CREATING AN NSF DATA MANAGEMENT AND SHARING PLAN

NSF-required data management and sharing plans are becoming more important when determining funding awards. Proposals must include a document of no more than **two pages** uploaded under "Data Management and Sharing Plan." This should describe how the proposal will conform to NSF policy on the dissemination and sharing of research results and may include (from *NSF Proposal and Award Policies and Procedures Guide,* effective May 20, 2024):

1. types of data, samples, physical collections, software, curriculum materials, and other materials to be produced in the course of the project;
2. 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 preserving access to them.

Research Data Services at UW–Madison shows how to create a management plan (https://researchdata.wisc.edu/how-to-create-a-dmp/) and provides tips for best practices. UW–Madison is a member institution with DMPTool, which is a free, open-source, online tool that helps researchers structure and draft data management plans as well as share their plans with others. DMPTool provides data management plan templates based on specific requirements of funders, but researchers should always follow the guidelines provided by program solicitation and the funder's policy documents when creating a DMP. A quick guide to using the tool is here.

Good Practices and Things to Avoid

Good practice

- Share your data upon publication.
- Share your data in an open, accessible, and machine-readable format (e.g., PDF [portable document format], .csv [comma-separated values], .docx [Microsoft Word], .txt [plain-text].
- Deposit your data in a subject or institutional repository so your colleagues can find and use it.
- Deposit your data in your institution's repository to enable long term preservation.
- License your data so people know what they can do with it.
- Tell people how to cite your data.
- When choosing a repository, ask about the support for tracking its use. Do they provide a handle or DOI? Can you see how many views and downloads? Is it indexed by Google, Google Scholar, the Data Citation Index?

Avoid

- "Data available upon request" is *not* sharing the data.
- Sharing data in PDF files.
- Sharing raw data if the publication doesn't provide sufficient detail to replicate your results.

To help you prepare a data management plan, the following model may be adapted to your proposal needs. NOTE: text in Roman is standard boilerplate material that can be modified. Text in *blue italics* offers suggestions on filling in details to satisfy NSF requirements.

Data Management Planning Tool

Part 1: General Information

Short description of the type of data collected and used during the project. Describe the process by which you will collect and/or generate this data.

Example data types: spatial, temporal, observational, experimental, simulations or models, survey responses, etc.

Collection and analysis process: how will observations be recorded (ex: audio, video, field notes)? What kind of equipment or software will you use?

Part 2: Standards and Methods of Description

Describe the file formats your research will include. List any descriptive standards, internally generated or pre-existing, that will be used. Describe the methods and best practices you will follow when describing your data.

Think about documentation necessary for others to understand your data: description of data capture methods, explanation of analysis, human readable field names, discipline-specific descriptive standards.

Is your metadata embedded in your data – such as field names in a spreadsheet or database? Or is it in an external document such as a controlled vocabulary list or an encoding key?

Part 3: Access and Security

A. The following is boilerplate text about WCER networks and storage. You may add additional information at the bottom of the boilerplate section.

All digital data collected by WCER research groups will be stored in a networked file storage system. Data stored on WCER and UW–Madison file systems are encrypted at rest and backed up regularly to a secondary location. Access to any research data located on the WCER network is controlled through use of active directory permissions. Requests for access to specific data items must be approved by the principal investigator or her/his designate. Access approval will be consistent with the project's approved research protocol and any applicable federal and State of Wisconsin laws. The networked file storage system is protected by a firewall, which blocks access outside of WCER and is managed by the Research IT staff. If required, researchers may access data remotely through VPN or through a secure terminal server that is encrypted with SSL technology. All communications are encrypted with SSL and access to specific data items, once approved by the principal investigator or her/his designate, is overseen by the WCER Technical Services staff.

B. Describe conditions under which your research data will or will not be shared. Address methods for facilitating sharing where appropriate.

Note any restrictions on sharing based on privacy or security requirements. If raw data cannot be shared or retained, will cleaned files be created and made available?

What research protocol and applicable laws will govern access to the data? Who will be responsible for approving requests to gain access to the data?

WCER requires individual projects to redact sensitive information or aggregate data to an appropriate level so that individual research subject identities are protected.

Part 4: Sharing and Distribution

NSF is less inclined to fund projects that have limited dissemination plans (i.e., only papers and conferences) and a passive data sharing plan (e.g., language such as "data available upon request" is a red flag for NSF that your data sharing plan is not active). Thus, NSF data sharing and dissemination must be strong for your proposal to be competitive.

Describe how you anticipate your data being reused in the future, either by others within your research group or by scholars at large

WCER hosts a web server for individual projects that can serve as a means to disseminate project information in addition to publications in journals.

There is a movement for publicly funded data to be of wide benefit. How might your data be used beyond your project? Is such use appropriate or common practice within your discipline?

Explain how your choices of file format, technologies, and access policies will facilitate reuse where appropriate.

Data Sharing

1. Note in your proposal that WCER Tech Services will create a project website at no cost. This website will help in publicly sharing project findings and data.

2. Explore and note tools that will help you share data with other users, such as:

- Inter-University Consortium for Political and Social Research (ICPSR) for data sharing and data analysis. The site helps you manage your research data and prepare and archive it. The site has a data prep guide and data deposit page. If you have questions, the UW–Madison official rep is Aaron Crandall, director of Social Science Research Services, 4402 Social Sciences Building, 1180 Observatory Drive. Email aaron.crandall@wisc.edu; phone: 608-262-1131.
- NCSES Survey Data can be manipulated. Researchers might be allowed to deposit data from their NSF projects.

Dissemination

In addition to your plan for papers and conference presentations, consider plans for reaching other stakeholders, such as parents, teachers, administrators, policymakers. Language on tools available that might aid dissemination (if appropriate to your project) include:

WCER has an ongoing agreement with the Education Resources Information Center (ERIC) to contribute digital versions of its working papers to the ERIC database. This allows educators, researchers, and the public free access to work produced by WCER researchers. Our project reports can be easily tailored for inclusion in the WCER working paper series, thus widening their audience. All final papers produced under this grant will be submitted to ERIC.

Part 5: Preservation and Storage

Describe what data will be kept in long term storage after the project is completed.

WCER has the ability to maintain data after project conclusion to allow continued dissemination to the public. Archival storage for the long-term preservation of research data is available.

You may wish to preserve all, none, or a selection of your data for the long term. How long do you need to preserve raw data versus derived data? What criteria will be used to determine what is kept or discarded at the end of your funding period?

Do you need to plan for migration to new file formats in the future (ex: technologies or storage devices that may become obsolete)?