Writing a Data Management Plan

Tom Melvin, Anita Schwartz, and Leigh Botner
University of Delaware

November 10, 2016
What Will We Cover?

- Research Guide on Data Management
- Locating information on funding agency requirements
- Locating sample data management plans
- Directories for metadata schemas and data repositories
- DMP Tool
- Data Management Budgeting
- Data attribution, ethical and legal issues
Staff Directory

The Research Office is available to assist UD researchers every step of the way in the development of proposals, to the patenting and reporting of results. The Research Office includes two major units: Research Management and Operations, which supports proposal processing and contract-and-grant administration; and Intellectual Property, Technology Transfer and Compliance, which oversees research integrity and compliance, patents, copyrights, trademarks, and technology licensing. UD Research Communications, based in Communications and Public Affairs, also works closely with the office, helping to translate and promote UD research to the public.
Morris Library Research Guide for Research Data Management

http://guides.lib.udel.edu/c.php?g=371489
Well organized, documented, preserved and shared data advance scientific inquiry and increase opportunities for learning and innovation.
Government Mandates

October 1, 2003: NIH applications over $500,000 required a DMP or state why it was not possible.

January 18, 2011: NSF required all grant applications to submit a two page DMP

December 2011: Research Works Act (RWA)

February 2012: Federal Research Public Access Act (FRPA)

May 2012: “Require free access over the internet to scientific journal articles arising from taxpayer funded research” petition

February 2013: OSTP Memorandum

May 2013: White House Executive Order 13642

February 2015: HHS releases plans to comply with 2013 memorandum
Why is Managing Data Important?

- Stay organized throughout the research process
- Comply with requirements (journal or funder)
- Ability to validate and reproduce research
- Ability to disseminate for other’s use
- Conform with legal and ethical responsibilities
Legal and Ethical Issues

Intellectual Property Issues

a. Who owns the data?
b. How can the data be used?
c. Collected data vs. system used for storage.
d. Does a database have copyright protection?
1. Public Domain Dedication and License

dedicates the database and it’s content to the public domain, free for everyone to use as they see fit.

2. Attribution License: Users are free to use the database and its content in new and different ways provided they provide attribution to the source of the data and/or database.

3. Open Database License: stipulates that any subsequent use of the database must provide attribution, an unrestricted version of the new product must always be accessible, and any new products made using OPDBL must be distributed using the same terms. Most restrictive.
Creative Commons Licenses

1. **CCO License**: When an owner wishes to waive the copyright and/or database rights, they can use the CCO mark. It effectively places the database and data into the public domain. (Equivalent to a ODC PDDL license).

2. **Public Domain Mark**: Used to mark works that are in the public domain and for which there are no known copyright or database restrictions. It is possible to flag factual data as PDM in a database in order to make it clear that it is free to use.
First Steps in Managing Your Data

1. Figure out your criteria for keeping data

2. Consider the metadata you want to collect to document your datasets - (your data can only be cited if you provide suitable metadata giving details of the dataset and where it can be found.)

3. Think about where you want to store your data
Metadata Directories

Society of American Archivists

http://www2.archivists.org/groups/metadata-and-digital-object-roundtable/metadata-directory

DataCite (re3data.org)

https://schema.datacite.org/
Where Should You Keep Your Data?

Storing your data during the project

versus

Storing your data long-term
(preservation and sharing)
Data Storage and Backup Services at UD

- Check with your college IT professional for departmental storage and backup solutions.
  - Custom (Research) solutions: Central IT can work with you and/or your college IT professional to offer fee-based custom storage solutions

- Understand your requirements for your data
  - Size specifications or limitations
  - Encryption for sensitive data
  - Access type (mapped drive, web access, etc.)
  - Speed of drives for read/write
  - Licensing or grant restrictions
Central IT Storage Services at UD

- **Win Domain File Services**
  
  The Win Domain File Services (Win DFS) is a subscription-based, native Windows storage solution offered by UD IT for University departments to locally manage University information.

- **Google Apps@UD**

  Google Drive A cloud-based storage solution offered free for current faculty, students, and staff that syncs files between your computer and the web and allows for real-time online editing.

  Careful consideration should be taken before storing data on any non-UD server as described on Cloud Services at the University.
Central IT Backup Services at UD

IT Backup Services provides

- Data and email backup and recovery services for centrally managed systems and servers.
- Fee based backup and recovery services for UD College- and department-owned servers only (not available for desktop systems).

CrashPlan Personal (students and employees)
Long Term Data Storage

Preserve and share your research data in order to:

- Further science as a whole
- Further your research/reputation
- Enable new discoveries with your data
- Comply with funder or publisher data sharing requirements

Will your data be accessible and usable when it is needed?
Data Repositories

Data repositories may or may not include preservation

Domain repositories
Some funders require you to deposit data in a specialist archive where other researchers in your field can easily access and cite your data.

Institutional repositories (e.g., UDSpace)

http://udspace.udel.edu
How to Choose the Right Repository

Who can access the data?

Is it Open Data?

“Open data is data that can be freely used, reused and redistributed by anyone - subject only, at most, to the requirement to attribute and sharealike.”

- Open Data Handbook, Open Knowledge
Repository Directories

Repositories for Open Data
http://oad.simmons.edu/oadwiki/Data_repositories

DataCite (re3data.org)
https://www.datacite.org/services/find-repository.html
Why Have a Data Management Plan?

1. Meet funding agency requirements and gives you an edge in the grant selection process.
2. Increases your citation count.
3. Ensure research integrity and replication.
4. Facilitates the sharing of data.
5. Helps manage the data.
6. Allows you to budget for data retention.
Roles and Responsibilities

How will the responsibilities for managing the data be delegated?
Who will monitor the data management plan?
Who will have responsibility over time for the decisions about the data?

Example:
The primary responsibility for managing the data will belong to the PI, Lizabeth Schlemer. If the PI is not available, both Linda Vanasuppa and Aaron Estrada will be able to access the data and continue to manage it. It is our goal to house the data in the Digital Commons at Cal Poly. The data will be uploaded at the completion of the grant and then the responsibility for the information will be transferred to the Library. Any data not deemed public will be kept on the computer belonging to the PI, which is backed up periodically.
Expected Data

Types of data, samples, physical collections, software, curriculum materials.

Describe the expected types of data to be retained.

Describe file types, dataset size, number of expected files or sets, content.

Example:

The data for this project will include 60 interviews with engineering faculty and a set of survey data. The interviews will be stored with both audio files and transcribed text files. The surveys will be stored in excel spread sheets and include survey answers and demographic information. Transcribed text will be verified to the audio files. No fewer than two copies of the data will be kept at all times and will be located in different locations. In addition, data for public access will be anonymized and stripped of identifying information. This process of anonymizing will be documented.
Period of Data Retention

Describe the period of data retention.

Exceptions requiring longer retention periods.

Estimate how long your data will be kept after the completion of research.

Example:
The data will be available for a 10 year period after completion of the grant.
Describe specific data formats, media, including any metadata.

Example:
There will be three file types: audio, text and spreadsheet. For each of these we will include an ID for the individual interviewed along with the date, location, file size and the level of anonymization that occurred (for public files). We will create a naming system for the files that will indicate the meta data. For files uploaded to the Digital Commons will include a complete descriptor with the file.
Data Dissemination and Policies for Public Access, Sharing, and Publication Delays

Articulate how sharing of primary data is to be implemented.

Policies for public access and provisions for protection of privacy.

How will data be shared and managed with stakeholders.

Clearly state publication delay policies.

Example:

It is the goal to house the data in the Digital Commons at Cal Poly. The data will be uploaded at the completion of the grant. The individual’s privacy will be insured by removing any identifiable information. The data will be publicly available through Cal Poly’s Digital Commons and through our research website. The access to the data will comply with the IRB requirements. Although we will publish findings based on the data, it is hoped that the data will be a rich source of inquiry for anyone interested in educational research. We only request attribution in the use of the data.
Data Storage and Preservation of Access

Describe physical and cyber resources and facilities used for retention.

Long term strategy for storing, archiving, and preserving data.

What archives/repository/database will the data be deposited.

What procedures do they have for preservation and backup.

Example:
The data will be housed long term on Cal Poly's Digital Commons with all the guarantees afforded data curation on our campus. It is hoped the data will be kept for 10 years after completion of the research.
Could the budget and its justification specifically address the costs of implementing the Data Management Plan?

Yes. As long as the costs are:

- In accordance with applicable Cost Principles,
- Necessary to implement the Data Management Plan,
- Included (typically on Line G2) on the proposal budget, and
- Justified in the budget justification.
Cost Principles

Reasonable and necessary
- Prudent person test

Allowable under regulations or award provisions
- Reasonable & Necessary
- Allocable
  - Within Terms & Conditions of Award

Allocable and verifiable
- Benefits the project
- Easily identified and assigned
  - Proportional to the relative benefit received by project

Consistently treated
- Consistency in estimating, charging and reporting of direct costs and F&A
Publication/Documentation/Dissemination

Line G2 on the NSF Proposal Budget

The proposal budget may request funds for the costs of documenting, preparing, publishing or otherwise making available to others the findings and products of the work conducted under the grant.

This generally includes the following types of activities: reports, reprints, page charges or other journal costs (except costs for prior or early publication); necessary illustrations; cleanup, documentation, storage and indexing of data and databases; development, documentation and debugging of software; and storage, preservation, documentation, indexing, etc., of physical specimens, collections or fabricated items.

Work with your departmental administrator and RO contract and grant specialist to correctly budget for these costs.
DMP Tool

Watch a 90-second animated video providing a brief overview of DMPTool, its basic features and capabilities, and why it's a valuable application for researchers from all disciplines.
Data Management Planning Tool
Create, review, and share data management plans that meet institutional and funder requirements.

DMPTool NEWS
Latest information about data management and the DMPTool.
- New DOE “Generic” template
- New template: NIH Genomic Data Sharing
- New templates for IMLS
- Two New Templates: NIFA and NOAA
- NSF and NIH Update for the DMPTool
Data Management Planning Tool
Create, review, and share data management plans that meet institutional and funder requirements.

Click either to get started.
Log in through your institution

Select your institution below and you will be directed to your institutional log in page.

Select your institution

If you do not see your institution in the list, please select "Not in List" and click Next.
INSTITUTION LOG IN

Log in through your institution

Select your institution below and you will be directed to your institutional log in page.

Select your institution

- University of California, San Francisco
- University of California, Santa Barbara
- University of California, Santa Cruz
- University of Central Florida
- University of Chicago
- University of Cincinnati Main Campus
- University of Colorado Anschutz Medical Campus
- University of Colorado Boulder
- University of Connecticut
- University of Delaware
- University of Florida
- University of Georgia
- University of Hawaii
- University of Helsinki, Finland
- University of Houston
- University of Idaho
- University of Illinois at Chicago
- University of Illinois at Urbana-Champaign
- University of Iowa

California Digital Library

Privacy Policy | Accessibility Policy | Terms of Use | Contact Us | About
Log in through your institution

Select your institution below and you will be directed to your institutional log in page.

University of Delaware

Click Next >>

If you do not see your institution in the list, please select "Not in List" and click Next.
Please complete filling in your profile information.

**MY PROFILE**

**Personal Information**
- **Username**: trains@udel.edu
- **First name**: 
- **Last name**: 

**Contact Information**
- **Email**: trains@UDel.Edu
- **Institution**: University of Delaware

**ORCID**
- ORCID: [http://orcid.org/xxxx-xxxx-xxxx-xxxx-xxxx](http://orcid.org/xxxx-xxxx-xxxx-xxxx-xxxx)

**API Token**
- **My Token**: 03a20f1607044a2e0002e4e098d2100
- **Generate New Token**
- **Remove Token**

[Save] [Cancel]
ORCID (orcid.org)

Register to get your unique ORCID identifier

ORCID provides a persistent digital identifier that distinguishes you from every other researcher and, through integration in key research workflows such as manuscript and grant submission, supports automated linkages between you and your professional activities ensuring that your work is recognized.

The goal is to prevent research ambiguity and distinguish one botanist names James A. Cook from another.
Please complete filling in your profile information.

MY PROFILE

Personal Information  Notification Preferences

Check or uncheck boxes for email notifications.

Select / deselect all

Users

- A new user role is granted

DMP owners and co-owners:

- A new comment has been added to my DMP
- A submitted DMP is approved, rejected or finished informal review
- My DMP's visibility has changed
- I have been made a co-owner of a DMP
- A DMP is completed

Save  Cancel
1. **My Dashboard** is your best starting point for managing your DMPs. From here you can get an overview of all your DMPs.

2. **My DMPs** lists of all your DMPs, noting which you own, co-own, and their status.

3. **Create New DMP** allows you to start from a template or you can copy an existing DMP.

4. **My Profile** is where you can update your personal information, including your ORCID. You can also modify your notification preferences.
### DMP Requirements

Use the A-Z links below to narrow down the list by institution or use the search box to search for specific DMP Templates. Sample plans are not necessarily from the funder.

<table>
<thead>
<tr>
<th>Template</th>
<th>Funder</th>
<th>Funder Links</th>
<th>Sample Plans (if available)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCO-DMO NSF OCE: Biological and Chemical Oceanography</td>
<td>National Science Foundation</td>
<td>NSF OCE Sample and Data Policy; May 2011 (PDF)</td>
<td>NEH-ODH Sample</td>
</tr>
<tr>
<td>NIH-GDS: Genomic Data Sharing</td>
<td>National Institutes of Health</td>
<td>Guidance</td>
<td></td>
</tr>
</tbody>
</table>
Create New DMP:

University of Delaware

CREATE NEW DMP

Start with a DMP Template

To create a new DMP, select a funder or institutional template.

Select Template >>

Copy an Existing DMP

The existing DMPs in this list are either publicly shared by any user, shared within your institution by other DMP creators, or are plans that you have previously created.

Select an existing DMP below and add text to the template.

<table>
<thead>
<tr>
<th>Name</th>
<th>Owner</th>
</tr>
</thead>
<tbody>
<tr>
<td>USGS CDR/ECV DMP</td>
<td>Call Jenkerson</td>
</tr>
<tr>
<td>A unified approach to preserving cultural software objects and their development histories</td>
<td>DMP curator</td>
</tr>
<tr>
<td>Atmospheric CO2 Concentrations, Mauna Loa Observatory, Hawaii, 2011-2013</td>
<td>DMP curator</td>
</tr>
<tr>
<td>Daymet Follow-On: Surface Weather Data with Uncertainty Quantification for Terrestrial Ecosystem Process Models</td>
<td>DMP curator</td>
</tr>
</tbody>
</table>
SELECT DMP TEMPLATE

Select one of the funded DMP Templates listed to proceed to the next step. The type of template chosen can affect what information you will need to provide in the following steps.

- Alfred P. Sloan Foundation
- Department of Energy (DOE)
- DMP Template from DCC
- GoMRI Research Consortia DMP Template 2015
- Gordon and Betty Moore Foundation
- Institute of Education Sciences (US Dept of Education)
- Institute of Museum and Library Services
- Joint Fire Science Program
- National Institutes of Health
- National Science Foundation
- NEH-ODH: Office of Digital Humanities
- NOAA Data Sharing Template
- U.S. Geological Survey DMP Guidance
SELECT DMP TEMPLATE

Select one of the funder DMP Templates listed to proceed to the next step. The type of template chosen can affect what information you will need to provide in the following steps.

A - F  G - L  M - S  T - Z  All

- Alfred P. Sloan Foundation
- Department of Energy (DOE)
- DMP Template from DCC
- GoMRI Research Consortia DMP Template 2015
- Gordon and Betty Moore Foundation
- Institute of Education Sciences (US Dept of Education)
- Institute of Museum and Library Services
- Joint Fire Science Program
- National Institutes of Health
- National Science Foundation
- NSF-AGS: Atmospheric and Geospace Sciences
- NSF-AST: Astronomical Sciences
- NSF-BIO: Biological Sciences (2013- )
- NSF-CHE: Chemistry Division
- NSF-CISE: Computer and Information Science and Engineering
- NSF-DMSE: Materials Research
- NSF-EAR: Earth Sciences
- NSF-EHR: Education and Human Resources
- NSF-ENG: Engineering
- NSF-GEN: Generic
- NSF-SBE: Social, Behavioral, Economic Sciences
- NEH-ODH: Office of Digital Humanities
- NOAA Data Sharing Template

Search
Plan was successfully created.

NSF-GEN: Generic

Click on a section below to edit it at any time

- = Complete
* = Mandatory

Template Outline

- Types of data produced
- Data and metadata standards
- Policies for access and sharing
- Policies for re-use, redistribution
- Plans for archiving & preservation

Instructions

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

Guidance

Provide a description of the data you will collect or re-use, including the file types, dataset size, number of expected files or sets, and content. Data types could include text, spreadsheets, images, 3D models, software, audio files, video files, reports, surveys, patient records, etc. Data that underpin the findings reported in a journal article or conference paper should be...
Plan was successfully created.

NSF-SBE: Social, Behavioral, Economic Sciences

Click on a section below to edit it at any time

- Complete
- Mandatory

Template Outline

- Roles and responsibilities
- Expected data
- Period of data retention
- Data format and dissemination
- Data storage and preservation of access
- Additional possible data management requirements

Instructions

The Data Management Plan should outline the rights and obligations of all parties as to their roles and responsibilities in the management and retention of research data. It must also consider changes to roles and responsibilities that will occur should a principal investigator or co-PI leave the institution.

Guidance

(box size: small | medium | full)

Outline the staffing/organizational roles and responsibilities for implementing this data management plan.
1. **DMP Preview** button opens a new window and displays the contents of your entire DMP. You can click this at anytime to preview your plan. Be sure to click “Save Response” before clicking DMP Preview.

2. **DMP Preview** displays your entire DMP, including section headers and your responses.

3. **Export or Print** your plan as PDF or DOCX format, or print your plan.

4. While in DMP Preview, you must click “Done” to change your DMP status to ‘Completed’. If your organization has set up a “Review” process, you will instead see a “Submit for Review” button. Note: you can always come back and change or add information by clicking on DMP Overview or DMP Details in the navigation.
Ask Us!

Tom Melvin <tmel@udel.edu>
UD Library

Anita Schwartz <anita@udel.edu>
UD IT Client Support and Services

Leigh Botner <lbotner@udel.edu>
UD Research Office

http://www.udel.edu/research/about/directory.html

Some workshop content was repurposed from the MIT Libraries Research Data Management 101 workshop which is shared under the Creative Commons License.