

Data Management 101:

**Tips and Tools for Successful Data
Management Planning**

Introductions

Andrew Johnson

Research Data & Metadata Librarian

andrew.m.johnson@colorado.edu

- Data management consultations
- Teaching data management
- Campus-level research data services:

<http://data.colorado.edu>

data-help@colorado.edu

Introductions

You?

- Name?
- Status (e.g., grad student, faculty)?
- Department/institute?
- Most common data type/format?

More Questions!

- Has anyone created a data management plan?
- What strategies do you use for:
 - Organizing data?
 - File naming?
 - Storing and backing up data?
 - Metadata/documentation?
 - Sharing data?
 - Archiving data?

Data Management Plans

- Required for all NSF proposals since 2011
- February 22, 2013 White House OSTP Memo:
 - DMPs likely to be required for proposals to all major funding agencies
- A good idea regardless of funding agency requirements

General DMP Requirements

1. Types of data produced:

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

2. Data and metadata standards:

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)

3. Policies for access and sharing:

Policies for access and sharing; Provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements.

4. Policies for re-use, redistribution:

Policies and provisions for re-use, re-distribution, and the production of derivatives.

5. Plans for archiving and preservation:

Plans for archiving data, samples, and other research products, and for preservation of access to them.

DMPTool

- Tool for creating custom DMPs to meet funding agency requirements/guidelines
- Available at: <https://dmp.cdlib.org/>
- No need to register to use:
 - Click “Get Started”
 - Select “University of Colorado Boulder”
 - Log in with IdentiKey

PetaLibrary Storage

- Available to any US-based researcher affiliated with the University of Colorado Boulder
- Modest fees based on cost of disk or tape and overhead
- Enables researchers to store large and growing data sets
- More information: <https://www.rc.colorado.edu/services/storage/petalibrary>

Quick Tips for Data Management

- Consistency in organization, file naming, metadata (have a plan and stick to it)
- Keep two onsite copies and one offsite/cloud copy of data
- Document everything for yourself, your collaborators, researchers in your discipline, researchers outside your discipline
- Provide access to as much as possible:
 - **Databib** (repository directory): <http://databib.org/>
 - **figshare** (free data sharing tool): <http://figshare.com>

Thank You!

andrew.m.johnson@colorado.edu