

NSF Data Management Plan – Instructions and Template

For Use by UVU Faculty

The Data Management Plan (DMP) addresses the Merit Review Criteria of the proposal by indicating how the results and products of the project will be collected and preserved (*Intellectual Merit*) and how they will be shared efficiently with others to effect a *Broader Impact*. Each DMP should be appropriate for the data, software/code, and materials being generated. Each should reflect the best practices and standards for the proposed research and types of data being generated. Generic DMPs should be avoided. While norms and expectations about the dissemination, publication, and archival deposit varies by data type and across research communities, PIs are encouraged to consider how data management will facilitate robust and reliable research and improve public access to research results. Remember that data management practices are rapidly evolving. NSF provides the following guidelines:

NSF Data Management Plan Requirements: All NSF proposals now require a supplementary document of no more than two pages labeled *Data Management Plan*. This supplementary document should describe how the proposal will conform to NSF policy on the dissemination and sharing of research results. The Data Management Plan will be reviewed as an integral part of the proposal and considered under Intellectual Merit or Broader Impacts or both, as appropriate for the scientific community of

NSF Data Sharing Policy: Investigators are expected to share with other researchers, at no more than incremental cost and within a reasonable time, the primary data, samples, physical collections and other supporting materials created or gathered in the course of work under NSF grants. Grantees are expected to encourage and facilitate such sharing.

Requirements by Directorate, OFFICE, Division, Program, or Other NSF Unit: Some Directorates, Offices, Divisions, Programs, or other NSF units, have specific guidance for the preparation of Data Management Plans. The following NSF page provides links to these documents: <https://www.nsf.gov/bfa/dias/policy/dmp.jsp> . If guidance specific to the program is not available, then the requirements established in this section apply. If a specific program solicitation provides guidance on preparation of data management plans, such guidance must be followed.

Proposal & Award Policies & Procedures Guide, pp. II-24-25, NSF 18-1

Plans for data management and sharing of the products of research. Proposals must include a document of no more than two pages uploaded under “Data Management Plan” in the Supplementary Documentation section of FastLane. This supplementary document should describe how the proposal will conform to NSF policy on the dissemination and sharing of research results (see Chapter XI.D.4), and may include:

1. The types of data, samples, physical collections, software, curriculum materials, and other materials to be produced in the course of the project;
2. The standards to be used for data and metadata format and content (where existing standards are absent or deemed inadequate, this should be documented along with any proposed solutions or remedies);
3. Policies for access and sharing including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements;
4. Policies and provisions for re-use, re-distribution, and the production of derivatives; and
5. Plans for archiving data, samples, and other research products, and for preservation of access to them.

A valid Data Management Plan may include only the statement that no detailed plan is needed, as long as the statement is accompanied by a clear justification.

The Data Management Plan will be reviewed as an integral part of the proposal, considered under Intellectual Merit or Broader Impacts or both, as appropriate for the scientific community of relevance.

This document has been abstracted from the following source:

Introduction to Data Management Plans

Dr. Andrew Stephenson, Penn State University

<https://www.e-education.psu.edu/dmpt/node/644>

What is a Data Management Plan (DMP)?

A data management plan is a document that tells how a researcher will collect, document, describe, share, and preserve the data that will be generated as part of a project.

Dr. Andrew Stephenson is Distinguished Professor of Biology and Associate Dean for Research and Graduate Education in the Eberly College of Science at Penn State. As an active researcher, he has generated and collected data for many years and served on many a panel reviewing grant proposals. From his perspective, data management plans make good sense.

Many funding agencies are now requiring that grant applicants provide information about their data management plan (DMP) as part of their grant proposal. Since 2011 the National Science Foundation (NSF) has required researchers to include DMPs with their grant proposal applications.

Why Do You Need a Data Management Plan?

Obviously, the foremost reason for needing a plan is that agencies such as the NSF, the National Institute of Health (NIH), and the National Endowment for the Humanities (NEH) are requiring DMPs. There are other reasons, however, why formulating a plan for managing research data is important.

Reason One: First, a DMP helps you plan and organize your data collection by having you think through the questions that will arise as you gather data. A DMP essentially documents key activities in the research data lifecycle, such as the collection, description, preservation, and access or discovery of data. Such documentation is crucial to reproducibility of research results. By laying out the blueprint for lifecycle management of data, a DMP provides valuable details, such as how the data will be preserved for the long term, how and where the researcher will make the data available for sharing, and whether reuse of the data, including derivatives, will be allowed.

Reason Two: Second, related to reproducibility, a DMP can help prevent or reduce the likelihood of mishaps such as data loss, data errors, and unethical uses of data. In effect, a DMP fosters improved communication and accountability for data.

Reason Three: Third, data that has been generated by a federally funded project is publicly funded data - that is, data that has been made possible by taxpayer dollars. As such, unless there are restrictions or sensitivities about the data, these are data that should be made available to the public for broad sharing and accessibility. Finally, having a DMP reflects an understanding that the collected data have intrinsic value, as illustrated in the video below. It can be another source of attribution and further investigations. Sometimes the collected data is all that remains for further investigations.

Components of a Typical Data Management Plan

A DMP basically consists of five parts, in which the following aspects of data are addressed:

Part 1: The types of data to be collected or produced during the project, and the processes or methodology for doing so:

- Types of data that will be generated by your research (e.g., human subjects related surveys, field data, samples, model output data)

- Data format(s) and file types (e.g., .txt, .pdf, .xls, .csv, .jpeg, etc.)
- How the data will be collected or accessed (if using existing data)

Part 2: The formats for the data and the standards that will be followed for documenting and describing the data:

- Information about your data you will need to save (i.e., experimental design, environmental conditions, global positioning information, etc.)
- What metadata standard you will use to document your data (i.e., some research domains have widely accepted formats, others may not and you may target how that decision may be made in the project)
- How you plan to record your metadata

Part 3: The availability of the data, including information about ways in which the data will be accessed, and whether there are any issues related to privacy and/or intellectual property:

- Expected availability of the data during the project period
- List/Explain any ethical or privacy issues incurred by the data
- Address any intellectual property rights issues (e.g., who holds the rights to these data?)
- If you are collecting sensitive data (e.g., data stemming from human-subject research), then sharing such data will likely require different types or levels of access. Are higher levels of security required? Will an embargo be needed?

Part 4: The guidelines, procedures, or policies for data reuse and/or redistribution, attribution, as well as for the creation of derivatives from the data:

- What you will permit in terms of reuse and redistribution of the data, based on policies for access and sharing
- Think about what other researchers (whether in your subject domain or others) may find your data useful
- Identify the lead person or committee on the project who will make the decisions on redistribution on a case-by-case basis
- Where the data will be deposited (e.g., data repository, repository service at your institution, etc.)

Part 5: The measures that will be taken to help ensure the long-term preservation of, and access to, the data - including possible mention of factors such as format migration and who will be responsible for managing the data for the duration of the project:

- Will all of the data produced on your project be preserved, or only some?
- Context for your data (e.g., tools, project documentation, metadata etc.) required to make it accessible and understandable
- Anticipated transformations of the data in order to deposit it and make it available
- The length of time the repository will be available to the public and/or maintained (some directorates have a suggested minimum for the time after a project ends or after publication of certain data)

Again, remember: Data management plans submitted with NSF proposals cannot be longer than two pages.

DATA MANAGEMENT PLAN **TEMPLATE***

- 1. Types of data, samples, physical collections, software, curriculum materials, and other materials to be produced in the course of the project.**

Provide a detailed description of data sets and other materials generated; specify how much data will be produced and whether it will change or be updated. If data that will be collected are sensitive in nature, note here.

- 2. The standards to be used for data and metadata format and content (where existing standards are absent or deemed inadequate, this should be documented along with any proposed solutions or remedies).**

Identify the format in which the data will be stored as well as the type(s) of metadata that will be included. These standards will often be determined by the accepted best practices of your discipline.

- 3. Policies for access and sharing including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements.**

Detail specific mechanisms that will be utilized to provide the requisite level of access. For data of a sufficiently sensitive nature, describe the means by which granular control and access will be achieved.

- 4. Policies and provisions for re-use, re-distribution, and the production of derivatives.**

Identify any potential limitations on the data requestor's ability to re-use or re-distribute the data or materials. This could be because of the type of sensitivity identified above or because the source of some of the data collected for the project has placed limitations on the re-distribution of those data and materials.

- 5. Plans for archiving data, samples, and other research products, and for preservation of/access to them.**

If your data is extensive, you may need to purchase data storage for your project and include it in the project budget.

* Remove red text when preparing your Data Management Plan. Remove footer.