Supporting Data Management Plans: Guides for Grant Writers and Reviewers

David Fearon, Betsy Gunia, Tim DiLauro   Johns Hopkins University Sheridan Libraries   datamanagement@jhu.edu http://dmp.data.jhu.edu

The U.S.'s National Science Foundation & other funders require Data Management Plans (DMPs), but they provide limited guidelines for proposal writers *AND* proposal reviewers. JHU DMS developed two guides for DMP creation, evaluation & consulting

Grant proposal writer needs…

- help with organizing and formatting the DMP document
- Beyond the DMP template: what *expresses* my project?
- What extra (competitive) efforts at sharing & preservation are feasible?

Shared DMP questions about the proposed project's lifecycle

- What are the proposed project's data management requirements?
- Who will re-use shared data?
- What data products & metadata??
- Which data will be preserved?
- Using a data archive?
- How will data be shared?
- How to manage data during the project?

Grant proposal reviewer needs…

- a checklist: what to look for
- Quick comparisons among proposals
- What sets a DMP apart from others?

Questionnaire for Help Preparing a Data Management Plan

How to use this document
This questionnaire distills NSF's guidelines for what to address in your data management plan. Note that some questions may not apply to your study…

Data Products and Standards

Research Outputs

	Data Product	Data Source	Format(s)	Estimated Amount
1				
2				
3				

Data and Metadata Standards   *(Sample selection of questions)*

1. Do the listed data products use standards for formats or metadata, and why are you using them? If not, will your project develop and maintain standardized formatting and metadata?
2. What details (metadata) are necessary for others to use your data?
3. How will metadata be generated (automatically, manually, or both)?

Data Storing and Long-Term Preservation 3

Storage during project

8. What digital and non-digital data will be retained during the project?
9. How and where will data be backed up and who is responsible for it? 5

Data Sharing

19. Of the data products generated during the project, which data will be shared? Which data will be publically-accessible? ⑦
20. Who is expected to use the shared data? *(Continues to page 2).*

The questionnaire format encourages researchers to add project details in their own words, rather than copying template statements.

Both DMP guides share these features…

A table allows listing of research products, highlighting each item's relevance for data management, and can be ordered by the research workflow.

We created section headings for both tools help unify diverse guidelines among NSF directorates

On MS Word, pop-up text shows guidelines and sample DMP text:

Sharing can include direct release to interested parties upon request. Accessible generally means unmediated distribution of data through an online resource or database. Ctrl+Click to follow link

"All data types will be made publically accessible after de-identification, and will use a unique durable identifier…" Ctrl+Click to follow link

Both include guidelines, tips and examples of DMP content on subsequent pages.

Reviewer's Worksheet for NSF Data Management Plans

	Research product	Source	Format	Size	Preserved (how?)	Shared (how?)
	E.g., Tables, images, computer code, curriculum items, physical samples	Data repository, Instrument, interviews, PI's prior project	JPG, MATLAB, Excel table, proprietary	>1TB, 20K files	Discarded, PI retains, data archive	By request, website, repository
1	processed fMRI	MRI scanner	.rec	50GB	PI server	JHU Data Archive
2	survey data	Qualtrics	.xlsx	N=25	PI	By req
3	filter algorithm	code	Python		JHU DA	NITRC

Data Sharing   *i* *eg.*

- ☒ Is data publically accessible?
- ☐ When will data be shared?
- ☐ Who administers?
- ☐ ☆ **Describes audience to benefit.**

Preparation of data for sharing: *i* *eg.*
- ☒ Uses research field's metadata standards
- ☐ AND/OR creates description sufficient for re-use
- ☐ Metadata or supplementary files explaining: Content, file structure, procedures, codebook
- ☒ ☆ **Metadata associated with digital files**
- ☐ ☆ **Will convert files to non-proprietary formats**

Data sharing policy: *i* *eg.*
- ☐ Gives conditions for re-use
- Accounts for:
- ☐ ☆ **privacy (personal identifiers)/security issues**
- ☐ ☆ **intellectual property (copyrights, patents)**

- ☐ If a plan states there is no data to manage or share, have they justified it?   *i*

Data management during project:
- ☒ Storage: has a backup plan   *i, eg.*
- ☐ Location & media used:
- ☐ ☆ **2+ copies with 1 off-site**
- ☐ ☆ Specifies who is responsible
- ☒ ☆ Data security/ access controls   *i eg*
- ☐ ☆ Has conventions for naming & organizing files   *i eg*
- ☐ ☆ Version control
- ☐ ☆ Collaboration coordination

Data retention after the project: *i*
- ☒ Where is data preserved?
- ☐ How long?
- ☐ Who administers?
- ☐ ☆ Gives reasons for preserving data
- ☒ ☆ **Using an archive or repository?**

Services of archive (if specified for preservation and/or sharing data)   *i*

Archive type	Preservation activities	Data sharing services
☒ PI's Institutional repository	☐ Data integrity checking	☒ Public access to data files
☐ ☆ **Data repository**	☐ ☆ **Format & media migration**	☒ Persistent data citation

---

Reviewers can use the worksheet as a printout or in Microsoft Word, either by entering information or as a visual reference when comparing proposals.

☆ Stars indicate 'extra credit' details of more thorough plans

A section on archiving highlights the projects' extra efforts at data sharing

---

Using our DMP resources for consulting:

JHU DMS Data Management Consultants directly consult with researchers who are preparing DMPs, using the questionnaire to guide the conversation and highlight best practices of data management that researchers infrequently consider during the research process.

Consultants use a "feedback" version of the questionnaire to share their notes from DMP consultation meetings with researchers, offering suggestions customized to their research project.

The Reviewer's DMP Worksheet is also a useful topic checklist for consultation discussions with researchers about data management for DMP preparation, and their research projects more generally.