What is a Data Management Plan?

The Data Management Plan is a document that formulizes how the research project, data will be collected, generated, analyzed, shared and stored and includes a timeline of how the work will be conducted. It should answer the 5Ws- what, how, why, where and when.

Phases of data life cycle

There are several key phases in the data lifecycle – these include: creation, storage, use, sharing, archiving and destruction. Each of these should be addressed in the DMP you provide to the sponsor if one is required at the time of proposal submission.

Why do I need to know?

Over the past several years large funding entities such as NIH and NSF have been updating policies and procedures to ensure that data and samples collected as part of federally funded studies are shared within a reasonable period of time. The idea is that sharing scientific data in a timely manner, will accelerate research discovery and enhance rigor and reproducibility. As a result you will now need to complete a DMP.

FAIR Principles- are guidelines for scientific data management that were issued in 2016, in the Journal Scientific Data; to improve the Findability, Accessibility, Interoperability and Reuse of digital information by humans and machines. These are now becoming adopted as standards by many sponsors.

Effective January 25, 2023 - NIH and NSF, NASA, IMLS, DOD, DOJ amongst others will require a full Data Safety Monitoring Plan: to be provided at the time of proposal submission.

Helpful Resources

The University Library has several resources to help you develop a strong DMP. It also provides a link to the DMPTool; a great application, with directions, templates and examples. You can log onto this website with your Mason ID and password. THE DMP Tool has a help function that will walk you through the steps.

Other helpful links include:

» NIH Data Management and Sharing Basics

*For most up-to-date information, please visit the sites listed under “Helpful Resources”
What to include in a DMP
The plan should include the following elements:

- **Type of Data to be shared**—*Note: researchers are not required to preserve and share all scientific data.*
  The plan should explain what data will be preserved and what will not and why. Data should describe the type of data e.g. observational, experimental, simulated, compiled and from how many subjects or data points it was collected.

- **Data Format**
  Numeric, text, video, and/or audio.

- **File format**
  Excel, PDF, video, images, JPEG or PIC files.

- **Size of total data and how it will be stored**
  Describe how data will be stored—for example, repository or hard drives—and include the plan to back-up the data.

- **For relevant data associated documentation should also be preserved**
  For example, include description of how study protocols and data collection instruments will be preserved.

- **Define method of data collection**
  Define methods of data collection and define what data will be preserved.

- **Timeline of data collection**
  Describe.

- **Data management**
  Describe procedures for managing data. Include how and where data will be stored, who has access, any security provisions, and timelines for keeping data.

- **Roles and responsibilities of project staff with data management and retention**
  Describe the assignment of project staff for data management: define each role and their access rights. In the plan include contingency for if PI leaves the institution.

- **Schedule for sharing data**
  Describe how data will be shared and with whom, include planned publication dates.

- **Format of final data**
  Describe what data will be shared, in what format. For example: will it be aggregate data and how will it be shared?

- **Describe methods of sharing data**
  Describe how data will be shared and for how long, and if there will be any restrictions, such as HIPPA, FERPA, state law or informed consent restrictions. Disclose any data that will be proprietary and not be shared. Will data sharing agreements be needed?

- **Describe how documentation will be prepared to let other researchers know how information will be collected and shared.**
  Procedures should define all steps of the data collection so it should be possible to replicate the research of the original team. DMP- should also include the purpose of data collection.

- **Data storage and security**
  Describe in detail the process of mitigating risk including restricted data access.
  
  If a data repository is used define the process of accessing data and mitigating risk. Write out a plan of how the data repository will be used and data access process. Also include an outline of the timeline for data preservation. If there are restrictions to data access—for example informed consent, privacy or confidentiality regulations—please include that in the description.

- **Costs for data sharing and data storage can be included in the proposal budget for some sponsors.**
  **Budgeting for Data Management and Sharing**—Effective January 2023 costs for data management and sharing can be included in the budget. This link will take you to what NIH considers as reasonable and allocable costs. Other sponsors may provide similar guidance; you will need to check their website for updated information.

- **Oversight of Data Management and sharing**
  Include how plan compliance will be monitored.

- **Data Destruction and Preservation**
  Outline how data will be preserved and the plan for safe destruction where appropriate.