Guidelines on NIH Policy for Data Management and Sharing Plan

A new NIH policy will take effect January 25th, 2023, that requires inclusion of a Data Management and Sharing (DMS) Plan. The specific details of the data management and sharing policy have been published on the NIH Scientific Data Sharing Website and presented here.

The DMS Policy applies to all research, funded or conducted in whole or in part by NIH, that results in the generation of scientific data. This includes research funded or conducted by extramural grants, contracts, Intramural Research Projects, or other funding agreements regardless of NIH funding level or funding mechanism. The DMS Policy does not apply to research and other activities that do not generate scientific data, including training, infrastructure development, and non-research activities.

I. Writing a Data Management and Sharing Plan

Under the new DMS policy, NIH expects researchers to maximize the appropriate sharing of scientific data, considering factors such as legal, ethical, or technical issues that may limit the extent of data sharing and preservation. NIH requires all applicants planning to generate scientific data to prepare a DMS plan that describes how the scientific data will be managed and shared. NIH encourages DMS plans to be consistent with the FAIR (Findable, Accessible, Interoperable, and Reusable) data principles.

Elements to Include in a Data Management and Sharing Plan

DMS plans should address the following recommended elements and should be two pages or less in length. NIH has developed an optional DMS plan format page that aligns with the recommended elements of a DMS Plan. A draft DMS plan template is available and appended here, with a final fillable format version pending.

Element 1. Data Type: Briefly describe the scientific data to be managed and shared:

A. Summarize the types (for example, 256-channel EEG data and fMRI images) and amount (for example, 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.

B. 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.

C. 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.

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. If applicable, specify how needed tools can be accessed.

Element 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).

Element 4. Data Preservation, Access, and Associated Timelines: Give plans and timelines for data preservation and access, including:

A. The name of the repository(ies) where scientific data and metadata arising from the project will be archived. See Selecting a Data Repository for information on selecting an appropriate repository.

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

C. When the scientific data will be made available to other users and for how long. Identify any differences in timelines for different subsets of scientific data to be shared.

Element 5. Access, Distribution, or Reuse Considerations: Describe any applicable:

A. 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.

B. Whether access to scientific data will be controlled.

C. Protections for privacy, rights, and confidentiality of human research participants.

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. The NIH IC will assess whether an applicant's DMS plan appropriately considers and describes these factors.

Element 6. Oversight of Data Management and Sharing: Indicate how compliance with the DMS plan will be monitored and managed.

A. It is expected that PIs will become proficient in developing DMS plans and ensure that every member of the research project team is involved in the planning, implementation, and maintenance of data management policies and procedures.

B. The PI is ultimately responsible for adherence to the DMS plan and periodically revising the plan as needed.

C. Once awarded, PIs will be required to submit their NIH approved DMS plan to RFUM's designated Research Data Management Coordinator (Dr. Carl White) and attest that they will remain in compliance with the NIH DMS Plan and University Policies.

Additional Considerations: Note that funding opportunities or ICs may have specific expectations (for example: scientific data to share, relevant standards, repository selection). View a list of NIH Institute or Center data sharing policies.

II. Budgeting for Data Management and Sharing

As outlined in the NIH Guide Notice Supplemental Policy Information: Allowable Costs for Data Management and Sharing, NIH recognizes that making data accessible and reusable for other researchers may incur costs. For that reason, investigators may request funds toward data management and sharing in the budget and budget justification sections of their applications.

The National Academies of Science has developed a resource "Forecasting Costs for Preserving, Archiving, and Promoting Access to Biomedical Data" that may be useful when budgeting for data management and sharing costs.

RELATED RESOURCES

NIH-OD-21-013 Final NIH Policy for Data Management and Sharing

NOT-OD-21-014 Supplemental Information to the NIH Policy for Data Management and Sharing: Elements of an NIH Data Management and Sharing Plan

NOT-OD-21-015 Supplemental Information to the NIH Policy for Data Management and Sharing: Allowable Costs for Data Management and Sharing

NOT-OD-21-016 Supplemental Information to the NIH Policy for Data Management and Sharing: Selecting a Repository for Data Resulting from NIH-Supported Research

NOT-OD-22-213 Supplemental Information to the NIH Policy for Data Management and Sharing: Protecting Privacy When Sharing Human Research Participant Data

Sample plan created by the NIH DMSP Guidance Working Group.

Ten simple rules for maximizing the recommendations of the NIH data management and sharing plan

FASEB statement and recommendations

AAMC resources

FAQ

How do I get assistance with data storage, curation and management? My lab generates lots of data and data types, e.g. images, DNA/RNA sequence, etc. Where am supposed to store it and manage uploading it to a repository when the time comes?

The office of the EVP for Research is currently developing strategies, policies and resources to guide and enable researchers implement their DMS plans.

How is the DMS plan reviewed at NIH?

DMS plans are assessed at the programmatic level and not by study section members. There will be an opportunity to revise a DMS plan if needed.

When should data be shared?

Note that NIH encourages scientific data to be shared as soon as possible, and no later than the time of an associated publication or end of the performance period, whichever comes first. NIH also encourages researchers to make scientific data available for as long as they anticipate it being useful for the larger research community, institutions, and/or the broader public.

What about genomic data?

To comply with the genomic data sharing policy, NIH expects that investigators and institutions develop and provide a plan for sharing genomic data as a part of their DMS Plan.

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.

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).