

# National Institutes of Health: Data Management and Sharing Plan

## Data Management and Sharing Plan

### Data Type

#### Guidance:

Briefly describe the scientific data to be managed and shared:

- 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.
- 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

#### For data subject to the GDS Policy:

- Data types expected to be shared under the 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". For more information on the data types to be shared under the GDS Policy, consult [Data Submission and Release Expectations](#).

### Related Tools, Software and/or Code

#### Guidance:

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.

### Standards

#### Guidance:

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).

### Data Preservation, Access, and Associated Timelines

#### Guidance:

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. See [Selecting a Data Repository](#) for information on selecting an appropriate repository.
- 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 and for how long. Identify any differences in timelines for different subsets of scientific data to 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.

#### For data subject to the GDS Policy:

- For human genomic data:
  - Investigators are expected to submit data to a repository acceptable under the Genomic Data Sharing Policy. See [Where to Submit Genomic Data](#).
  - 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.

### Access, Distribution, or Reuse Considerations

#### Guidance:

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. The NIH ICO will assess whether an applicant's DMS plan appropriately considers and describes these factors. For more examples, see [Frequently Asked Questions](#) for examples of justifiable reasons for limiting sharing of data.

#### **Expectations for human genomic data subject to the GDS Policy:**

- Informed Consent Expectations:
  - For research involving the generation of large-scale human genomic data from cell lines or clinical specimens that were created 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 anticipating sharing according to the criteria of the [Institutional Certification](#).
  - 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), as described in [NOT-OD-19-023](#).

#### **Oversight of Data Management and Sharing**

##### *Guidance:*

Indicate how compliance with the DMS Plan will be monitored and managed.