests must be incurred (e.g., curation fees, data repository fees) **during the performance period**, even for scientific data and metadata preserved and shared beyond the award period.
- Prior approval from NIH may be required if data management and sharing costs require changes.

Source: https://sharing.nih.gov/faqs#/data-management-and-sharing-policy.htm?anchor=56553 (Updated 2/2/23)

Q&A

LVX
VERITAS
VIRTVS

What about costs associated with data storage or related activities after the end of the award? How should these be handled?

- Include in the budget for the last year of the award.
- Funds must be expended BEFORE the end of the performance period.
- NEU will need to consider developing a process to manage pre-payment to cover cost in out years to the term when storage costs will be incurred.

Does uploading data in a repository alleviate a PI's need to retain data?

- PIs are still required to retain and maintain NEU data, including original data, with all identifiers where applicable, collected under a NEU research project or its auspices or collected using NEU resources.

- These data are owned by NEU and must be retained by the PI in a location that meets NEU Information Security requirements and in accordance with the Institute's research data retention policy.

  See Northeastern Data Classification Guidance:

  Data Classification - Home1 (sharepoint.com)

Is the new DMS Plan separate or in addition to the existing Resource Sharing plan? Are both required? Where should they be uploaded in grant application?

- The DMS Plan is separate from Resource Sharing.

- The DMS Plan, including Genomic Data Sharing where applicable, is now submitted as an attachment under “Other Plans.”

The DMP Tool format does not create a document in the NIH format (e.g., wrong size font, headers, footers.) How can we modify the document to meet NIH formatting?

- The DMP Tool defaults to PDF but includes an option to use DOCX which allows the DMP to be reformatted to meet NIH requirements.

In addition to the DMP Tool, are there DMS Plan templates for PI use?

- NIH posted resources to assist PIs in drafting their DMS Plan. These range from identification of the six required elements with detailed information on each element to six sample plans in the areas listed below

  - Clinical and/or MRI data from human research participants
  - Genomic data from human research participants
  - Genomic data from non-human source
  - Secondary data analysis
  - Human genomic data
  - Technology development

- NIH sample plans may include specific I/C expectations. PIs should consult the FOA to determine the expectations of the I/C that will fund their project.
- Link: Writing a Data Management & Sharing Plan | Data Sharing (nih.gov)

How does the new NIH Data Management and Sharing Policy apply to applications where NEU is not the Prime?

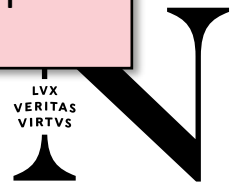
- Prime is responsible for DMS Plan development, application submission, and implementation throughout the award
- Generally, the subaward PI is not required to include DMS Plan in Statement of Work
- Subaward PI is responsible for familiarizing themselves with and implementing the Prime's DMS Plan at NEU once an award is made.

What about when NEU is the Prime?

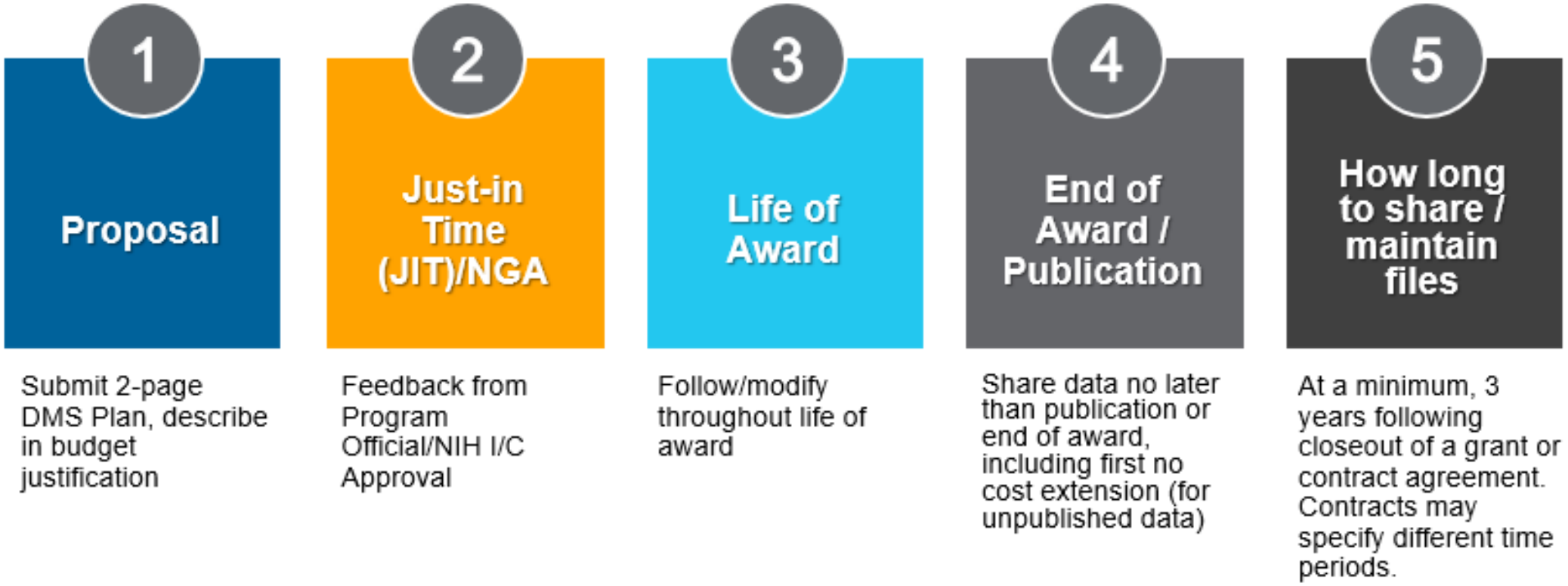
- NEU PI is responsible for
  - Developing the DMS Plan to meet NEU needs and those of collaborators
  - Including costs in budget
  - Submitting Plan and budget with application
  - When award is made
    - Implementing locally
    - Educating subaward collaborators on Plan requirements
    - Monitoring compliance with the plan generally, including sub-recipient activities.

Quick Summary

When	The effective date of the DMS Policy is January 25, 2023.
What	Plan for managing and sharing scientific data that addresses 6 recommended elements. Applications subject to Genomic Data Sharing should include GDS.
Who	All NIH funded projects that to generate scientific data. Complete list of NIH activity codes subject to the DMS Policy application.
Where	NIH strongly encourages investigators to use an established repository to maximize the sharing of scientific data. Note that some funding opportunities or programs may require the use of a specific repository.



PI Quick Guide: Applicability and Timing



NIH Implementation Details:

2023 NIH Data Management and Sharing Policy | NIH Office of Intramural Research

NOT-OD-22-189: Implementation Details for the NIH Data Management and Sharing Policy

NEU Resources and Help

- Data Classification Guidance:
  - https://northeastern.sharepoint.com/sites/DA-DataClassification
- Libraries:
  - https://subjectguides.lib.neu.edu/datamanagement/DMPs
  - There is also a training on the new policy
  - NU library has a repository
- Research Computing:
  - https://rc.northeastern.edu
- Research Compliance:
  - https://research.northeastern.edu/nu-res/compliance/

Questions/Suggestions

Thank you!