NIH Data Management & Sharing Policy

Changes Effective January 25, 2023

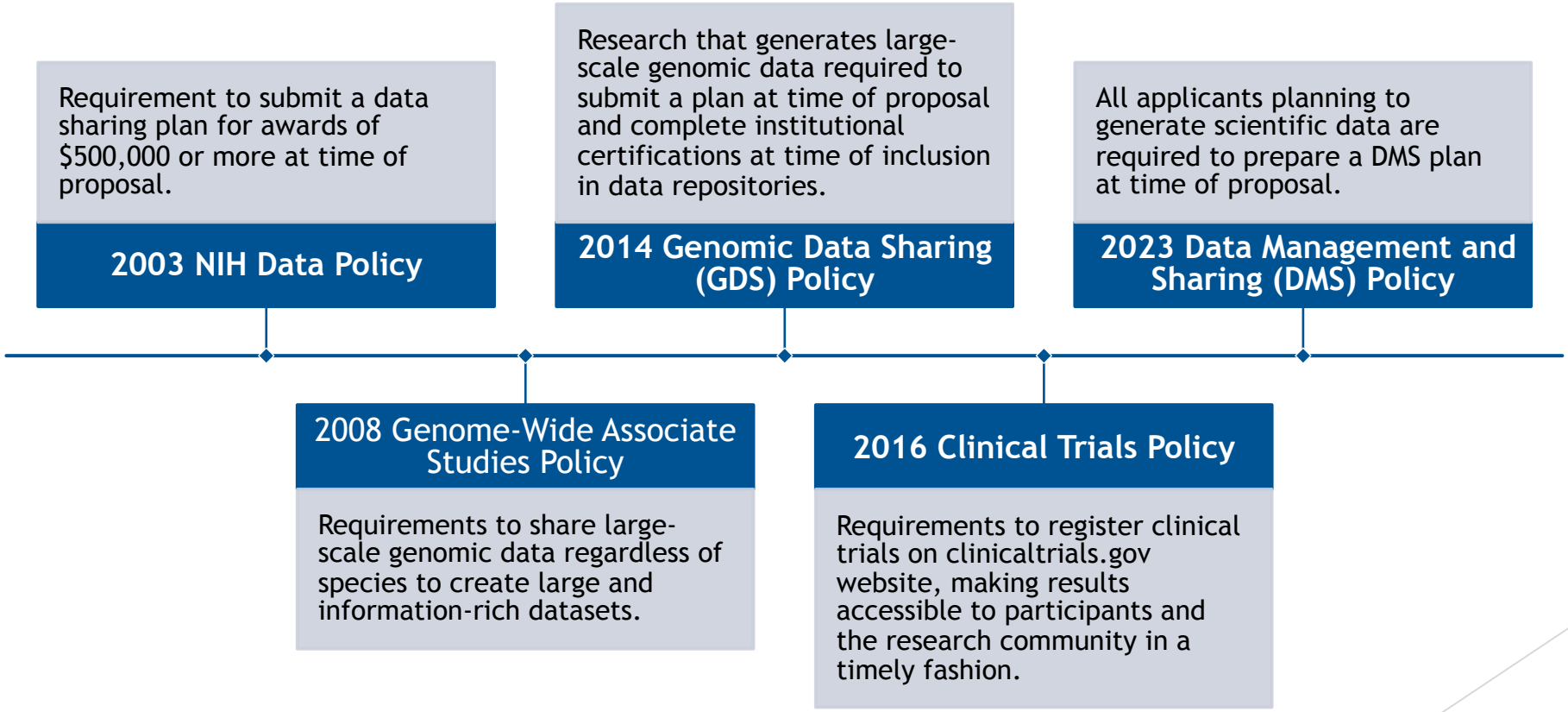
Sarah Saxton
Executive Director, Office of Sponsored Programs
Cassandra Myers
Associate Vice Provost of Research Compliance and Regulatory Affairs

DREXEL UNIVERSITY
Office of
Research &
Innovation

Objectives

- Provide regulatory background, and how to determine if your ongoing or proposed research is subject to the DMS Policy.
- Provide steps and tools to identify and develop appropriate methods/approaches and repositories for managing and sharing scientific data.
- Discuss how to estimate and request funds for data management and sharing, and submit with proposal.
- Discuss when to provide updates and incorporate changes over the course of the project.

NIH Policy Progression



Goals of 2023 NIH Data Sharing & Management Policy

- "The National Institutes of Health (NIH) Policy for Data Management and Sharing reinforces NIH's longstanding commitment to making the results and outputs of NIH-funded research available to the public through effective and efficient data management and data sharing practices.
- 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 FAIR data principles."

-Final NIH Policy for Data Management and Sharing (Notice Number: NOT-OD-21-013

NIH Data Management and Sharing Policy

Brief Overview

2023 Data Management and Sharing Policy

Purpose	"making the results and outputs of NIH-funded research available to the public through effective and efficient data management and data sharing practices."
Application	"applies to all research, funded or conducted in whole or in part by NIH, that results in the generation of scientific data"
Proposal/Plan	"required to submit a Plan to the funding NIH ICO" as part of the proposal – varied by funding instrument – and "should explain how scientific data generated by research projects will be managed and which of these scientific data and accompanying metadata will be shared."
Enforcement	"non-compliance with the NIH ICO-approved Plan may be taken into account by NIH for future funding decisions for the recipient institution"

Source: **NIH** NOT-OD-21-013

New NIH Data Management and Sharing Policy

- NOT-OD-21-013
- Creates a consistent minimum expectation for all research supported by NIH
- Applies to all research, funded or conducted in whole or in part that generates scientific data
  - **NIH Definition of Scientific Data:** the recorded factual material commonly accepted in the scientific community as of sufficient quality **to validate and replicate research findings**, regardless of whether the data are used to support scholarly publications

Exclusions from Data Management and Sharing Policy

- Scientific Data do **NOT** include:
  - Data **not** necessary for or of sufficient quality to validate and replicate research findings
  - Laboratory notebooks
  - Preliminary analyses
  - Completed case report forms
  - Drafts of scientific papers
  - Plans for future research
  - Peer reviews
  - Communications with colleagues
  - Physical objects (e.g., laboratory specimens)

Projects Under DMS Policy

DMS Policy Applies	DMS Policy does NOT Apply
All research generating scientific data, including but not limited to:	Research projects not generating scientific data or non-research projects, including but not limited to:
• Research Projects (Rs)	• Training (Ts)
• Certain Career Development Awards (Ks)	• Certain non-research Career Awards (e.g., KM1)
• Research Centers	• Fellowships (Fs)
• Small Business SBIR/STTR	• Construction (C06)
	• Conference Grants (R13)
	• Resources (Gs)

Effective Date

- Policy applies to all NIH funding mechanisms

Extramural	• Competing applications submitted for January 25, 2023 and subsequent receipt dates • Non-competing awards will not immediately become subject to Policy
Contracts	• Proposals submitted on or after January 25, 2023
Intramural	• Projects conducted on or after January 25, 2023
Other funding agreements (e.g., Other Transactions)	• Executed on or after January 25, 2023, unless otherwise stipulated by NIH

What if my application is prior to 01/25/2023 or its already been awarded?

- Application for receipt prior to 01/25/2023 are required to continue to follow the 2003 Data Sharing Policy.
  - Data Sharing Plans are only required for requests for funding, grants, intramural research, contracts or other agreements if they are $500,000 or more per year.
  - **If you have a competing renewal** (i.e., the projects is extended with additional funds) whichever policy was in effect at time of the competing renewal submission will apply.
  - **If you have a non-competing renewal or no cost extension** (i.e., the projects is extended without additional funds) whichever policy was in effect at time of the last submission that included funds will apply.

Elements of Data Sharing and Monitoring Plan

Section Source: NIH-Writing a Data Sharing & Management Plan

Element 1: Data Type

Summarize

- Type (e.g., fMRI images, V02 max results, PHQ9 scores)
- Amount (e.g., 50 human participants, 100 hispid cotton rats, 30 cell lines)
- Data Modality (e.g., survey, imaging, genomic)
- Level of Aggregation (e.g., individual, summarized, aggregated)

Preserved/Shared

- Described what scientific data will be preserved and shared.
- NIH does not expect all scientific data to be preserved and shared in a repository.
- Researchers should consider ethical, technical, and legal factors

Facilitate Interpretation

- Brief list of metadata,
- Other related data, and
- Associated documentation (e.g., protocol, data collection instrument) that will facilitate interpretation of the scientific data

Element 1: Data Type Examples

"Data to be preserved and shared"

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.

Demographic, clinical, and MRI, ^{1}H fMRS and fMRI imaging data will be acquired from 110 affected youth and 110 matched healthy controls (described in detail in sections C.3 and C.4 of this application). All data will be de-identified prior to receipt by the repository, but the information needed to generate a global unique identifier for the NIMH Data Archive (NDA) will be collected for each subject.

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.

Sufficient data from this project will be preserved to enable sharing via NDA data of sufficient quality to validate and replicate research findings described in the Aims. NIMH requires data measured from human subjects to be shared using the NDA.

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.

In addition to the subject level data described above, all ^{1}H fMRS and fMRI task related paradigm designs and experiment definitions will be deposited in the NDA.

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.

The data to be shared will include MRI images and clinical assessments from human research participants. This application is focused on secondary data analysis from existing data but will also deposit privately held data to a public repository. The existing data is available from the NIMH Data Archive (NDA) in collections 2134 (148 subjects) and 2433 (47 subjects). In addition, we have data from a previous study involving 155 research participants with major depressive disorder that have not yet been shared with the research community but will be uploaded to NDA during the second quarter of the first year of funding. As discussed in the application, structural MRI scans are available for time points before and after treatment along with relevant clinical data.

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.

This is a secondary data analysis application, so new data is not being measured. Much of the data is already available through NDA. Clinical and imaging data from 155 new subjects will be shared.

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.

Preparation for submitting existing data to NDA is largely complete. Within the first six months following the award, we will submit the Data Submission Agreement to NDA and will create the Data Expected list (see Standards section) in our new NDA Collection. The policies of our institution mandate that exact dates will not be shared (see Access section).

Element 2: Tools, Software, Code

"Tools & software needed to access & manipulate 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.

The basic statistical analyses described in the application will be done using R. We plan to use the MRI data analysis tools in the FMRIB Software Library (FSL) for multi-level modeling of group effects. BrainVoyager software will be used for anatomical segmentation to isolate regions of interest within individual subjects, and the AI-powered analyses described in the application will use custom code written with the PyTorch library for Python. R, FSL, Python, and PyTorch are all freely available to the research community. BrainVoyager is commercial software, with licenses available for purchase from https://www.brainvoyager.com/.

All R and Python code (including trained model weights) will be available on our lab Bitbucket page (located at https://www.thelab.edu/) no later than when publications are submitted. The Bitbucket page is publicly assessable and will be hosted for at least 5 years after the grant award ends.

Element 3: Standards

“Standards to be applied to scientific data & metadata”

Examples:

- Data Formats
- Data Dictionaries
- Data Identifiers
- Definitions
- Unique Identifiers
- Other Data Documentation

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.

Participant age, sex, ethnicity, height, weight, socioeconomic status, and other demographic data will be collected using the following instruments as defined in NDA:

1) Research Subject and Pedigree (ndar_subject01)
2) Demographics Short Form (demsf01)
3) Ethnic Group Questionnaire (ethgrp01)
4) Height and Weight (height_weight01)
5) Hollingshead Socioeconomic Rating Scale (ses01)
6) Pubertal Development Scale (pds01)
7) Edinburgh Handedness Inventory (edinburgh_hand01)
8) WASI-2 (wasi201).

In compliance with NOT-MH-20-067, the following data will be collected to facilitate aggregation of this data set with other data sets:

1) DSM Crosscutting for Youth (dsm5crossch01)
2) RCADS-25 (rcads2501)

The clinical assessments we plan to collect for this study include:

1) Kiddie-SADS-Present and Lifetime Version (ksads_pl01)
2) Children’s Yale-Brown Obsessive Compulsive Scale (cybocs01)
3) Schedule for Obsessive-Compulsive and Other Behavioral Syndromes (Hanna. Schedule for Obsessive-Compulsive and Other Behavioral Syndromes, Ann Arbor: University of Michigan, 2010, new data dictionary will be defined in NDA)
4) Dimensional Obsessive Compulsive Scale (docs01)
5) Yale Global Tic Severity Scale (yale01)
6) Child Behavior Checklist (cbcl01)
7) Multidimensional Anxiety Scale for Child Parent and Self (masc_p01)
8) Conners 3 (conners3_ps01)
9) Adolescent Depression Rating Scale (doi:10.1186/1471-244X-7-2, new data dictionary will be defined in NDA)

^{1}H fMRS and fMRI data will be shared with the Image (image03), Imaging Work Flow (iwf01), and Imaging Collection (imagingcollection01) data dictionaries as defined in NDA.

Element 4: Data Preservation, Access, and Associated Timelines

"Repository to be used, persistent unique identifier, when/how long data will be available"

NOTE: NIH encourages data to be shared as soon as possible, but no later than the time of publication or end of performance period.

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).

All data will be deposited to NDA starting 12 months after the award begins and will be deposited every six months thereafter following the usual NDA data submission dates.

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.

Data will be findable for the research community through the NDA Collection that will be established when this application is funded. For all publications, an NDA study will be created. Each of those studies is assigned a digital object identifier (DOI). This data DOI will be referenced in the publication to allow the research community easy access to the exact data used in the publication.

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.

The research community will have access to data when the award ends. As required by NDA, studies will also be created that contain the data used for every publication. Those studies will be shared when the pre-print is available. NDA studies have digital object identifiers (DOI) to aid in findability. We will include that DOI in relevant publications. NDA will make decisions about how long to preserve the data, but that data archive has not deleted any deposited data up to now.

Element 5: Access, Distribution, or Reuse Considerations

"Description of factors for data access, distribution, or reuse"

Consider the following:

- Informed Consent
- Privacy and Confidentiality Protections
- Limitations by state, tribal, or federal law
- If Data Repository has appropriate controls

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).

All data will be deposited to NDA starting 12 months after the award begins and will be deposited every six months thereafter following the usual NDA data submission dates.

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.

Data will be findable for the research community through the NDA Collection that will be established when this application is funded. For all publications, an NDA study will be created. Each of those studies is assigned a digital object identifier (DOI). This data DOI will be referenced in the publication to allow the research community easy access to the exact data used in the publication.

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.

The research community will have access to data when the award ends. As required by NDA, studies will also be created that contain the data used for every publication. Those studies will be shared when the pre-print is available. NDA studies have digital object identifiers (DOI) to aid in findability. We will include that DOI in relevant publications. NDA will make decisions about how long to preserve the data, but that data archive has not deleted any deposited data up to now.

Element 6: Oversight of Data Management and Sharing

"Plan compliance will be monitored/managed & by whom"

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).

The Office of Research and Innovation at Drexel University will be conducting routine monitoring visits to a percentage of studies that have indicated a Data and Sharing Management Plan on a yearly basis and have received or been awarded funds through NIH and subject to the NIH Data Sharing Policy. The study personnel will also look for the NDA data DOIs from NDA Studies and will include that information in the annual progress report upon submission.

Validation Schedule (this section is required by NIMH)

If funded, within 6 months of the Notice of Award date we will submit a Data Submission Agreement signed by the principal investigators and an institutional business official, as well as define and complete the Data Expected section of this project. Uploads of all initial demographic, clinical, and raw structural MRI, ^{1}H fMRS and fMRI research data will be completed using the second submission cycle deadline following the Notice of Award date. Subsequent data uploads will be harmonized, validated, and submitted biannually on the standard January 15th and July 15th submission deadlines.

We also plan to use the NDA validation tool as a quality control measure in the laboratory. The data manager in charge of submitting data to NDA will help researchers in the group validate their data once every month.

Genomic Data Subject to DMS Policy

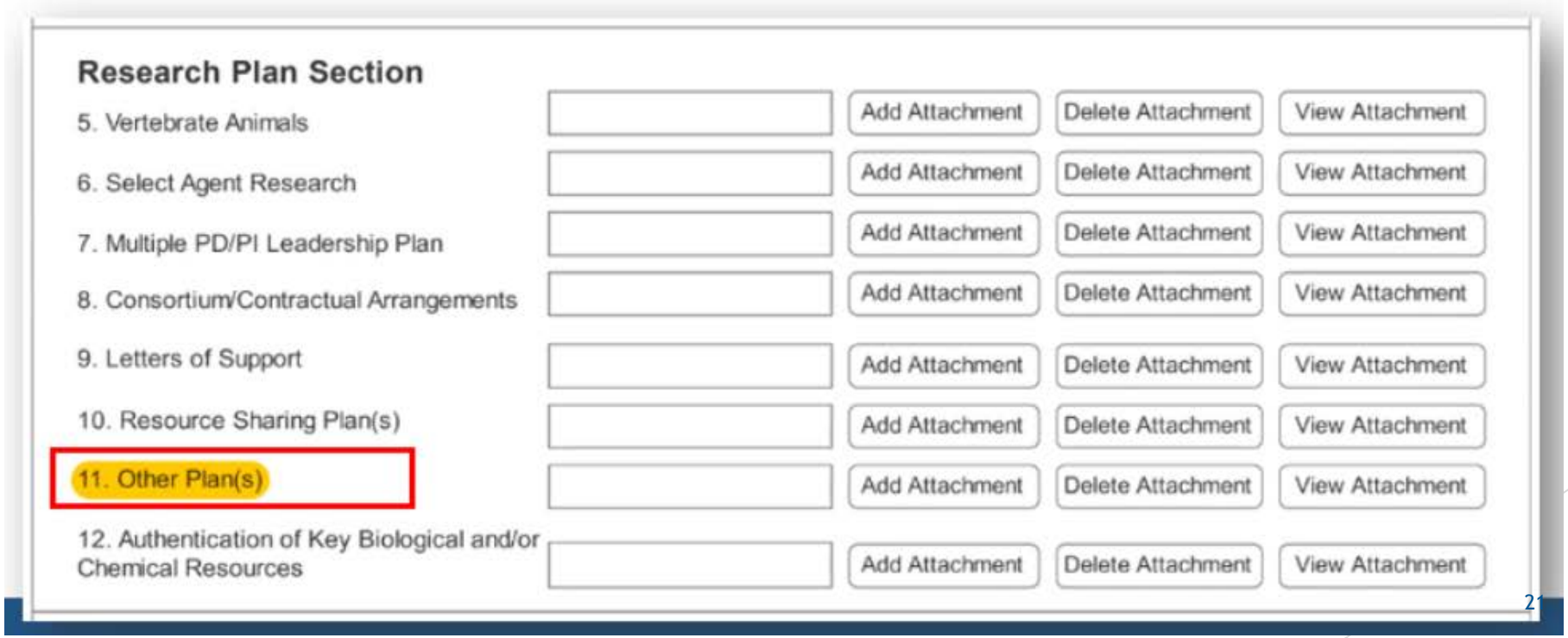
- For Human Data:
  - Investigators are expected to continue to submit data and follow the 2014 Genetic Data Sharing Policy, in addition to the DMS Policy, including Genetic Data Sharing language in the consent form (Template Language provided by Drexel HRPP)
  - This continues to include institutional certifications with pre 2015 and post 2015 Genetic Data (Review and Signature Required by Drexel HRPP)
- For Non-Human Genetic Data:
  - Investigators may submit data to any widely used repository
  - Expected to be shared as soon as possible, but no later than the time of an associated publication, or end of the performance period, whichever is first like other data subject to DMS Policy

Submission of DSM and Review by NIH

Section Source: NIH-Writing a Data Sharing & Management Plan

How to Submit

- New **"Other Plan(s)"** field will be added to the PHS 398 Form for extramural (grants)
  - Data Sharing Plans & Genomic Data Sharing Plans will no longer be submitted to the "Resource Sharing Plan(s)" field



DREXEL UNIVERSITY
Office of
Research & Innovation

Format & Location for Other Type of Awards

- Up to 2 pages
- Optional format page by Fall 2022
  - Will be added to list of Format Pages
  - Incorporated into FORMS-H application instructions
- **Location for non extramural grants**
  - **Extramural (contracts)-** As part of the technical evaluation
  - **Other Funding agreements-** Prior to the release of funds

Cost Considerations

- Costs associated with the DMS will be allowed within the project budget at time of proposal submission
  - Costs will be a line item within the project budget and a description will need to be included in the budget justification
- Costs must be incurred within the project period, when awarded

Allowable Costs	Unallowable Costs
• Curating data/developing supporting documentation	• Infrastructure costs typically included in indirect costs
• Preserving/sharing data through repositories	• Costs associated with the routine conduct of research (e.g., costs of gaining access to research data)
• Local data management considerations	

NIH Budget

- Data repository costs will need to be added under the "other direct costs" portion of the NIH R&R Budget

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 repository costs	10,000.00
9.		
10.		

What if I need to store data at a repository after the award ends and there is a continuing cost for storage?

- Costs must be incurred within the project period, when awarded
  - Budget for it in the proposal
  - Provide justification
  - At award, pay for the associated costs before the end of the project period by invoice, this can include future dates.

DMS Plan Approval

- **At proposal phase:** NIH staff determines if plan is acceptable or not
  - Peer reviewers only review budget to see if it is reasonable
  - DMS Plan will need to include a data repository
- **At award stage:** NIH must approve prior to award
  - Additional information or revised DMS Plan collected at JIT
- **Approved DMS Plan** becomes a **term & condition** of the award
  - Grantee reports progress of DMS Plan during RPPR (annual, interim, final)
  - NIH reviews annually for compliance

Additional Considerations

Section Source: Selecting a Data Repository

Selecting a Data Repository

- NIH, Institutes, Centers, or Funding Opportunity Announcements (FOA) may identify/require specific repositories.
  - When identified they should be used.
- When NIH, Centers, or FOA's don't designate or require a specific repository consider an NIH-supported repository, if applicable:
  - Over 100 repositories supported by NIH that are available to researchers
  - https://sharing.nih.gov/data-management-and-sharing-policy/sharing-scientific-data/repositories-for-sharing-scientific-data
- If no appropriate discipline or data-type specific repository is available from NIH, researchers should consider a variety of other potentially suitable data sharing options:
  - Small datasets (up to 2 GB in size) may be included as supplementary material to accompany articles submitted to PubMed Central (instructions).
  - Data repositories, including generalist repositories or institutional repositories, that make data available to the larger research community, institutions, or the broader public.
  - Large datasets may benefit from cloud-based data repositories for data access, preservation, and sharing.

Characteristics for Data Repositories

How can I appropriately store & share data/Tools Available at Drexel?

- Examples- (Also see “Drexel Resources”):
  - LabArchives (Existing)
  - Office 365 (Existing)
  - REDCap (Existing)
  - Other Tools/Software/Platforms previously assessed and designated by IT
  - Globus (Currently Being Evaluated)
  - Tools/Data Sharing/Platforms with Collaborators (Needs to be Assessed by IT prior to Use)

Consent Language for Human Data

- No “Required Language” from NIH, however NIH expects data management and sharing as described in the DMS plan be consistent with the application, and have suggested language based on study specifics:
  - NIH Repository or Non-NIH Repository
  - De-identified or Identifiable
  - Publicly Available Database or Restricted Database
  - Result Sharing
  - Ability to Withdraw
  - Risks of Broad Sharing
  - Optional Data Sharing vs. Required

- Drexel Consent Form Update (01/01/2023)
  - Incorporation of Suggested Data Sharing & Management Policy Language from NIH
  - Additional Minor Updates
  - Utilize the new template/don’t reuse past consent forms to mitigate back and forth with the IRB.

Do I need to submit a modification to the IRB for an existing study?

- As NIH has not set required language, you do not need to submit a modification or amendment unless your subject to the NIH DMS Policy (effective 01/23/2023) and your plan has changed and is no longer accurate in your consent or IRB Protocol.
  - Please ensure in your modification/amendment description why participants either should be reconsented or not reconsented as part of the modification.

- A modification/amendment is also required if you change and receive NIH approval to your DMS Plan and the Consent and IRB application is no longer accurate

NIH Resources

- **NIH Announcements**
  - New DMS Policy: NOT-OD-21-013
  - Additional Implementation Details: NOT-OD-22-189
  - Elements of an NIH Data Management & Sharing Plan: NOT-OD-21-14
- **Repository Information**
  - Data Repository Selection for NIH-Funded Research: NOT-OD-21-16
  - Selecting a Data Repository
  - Repositories for Sharing Scientific Data
- **Budgeting**
  - Allowable Costs for Data Management & Sharing: NOT-OD-21-15
  - Budgeting for Data Management & Sharing
- **Additional Information**
  - 2023 Data Management & Sharing Policy FAQs
  - NIH Data Sharing Website
  - Policy Decision Tool
  - NIH Institute and Center Data Sharing Policies
- **Training**
  - Recorded Webinars

Drexel Resources

- Data Management Plan Resources are available through the library
  - Research Data Management Plan Resources
  - DMP Tool
  - Exploro
- Electronic Repositories
  - Exploro (Libraries)
  - Electronic Research Notebooks
- To be developed
  - Metadata guides
  - Repository guidance
  - Storage/Security guidance
  - FAQs

Questions?