

# **Data Management Planning**

## **DataONE Community Engagement & Outreach Working Group**

# Lesson Topics

- What is a data management plan (DMP)?
- Why prepare a DMP?
- Components of a DMP
- Recommendations for DMP content
- Example of an NSF DMP



# Learning Objectives

- After completing this lesson, the participant will be able to:
  - Define a DMP
  - Understand the importance of preparing a DMP
  - Identify the key components of a DMP
  - Recognize the DMP elements required for an NSF proposal



*CC image by cybrarian77 on Flickr*

# The Data Life Cycle

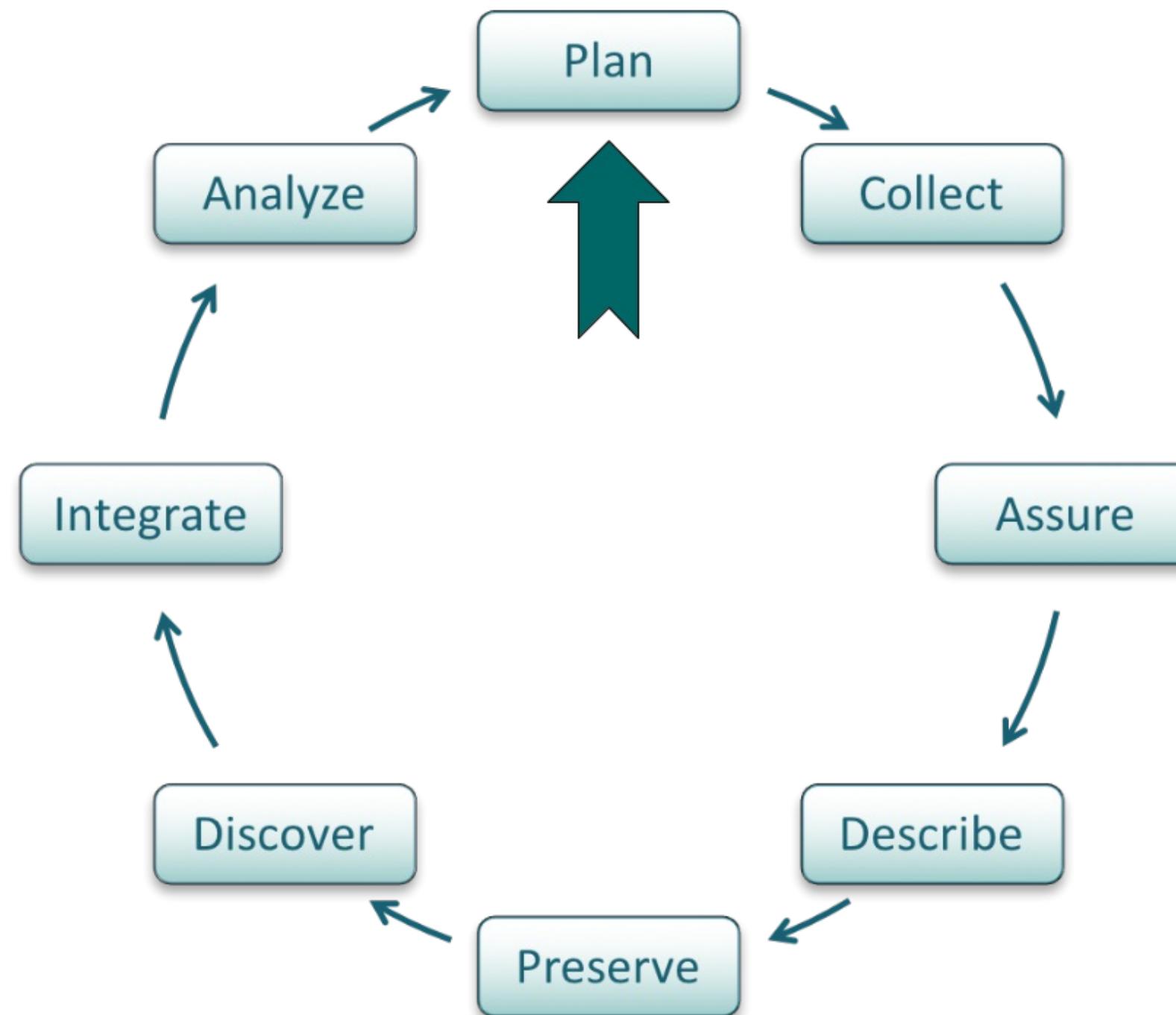


Photo Attribution: <https://www.dataone.org/data-life-cycle>

# What is a Data Management Plan?

- Formal document
- Outlines what you will do with the data **during** and **after** you complete your research
- Ensures the data is safe for the **present** and the **future**

-- *University of Virginia Library*



# Why Prepare a DMP?

Saves time

- Less reorganization later

Increase research efficiency

- Ensures you and others will be able to understand and use data in future

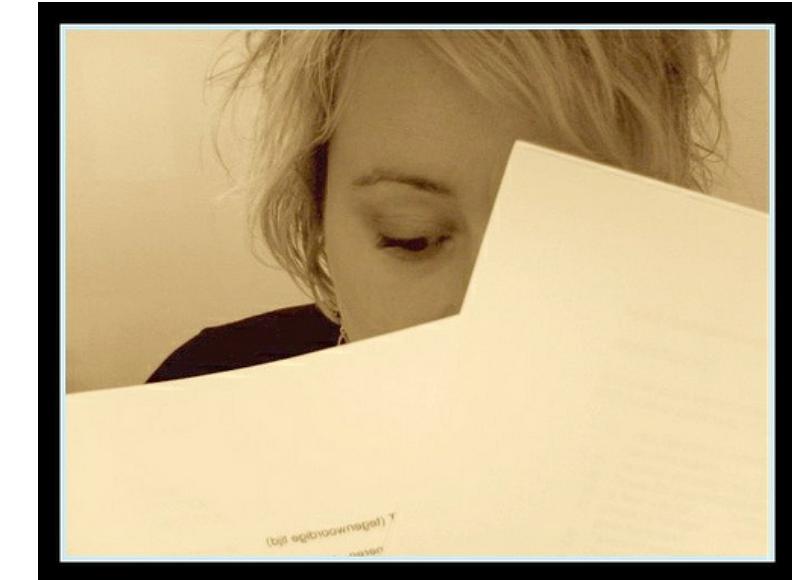


Photo Attribution: CC image Cathdew on Flickrby

# Why Prepare a DMP?, continued

- Easier to preserve data
- Prevents duplication of effort
- Can lead to new, unanticipated discoveries
- Increases visibility of research
- Makes research and data more relevant
- Funding agency requirement



*Photo Attribution: CC0 image from The Noun Project*

# Components of a General DMP

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



*Photo Attribution: CC0 image from The Noun Project*

# Information About Data & Data Format

Description of data to be produced

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



*Photo Attribution: CC image by Jeffery Beall on Flickr*

# Information About Data & Data Format, continued

How data will be acquired

- When?
- Where?



*Photo Attribution: CC image by Ryan Sandridge on Flickr*

How data will be processed

- Software used
- Algorithms
- Workflows

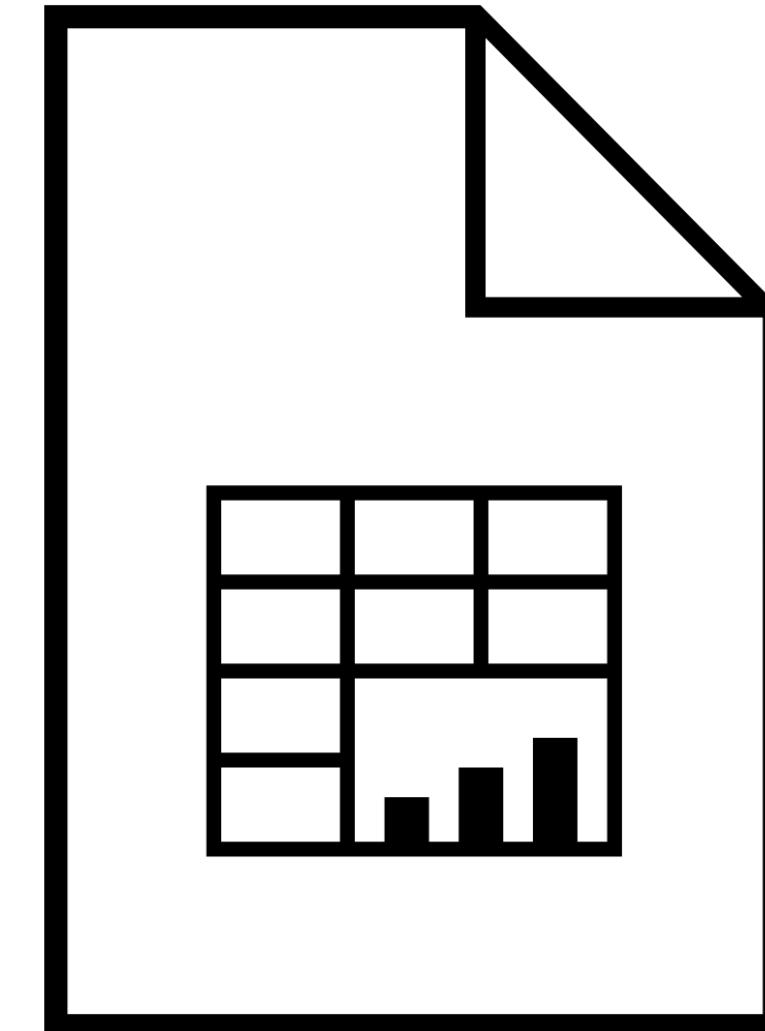
# Information About Data & Data Format, continued

File formats

- Justification
- Naming conventions

Quality assurance & control during

- sample collection, analysis, and
- processing



*Photo Attribution: CC0 image from The Noun Project*

# Information About Data & Data Format, continued

Existing data

- If existing data are used, what are their origins?
- Will the data be combined with existing data?
- What is the relationship between new data and existing data?

How data will be managed in short-term

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

# 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.



*Photo Attribution: CC0 image from The Noun Project*

# Metadata Content & Format, continued

What metadata are needed

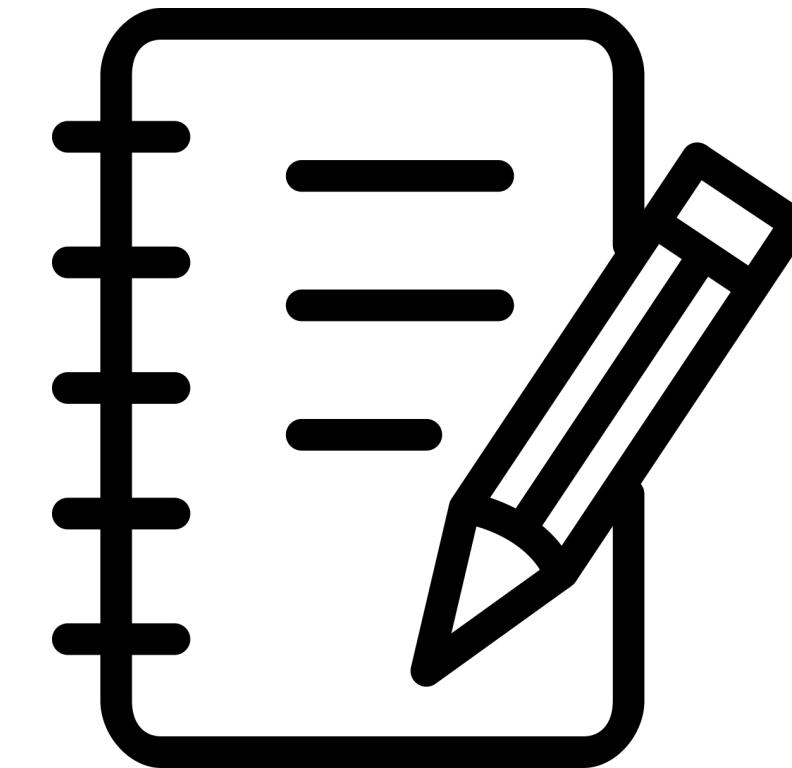
- Any details that make data meaningful

How metadata will be created and/or captured

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

What format will be used for the metadata

- Standards for community
- Justification for format chosen



*Photo Attribution: CC0 image from The Noun Project*

# Policies for Access, Sharing, Reuse

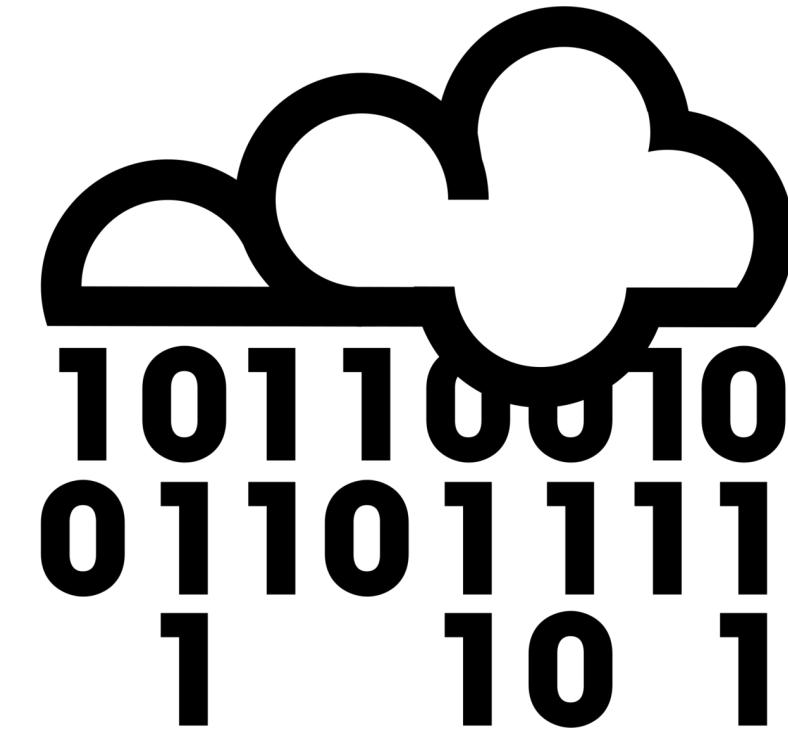
## Obligations for sharing

- Funding agency
- Institution
- Other organization
- Legal

## Details of data sharing

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

## Ethical/privacy issues with data sharing



*Photo Attribution: CC0 image from The Noun Project*

# Policies for Access, Sharing, Reuse, continued

Intellectual property & copyright issues

- Who owns the copyright?
- Institutional policies
- Funding agency policies
- Embargos for political/commercial reasons

Intended future uses/users for data

Citation

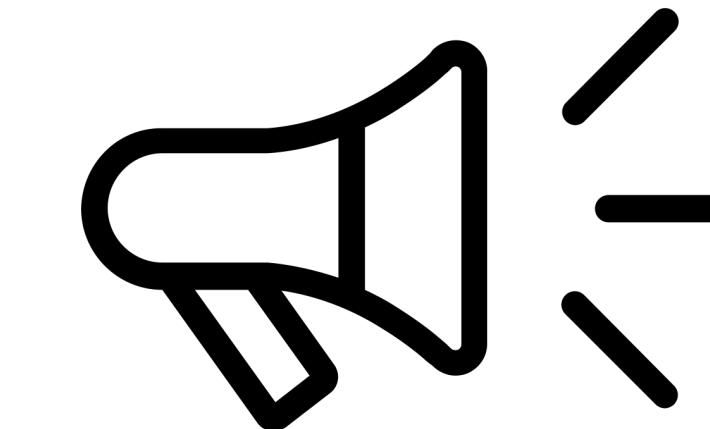


Photo Attribution: CC0 image from The Noun Project

- How should data be cited when used?
- Persistent citation?

# Long-term Storage & Data Management

- What data will be preserved
- Where will it be archived
  - Most appropriate archive for data
  - Community standards
- Data transformations/formats needed
  - Consider archive policies
- Who will be responsible
  - Contact person for archive

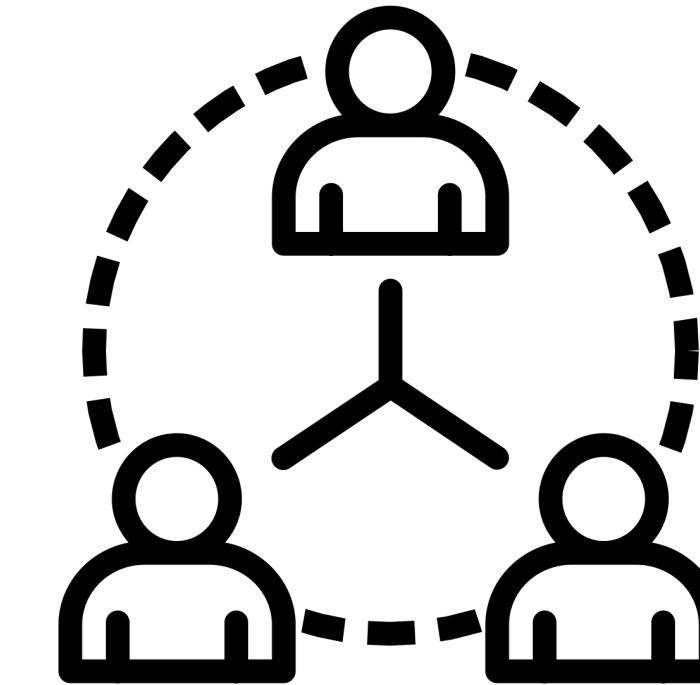


*Photo Attribution: CC0 image by Samuel Zeller on Unsplash*

# Roles and responsibilities

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

- For example:
  - 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?



*Photo Attribution: Chameleon Design on the NounProject*

# Budget

## Anticipated costs

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



## How costs will be paid

*Photo Attribution: CC0 image from The Noun Project*

# Tools for Creating Data Management Plans

## DMPTool

**PUBLIC DMPs**  
List of sample data management plans provided by DMPTool users.

- » UNDERSTANDING THE ROLE OF PHYSICIAN INTEGRATION WITHIN NURSING HOMES IN POST-ACUTE CARE OUTCOMES
- » A Political Ecology of Value: A Cohort-Based Ethnography of the Environmental Turn in Nicaraguan Urban Social Policy
- » A unified approach to preserving cultural software objects and their development histories

**DMPTOOL NEWS**  
Latest information about data management and the DMPTool.

- » New template: NIJ (DOJ)
- » NASA template update & bug fix
- » Review workflow enhancements
- » New templates: DOT and NASA
- » Getting our ducks in a row

**DMPTOOL HELP**  
Overview of how to use the tool, plus resources and guidance on data management.

- » Frequently Asked Questions
- » Create a DMP
- » Administer the DMPTool
- » Data management guidance
- » Community resources

## DMPOnline

**View plans** **Create plan** **About** **Future plans** **Help**

### My plans

The table below lists the plans that you have created, and any that have been shared with you by others.  
These can be edited, shared, exported or deleted at anytime.

Name	Owner	Shared?	Last edited	Select an action
My plan (Horizon 2020 DMP)	Me	No	13-06-2016	Edit Share Export Delete

**Create plan**

Contact us | Terms of use | DMPonline previous version  
© 2004 - 2016 Digital Curation Centre (DCC)

**DCC** **Jisc**

# NSF DMP Requirements

*From Grant Proposal Guidelines:*

- 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

[http://www.nsf.gov/pubs/policydocs/pappguide/nsf16001/gpg\\_2.jsp#dmp](http://www.nsf.gov/pubs/policydocs/pappguide/nsf16001/gpg_2.jsp#dmp)

# NSF DMP Requirements, continued

***Summarized from the Award & Administration Guide: Dissemination and Sharing of Research Results***

- Promptly publish with appropriate authorship
- Share data, samples, physical collections, and supporting materials with others, within a reasonable timeframe
- Share software and inventions
- Investigators can keep their legal rights over their intellectual property, but they still have to make their results, data, and collections available to others
- Policies will be implemented via
  - Proposal review
  - Award negotiations and conditions
  - Support/incentives

[http://www.nsf.gov/pubs/policydocs/pappguide/nsf16001/aag\\_6.jsp#VID4](http://www.nsf.gov/pubs/policydocs/pappguide/nsf16001/aag_6.jsp#VID4)

# Summary

DMPs are an important part of the data life cycle. They save time and effort in the long run, and ensure that data are relevant and useful for others. Most funding agencies (all federal\*) now require DMPs

Major components of a DMP:

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

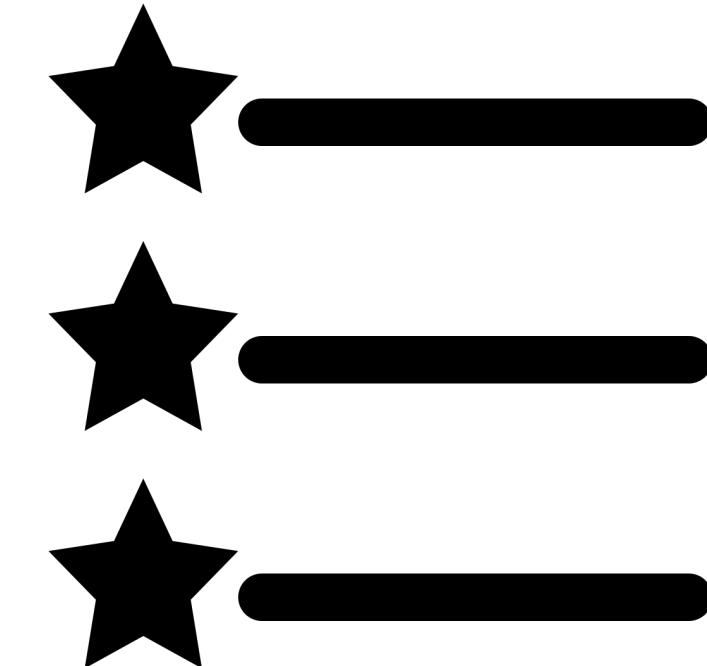


Photo Attribution: CC0 image from The Noun Project

# Resources

- University of Virginia Library. Accessed June 24, 2015 at <http://data.library.virginia.edu/data-management/plan/>.
- Digital Curation Centre Accessed June 26, 2015 at <http://www.dcc.ac.uk/resources/data-management-plans>.
- Oregon State University Library. Accessed June 27, 2015 at <http://guides.library.oregonstate.edu/dmp/policies>.
- NSF Grant Proposal Guidelines. Accessed June 26, 2015 at [http://www.nsf.gov/pubs/policydocs/pappguide/nsf16001/gpg\\_2.jsp#dmp](http://www.nsf.gov/pubs/policydocs/pappguide/nsf16001/gpg_2.jsp#dmp).
- Inter-University Consortium for Political and Social Research. Accessed June 26, 2015 at <http://www.icpsr.umich.edu/icpsrweb/ICPSR/dmp/index.jsp>.
- DataONE. Accessed June 26, 2015 at <https://www.dataone.org/data-management-planning>.

# About

Participate in our GitHub repo: [https://dataoneorg.github.io/dataone\\_lessons/](https://dataoneorg.github.io/dataone_lessons/)

**Suggested citation:** DataONE Education Module: Data Management. DataONE. Retrieved November 12, 2016. From XXXX

**Copyright license information:** No rights reserved; you may enhance and reuse for your own purposes. We do ask that you provide appropriate citation and attribution to DataONE.





