LAS 6292: In-class work - DMP

updated: 2021-03-05

Introduction

Today's in class we will start drafting the DMP assigned as part of your semester project. To do so we will use the DMP tool, to which we have access as UF Researchers.

1. Log on to the DMP

- The DMP Tool, Quick Start Guide, and Help

2. Chose the type of DMP you would liek to prepare.

3. Start filling in Boxes. It's that easy.

4. The hard work is thinking through **all** of the data you will be collecting during your thesis, how you will collect it, how much of it there will be, what format it will be in, how you will store it…It's best to start by making a list off-line rather than expecting to remember it all at this moment, so…

5. …I have created txt versions of the questions for the NEH DMP, NSF DMP, and a Generic DMP. You can download these (right-click, save as a .txt file to your computer) and work on them off-line.

Some sample DMPs

1. DataOneDMP samples for NSF

- NSF General: Mauna Loa example
- NSF General: Rio Grande example
- NSF General: HDF Map example
- NSF General: Nutrient Network example
- NSF BIO: E. afffinis example

2. Sample DMPs on the 'DMP Tool' website Public DMPs**(note: "Public DMPs are plans created using the dmptool service and shared publicly by their owners. They are not vetted for quality, completeness, or adherence to funder guidelines.")**

3. National Endowment for the Humanities: Data Management Plans From Successful Grant Applications (2015-2018) and (2011 - 2014) can be found as downloadable zip files on this page

4. A *really* nice DMP for a Digital Humanities Project by researchers at the University of Maryland, College Park.

Additional resources to help prepare your DMP

1. Qualitative Data have some unique issues regarding their description and storage, so the Qualitative Data Archive has put together a great checklist of things to consider.

2. Short Checklist & Questions to Consider when putting together a DMP

3. Very in-depth Checklist & Questions to Consider when putting together a DMP

4. UF Library Guidelines for preparing a DMP

5. Another tools for creating DMPs is the DMP Online

Reminder: What is a DMP?

- Formal document that lays out your plan for managing the data you will collect (or have collected).
- Outlines what you will do with your data during and after you complete your research
- Ensures your data is safe for the present and the future
- Video overview by RWTH Aachen University (12 min).

Why do one?

- Save time
- Less reorganization later
- Increase research efficiency
- Makes it easier to archive data downt he road (some data centers and repositories have specific requirements for data documentation, and knowing these requirements in advance saves time and effort that might be spent trying to correct data after they are collected).
- Increasingly: you don't have a choice. Many funding agencies (e.g., NSF, NIH, NEH) now require a DMP for proposal submission.

What are the components of a DMP?

1. Information about data & data format
2. Metadata content and format
3. Policies for access, sharing and re-use
4. Long-term storage and data management
5. Roles and responsibilities
6. Budget

Some funders or institutions may require specific elements in a data management plan; you should check in advance to make sure you meet all requirements

1. Information About Data & Data Format

1.1 Description of data to be produced

- Experimental
- Observational
- Raw or derived
- Physical collections
- Models and their outputs
- Simulation outputs
- Curriculum materials
- Software
- Images
- Other

1.2 How data will be acquired?

- When? Where? Method?

1.3 How data will be processed

- Software used

- Algorithms

- Workflows

- Data transformations/formats needed (Consider archive policies)

- *This step is important to consider before the project since it may affect how data are organized, what formats are used, and how much should be budgeted for hardware and software. Things to consider are what software may be used, what algorithms will be employed, how these fit into the overall workflow of the project.*

1.4 File formats

- Justification
- Naming conventions

1.5 Quality assurance & control during sample collection, analysis, and processing

1.6 Existing data

- If existing data are used, what are their origins?
- Will your data be combined with existing data?

- What is the relationship between your data and existing data?

1.7 How data will be managed in short-term

- Version control
- Backing up
- Security & protection
- Who will be responsible

2. Metadata Content & Format

- Metadata defined:
- Documentation and reporting of data
- Contextual details: Critical information about the dataset
- Information important for using the data
- Descriptions of temporal and spatial details, instruments, parameters, units, files, etc.

2.1 What metadata are needed

- Details that make data meaningful

2.2 How metadata will be created and/or captured

- Lab notebooks? GPS units? Auto-saved on instrument?

2.3 What format will be used for the metadata

- Standards for community
- Justification for format chosen

3. Policies for Access, Sharing, Reuse

3.1 Obligations for sharing

- Funding agency
- Institution
- Other organization
- Legal

3.2 Details of data sharing

- How long?
- When?
- How access can be gained?
- Data collector rights

3.2 Ethical/privacy issues with data sharing

3.3 Intellectual property & copyright issues

- Who owns the copyright? (? Institutions and/or funding agencies often have a policy for intellectual property and copyright. There may be other considerations such as embargos on data due to patents, politics, or journal requirements.)
- Institutional policies
- Funding agency policies
- Embargos for political/commercial reasons

3.4 Intended future uses/users for data

- This helps determine the most appropriate data center to archive your data.

3.5 Citation

- How should data be cited when used?
- Do you need a persistent citation (e.g. DOI)?

4. Long-term Storage & Data Management

4.1 What data will be preserved

4.2 Where will it be archived

- Most appropriate archive for data
- Community standards

4.4 Who will be responsible

- Contact person for archive

5. Roles and responsibilities

5.1 Outline the roles and responsibilities for implementing this data management plan.

- Who will be responsible for data management and for monitoring the data management plan? How will adherence to this data management plan be checked or demonstrated? What process is in place for transferring responsibility for the data? Who will have responsibility over time for decisions about the data once the original personnel are no longer available?

6. Budget

6.1 Anticipated costs

- Time for data preparation & documentation
- Hardware/software for data preparation & documentation
- Personnel
- Archive costs

6.2 How costs will be paid

NSF DMP Requirements

From Grant Proposal Guidelines: "Plans for data management and sharing of the products of research. Proposals must include a supplementary document of no more than two pages labeled "Data Management Plan". This supplement should describe how the proposal will conform to NSF policy on the dissemination and sharing of research results, and may include: the types of data, samples, physical collections, software, curriculum materials, and other materials to be produced in the course of the project 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) policies for access and sharing including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements policies and provisions for re-use, re-distribution, and the production of derivatives plans for archiving data, samples, and other research products, and for preservation of access to them."

REMINDER

The key components of a DMP:

- Information about data & data format
- Metadata content and format
- Policies for access, sharing and re-use
- Long-term storage and data management
- Budget