

NIH Data Management and Sharing Plan (DMSP) Guidance

Writing the DMSP

Use the recommended, but not required, template

Format page: <https://grants.nih.gov/sites/default/files/DMS-Plan-blank-format-page.docx>.

Add DMS costs to the budget and to the budget justification

Allowable costs include:

- Curating data and developing supporting documentation
- Local data management considerations
- Preserving and sharing data through established repositories

Allowable costs are described in this supplemental: <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-015.html>.

Reviewers may not see the grant's DMSP but will be able to see data management costs in the budget justification and detailed budgets.

Specify a repository where your data will be shared

Repository options:

1. Choose a subject-specific data repository if there is one available for your data type.
2. Refer to NIH's list of supported data repositories: <https://sharing.nih.gov/data-management-and-sharing-policy/sharing-scientific-data/repositories-for-sharing-scientific-data>.
3. Choose the CaltechDATA repository (<https://data.caltech.edu/>), which provides researchers with 500GB of free storage, makes it easy to describe and upload files with DOIs, and the Library takes care of access and preservation of the data. More information under the "Sharing Large Data" heading of this handout if you need more than 500GB of storage.
4. Contact the Library for a data repository recommendation at data@caltech.edu.

Describe genomic data sharing in the DMSP instead of a separate Genomic Data Sharing Plan

NIH's Genomic Data Sharing Policy is being subsumed into the new NIH Data Management and Sharing Policy. All genomic data sharing requirements stand but plans for genomic data sharing should be detailed in the DMSP instead of a separate Genomic Data Sharing Plan.

NIH Data Management and Sharing Plan (DMSP) Guidance

Sharing Large Data (>500 GB)

It can be more difficult to find repositories that will accept large data files, as these are costly to maintain for the long term. If your project generates a large volume of data, you'll want to make sure you have a good estimate of how much data your project will generate as well as a budget to pay for long-term storage.

CaltechDATA provides researchers with 500GB of free storage, but can accommodate storing larger volumes of data. Please reach out to Caltech Library at data@caltech.edu to discuss potential storage options and costs. While Caltech Library intends for all data uploaded to CaltechDATA to be available in perpetuity, charges for large data storage are based on the length of time that we guarantee data will be available.

Caltech Library will work with your research group to determine what storage options are the best fit for your project. All storage charges are one-time charges and generally processed when the data are uploaded and the supporting grant is still active. There are a number of potential models; here are some examples:

- Primary data files are stored utilizing the research group's Caltech HPC storage allocation. Groups with large storage volumes pay a one-time \$62.50 / TB charge to cover offsite backups for 5 years.
- The group pays a one-time \$300 / TB upload fee, which covers all storage costs for 5 years.
- Files are stored on a storage allocation utilizing national resources such as ACCESS (<https://access-ci.org/>), with offsite backups paid for utilizing AWS credits from CloudBank (<https://www.cloudbank.org/>). There is often no additional charge associated with these storage allocations, but they may only be available to certain types of data or grants. These options have a separate application and approval process outside of Caltech Library.

NIH Data Management and Sharing Plan (DMSP) Guidance

Data Types that Require Special Considerations in the DMSP

Data derived from human participants

The NIH DMS Policy wants researchers to responsibly share data derived from human subjects. NIH has the following best practices for protecting participant privacy when sharing data:

- Apply appropriate de-identification (relying on the Common Rule and the HIPAA Privacy Rule)
- Establish scientific data sharing and use agreements
- Understand and communicate legal protections against disclosure and misuse

Even with de-identification, data may need to be shared in a controlled way. NIH has a supplemental on "Protecting Privacy When Sharing Human Research Participant Data" containing more information: <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-22-214.html>.

Management and sharing of American Indian/Alaska Native participant data has extra guidance, detailed in this supplemental: <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-22-214.html>.

We encourage researchers to be in contact with Caltech IRB (<https://researchcompliance.caltech.edu/committees/institutional-review-board>) about complying with the NIH DMS Policy in a way that protects participant privacy.

Data that includes, or may be the basis of, intellectual property

Please be aware that sharing your unpublished research data through a repository may constitute a public disclosure that could be considered as prior art with respect to a later-filed patent application covering an invention to which the data pertains. If you are considering seeking patent coverage for anything related to the data in question, please submit an invention disclosure through the process described here — <https://innovation.caltech.edu/patents-licensing/report-an-invention> — or confer with the Office of Technology Transfer and Corporate Partnerships (OTTCP) (<https://innovation.caltech.edu/about-ottcp/contact-us>) for further advice, prior to making your data available to others.

NIH Data Management and Sharing Plan (DMSP) Guidance

Additional Resources

Caltech-specific resources

Caltech Library webpage about the NIH Data Management and Sharing Policy:

<https://library.caltech.edu/publish/data/nih-policy>.

Caltech example NIH DMSP based on the general template for the NIH Policy for Data Management and Sharing Policy: <http://doi.org/10.7907/x2g2-v221>.

DMSP sample language: <https://library.caltech.edu/publish/data/data-management-plans#language>

Contact the Library with questions about the policy and to review DMSPs at data@caltech.edu.

NIH resources

NIH webpage on data management and sharing: <https://sharing.nih.gov/>.

NIH-supported data repositories: <https://sharing.nih.gov/data-management-and-sharing-policy/sharing-scientific-data/repositories-for-sharing-scientific-data>.

Further resources

DSMP checklist to ensure that your NIH DMSP contains all of the necessary elements:

<https://osf.io/awypt>.

The DMPTool has templates for writing DM(S)Ps for various funding agencies: <https://dmptool.org/>.