NIH Policy for Data Management and Sharing

NIH's new Data Management and Sharing (DMS) policy applies to all NIH-supported research that results in the generation of scientific data. As such, a DMS Plan is required when applying for funding on or after January 25, 2023.[1] It is important to note that after this date, NIH will no longer be collecting separate genomic data sharing plans. Genomic data sharing considerations, such as where and when genomic data will be shared, will be expected to be addressed in DMS Plans.[2]

NIH expects institutions and investigators to carefully consider and craft a DMS Plan throughout the planning and proposal development stages (including budget development) as well as comply with the plan during the funding period. Data should become accessible "as soon as possible, and no later than the time of an associated publication, or the end of the award/support period, whichever comes first."

Furthermore, NIH encourages the use of established repositories that are discipline or data-type specific. See NIH's "Open Domain-Specific Data Sharing Repositories" for a list of repositories open for submitting and accessing scientific data.

Building an Appropriate DMS Plan

Case strongly recommends investigators utilize DMPTool.org to efficiently create data management plans that comply with the trans-NIH policy and individual funding agencies' requirements. These requirements are outlined on the NIH Institute and Center Data Sharing Policies webpage.

Overview of DMS Plan Components

Component	***Brief Description***
Data Type	Types and estimated amount of scientific data to be generated and/or used in the research and a description of which data will be preserved and shared
Related Tools/Software and/or Code if applicable	Specialized tools and software needed to access or manipulate data
Standards	Standards to be applied to the data and metadata
Data Preservation, Access, and Associated Timelines	Repository to be used, persistent unique identifier, and when/ how long data will be available
Access, Distribution, or Reuse Considerations	Factors affecting access, distribution, or reuse of data
Oversight of Data Management and Sharing	Explanation of how compliance with the DMS Plan will be monitored and managed and by whom

[1] Final NIH Policy for Data Management and Sharing (NOT-OD-21-013) and sample DMS Plans can be found on NIH's main webpage outlining the new policy.
[2] Implementation Changes for Genomic Data Sharing Plans (NOT-OD-22-198)

Detailed Description of DMS Plan Components

1. Data Type

Briefly describe the scientific data to be managed and shared:

- Summarize the types (e.g., 256-channel EEG data and fMRI images) and amount (e.g., from 50 research participants) of scientific data to be generated and/or used in the research. Descriptions may include the data modality (e.g., imaging, genomic, mobile, survey), level of aggregation (e.g., individual, aggregated, summarized), and/or the degree of data processing.
- Describe which scientific data from the project will be preserved and shared. NIH does not anticipate that researchers will preserve and share all scientific data generated in a study. Researchers should decide which scientific data to preserve and share based on ethical, legal, and technical factors. The plan should provide the reasoning for these decisions.
- A brief listing of 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.
- Data types expected to be shared under the Genomic Data Sharing (GDS) Policy should be described in this element. Note that the GDS Policy expects certain types of data to be shared that may not be covered by the DMS Policy's definition of "scientific data".[3]

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

3. Standards

Describe what standards, if any, 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). While many scientific fields have developed and adopted common data standards, others have not. In such cases, the DMS Plan may indicate that no consensus data standards exist for the scientific data and metadata to be generated, preserved, and shared.

[3] For more information on the data types to be shared under the GDS Policy, consult NIH's Data Submission and Release Expectations.

4. Data Preservation, Access, and Associated Timelines

Give plans and timelines for data preservation and access, including:

- The name of the repository(ies) where scientific data and metadata arising from the project will be archived.[4]
- How the scientific data will be findable and identifiable, i.e., via a persistent unique identifier or other standard indexing tools.
- When the scientific data will be made available to other users (ideally as soon as possible) and for how long. Identify any differences in timelines for different subsets of scientific data to be shared.

For human genomic data:

- Investigators are expected to submit data to a repository acceptable under the Genomic Data Sharing Policy.[5]
- Human genomic data is expected to be shared according to NIH's Data Submission and Release Expectations, but no later than the end of the performance period, whichever comes first.

For non-human genomic data:

- Investigators may submit data to any widely used repository.
- Non-human genomic data is 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.

5. Access, Distribution, or Reuse Considerations

Describe any applicable factors affecting subsequent access, distribution, or reuse of scientific data related to:

- Informed consent
- Privacy and confidentiality protections consistent with applicable federal, Tribal, state, and local laws, regulations, and policies
- Whether access to scientific data derived from humans will be controlled
- Any restrictions imposed by federal, Tribal, or state laws, regulations, or policies, or existing or anticipated agreements
- Any other considerations that may limit the extent of data sharing. Any potential limitations on subsequent data use should be communicated to the individuals or entities (for example, data repository managers) that will preserve and share the scientific data.

For human genomic data:

- Informed Consent Expectations:
    - For research involving the generation of large-scale human genomic data from cell lines or clinical specimens that were created

[4] When writing this section, remember to review the NIH Institute and Center Data Sharing Policies page for institute-specific requirements. Also, please note that CWRU is a member institution of the Inter-University Consortium for Political and Social Research (ICPSR). This repository is free for research teams from member institutions, and it is appropriate for a vast majority of studies in nursing and the social sciences. They have an application process to safeguard access to the data plus levels of access depending upon the sensitivity of the data. CWRU's ICPSR Representative is Mark Eddy (mark.eddy@case.edu; 216-368-5457).

[5] Frequently used repositories for sharing and accessing human genomic data

or collected AFTER the effective date of the GDS Policy (January 25, 2015): NIH expects that informed consent for future research use and broad data sharing will have been obtained. This expectation applies to de-identified cell lines or clinical specimens regardless of whether the data meet technical and/or legal definitions of de-identified (i.e., the research does not meet the definition of "human subjects research" under the Common Rule).
    - For research involving the generation of large-scale human genomic data from cell lines or clinical specimens that were created or collected BEFORE the effective date of the GDS Policy: There may or may not have been consent for research use and broad data sharing. NIH will accept data derived from de-identified cell lines or clinical specimens lacking consent for research use that were created or collected before the effective date of this Policy.
- Institutional Certifications and Data Sharing Limitation Expectations:
    - DMS Plans should address limitations on sharing by reviewing the criteria of the Institutional Certification, which is a form that the investigator and a CWRU signing official provide NIH during the Just-in-Time process. PIs will work with the Contracts Office to complete the Institutional Certification document after the proposal receives an awardable score from NIH.
    - In cases where it is anticipated that Institutional Certification criteria cannot be met (i.e., data cannot be shared as expected by the GDS Policy), investigators should state the Institutional Certification criteria in their DMS Plan, explaining why the element cannot be met, and indicating what data, if any, can be shared and how to enable sharing to the maximal extent possible (for example, sharing data in a summary format). In some instances, the funding NIH ICO may need to determine whether to grant an exception to the data submission expectation under the GDS Policy.
- Genomic Summary Results:
    - Investigators conducting research subject to the GDS Policy should indicate in their DMS Plan if a study should be designated as "sensitive" for the purposes of access to Genomic Summary Results (GSR).[6]

6. Oversight of Data Management and Sharing

Indicate how compliance with the DMS Plan will be monitored and managed, frequency of oversight, and by whom (e.g., titles, roles). CWRU investigators should include the following language in this section:

> The Principal Investigator will oversee the management and sharing of data during the study process. Effective January 25, 2023, Case Western Reserve University will require the Principal Investigator to certify at the times of annual progress report and final report that the NIH-approved

[6] NOT-OD-19-023

data sharing and management plan(s) has/have been followed. Case Western Reserve University's Office of Research Administration will periodically audit the NIH-approved data sharing and management plans for adherence.

Budgeting for Data Management and Sharing
Investigators may request funds toward data management and sharing by adding these costs to the budget and budget justification. Allowable costs include:

- 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.

The General Application Guide for NIH and other PHS agencies has been updated to reflect new budget guidance related to DMS plans.[7]

Assessment of the DMS Plan
Program staff at the proposed NIH Institute or Center will assess DMS Plans to ensure the required components have been addressed and to assess the adequacy/reasonableness of the plan.

Peer reviewers will not be reviewing DMS Plans if data sharing is not integral to the Funding Opportunity Announcement. However, they will be able to review a summary of the DMS Plan and a description of the requested Data Management and Sharing Costs, included in the budget justification attachment, and comment on the reasonableness of budget for data management and sharing. If data sharing were integral to the project design, then peer reviewers will review the DMS Plan attachment and may factor it into the score.

[7] PHS 398 Modular Budget Form with updated DMS guidance