National Institutes of Health: NIH template

Data Management and Sharing Plan

Data Type

Guidance:

Briefly describe the scientific data to be managed, preserved, and shared, including:

- A general summary of the types and estimated amount of scientific data to be generated and/or used in the research. Describe data in general terms that address the type and amount/size of scientific data expected to be collected and used in the project (e.g., 256-channel EEG data and fMRI images from ~50 research participants). Descriptions may 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 that has occurred (i.e., how raw or processed the data will be).
- A description of 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 that may affect the extent to which scientific data are preserved and shared. Provide the rationale 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.

Related Tools, Software and/or Code

Guidance:

An indication of 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.

Standards

Guidance:

An indication of what standards 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 Plan may indicate that no consensus data standards exist for the scientific data and metadata to be generated, preserved, and shared.

Data Preservation, Access, and Associated Timelines

Guidance:

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. NIH has provided additional information to assist in selecting suitable repositories for scientific data resulting from funded research (NOT-OD-21-016).
    - 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 (i.e., the larger research community, institutions, and/or the broader public) and for how long. NIH encourages scientific data be shared as soon as possible, and no later than time of an associated publication or end of the performance period, whichever comes first. Researchers are encouraged to consider relevant requirements and expectations (e.g., data repository policies, award record retention requirements, journal policies) as guidance for the minimum time frame scientific data should be made available. NIH 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. Identify any differences in timelines for different subsets of scientific data to be shared.

Access, Distribution, or Reuse Considerations

Guidance:

NIH expects that in drafting Plans, researchers maximize the appropriate sharing of scientific data generated from NIH-funded or conducted research, consistent with privacy, security, informed consent, and proprietary issues. Describe any applicable factors affecting subsequent access, distribution, or reuse of scientific data related to:

- Informed consent (e.g., disease-specific limitations, particular communities' concerns).
    - Privacy and confidentiality protections (i.e., de-identification, Certificates of Confidentiality, and other protective measures) consistent with applicable federal, Tribal, state, and local laws, regulations, and policies.
- Whether access to scientific data derived from humans will be controlled (i.e., made available by a data repository only after approval).
    - 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).
- Any other considerations that may limit the extent of data sharing.

Oversight of Data Management and Sharing

Guidance:

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