Enhancing DMPTool
Further Streamlining the Data Management Planning Process

Stephen Abrams
California Digital Library

Andrew Sallans
Center for Open Science
Data curation

“Policies and procedures for maintaining and adding value to trusted data resources for use by contemporary and future scholars”

— Adapted from DCC, www.dcc.ac.uk/digital-curation/what-digital-curation

- Value accretes to data over time through (re)use
- Integrity lies at the heart of the scholarly process
- Scholarship progresses through stepwise advances
Data curation

- Fundamental for the transparency, trustworthiness, and advancement of scholarship
  - Review, replication, and verification of results
  - Build on prior work; leverage prior investment
  - Facilitate collaboration and novel (re)use
  - Increase visibility, citations, impact, and credit
  - Avoid unnecessary duplication of effort
Data management planning

- Increasingly necessary as a prerequisite for

  ▶ Public or private grant funding

  “The Office of Science and Technology Policy (OSTP) hereby directs each Federal agency ... to develop a plan to support increased public access to the results of research funded by the Federal Government”

  ▶ Publication

  “A condition of publication in a Nature journal is that authors are required to make materials, data and associated protocols promptly available to readers without undue qualifications”

- Increasingly necessary to conform to

  ▶ Disciplinary best practices
  ▶ Institutional policies
The data management dilemma

- Many scholars feel ill-equipped to meet DMP requirements
  - What are the specific obligations I have to meet?
  - What are good data management best practices I should follow?
  - How can I find out about disciplinary and institutional resources I can use?
  - Where can I get help?

- The solution?
  - Online step-by-step wizard for creating submission-ready DMPs
DMPTool

DMPTool
Guidance and Resources for your Data Management Plan

Create ready-to-use data management plans for specific funding agencies.

The DMP Tool allows you to: 1 2 3 4

Recent DMP News
DMPTool 2 Feature Highlight: Co-Ownership & Administration
Call for Example Data Management Plans
What are your burning questions about the DMPTool?
More news>

DMPTOOL is a service of the University of California Digital Curation Center of the California Digital Library
Copyright © 2010-2013 The Regents of the University of California
Privacy Policy | Terms of Use | Photo Credits

dmptool.org
Customized instruction and help

- Guidance and best practice advice
- Links to useful disciplinary and institutional resources
Funder templates

- Support for 21 public and private funding agencies

- Gordon and Betty Moore Foundation
- Gulf of Mexico Research Initiative
- Institute of Education Sciences
- Institute of Museum and Library Services
- Joint Fire Science Program
- NEH Office of Digital Humanities
- National Institutes of Health
- National Oceanic and Atmospheric Administration
- NSF Astronomical Sciences
- NSF Atmospheric and Geospace Sciences
- NSF Biological Sciences
- NSF Chemistry Division
- NSF Computer and Information Science and Engineering
- NSF Earth Sciences
- NSF Education and Human Resources
- NSF Emerging Frontiers in Research and Innovation
- NSF Engineering
- NSF Materials Research
- NSF Physics
- NSF Social, Behavioral, Economic Science
Community supported

Created in 2011 through a collaborative effort

- California Digital Library
- DataONE
- Digital Curation Centre
- Smithsonian Institution
- University of California, Los Angeles
- University of California, San Diego
- University of Illinois, Urbana-Champaign
- University of Virginia
Adoption and use

- Over 100 institutions have provided customized guidance, or integrated local Shibboleth IdPs for SSO with institutional credentials.
Adoption and use

Since 2011, over 7,000 users from 1,400 institutions have completed 4,400 DMPs
There's always room for improvement

Good, but could be better, so ...
Everything the DMPTool does now, plus ...

- Plan co-ownership
- Self-service administrative functions
  - Create and edit DMP templates
  - Customize guidance and resource links
  - Maintain individual and institutional profiles
- Refactored UI
- Optional plan review for lifecycle management
- And more!

Formalized governance and advisory structure

Additional project collaborator: Purdue University

Funded by the Sloan Foundation and IMLS
How we got here

- Desire
  - *Expand growth, streamline operation, achieve sustainability*

- Listen
  - *Feedback from focus groups and the broader user community*

- Plan
  - *Detailed use cases for funding agencies and institutional policies*
  - *Wireframes / UI designs*

- Act
  - *Multi-institutional partnership*

- React
  - *Prototype evaluation by advisory boards and beta testers*
Governance

- Steering group
  - Initial project collaborators
  - Governing principles

- Administrative advisory board
  - CDL, Duke, Georgia Tech, Penn State, Smithsonian, Tennessee, UC Merced, UIUC, USGS, Virginia, Wisconsin, Woods Hole

- Researcher advisory board
  - CDL, George Mason, ICPSR, NCEAS, NEON, NYU, Oak Ridge, Smithsonian, UC San Francisco, UIUC, USGS, Virginia
Expanded user roles

- Dashboard and privileges personalized to role
  - Owner
  - Co-owner
  - DMP template editor
  - Resources editor
  - Institutional reviewer
  - Institutional administrator
  - DMPTool administrator
Create a DMP

- Start from a blank template
- Copy an existing plan
Select a DMP template

21 templates, including NEH, NIH, and most NSF directorates and divisions
Provide DMP details

- Easy access to customized
  - Instructions
  - Suggested answers
  - Institutional resources
Add co-owners

- Co-owners can
  - Edit
  - Preview
  - Download
Set visibility

- Owners/co-owners
- Institutional affiliates
- Public
Preview plan

Data Management Plan

E. affinis treatment conditions

1. Data Collected, Formats and Standards

Every two days, we will subsample E. affinis populations growing at our treatment conditions. We will use a microscope to identify the stage and sex of the subsampled individuals. We will document the information first in a laboratory notebook, then copy the data into an Excel spreadsheet. For quality control, values will be entered separately by two different people to ensure accuracy. The Excel spreadsheet will be saved as a comma-separated value (csv) file daily and backed up to a server. After all data are collected, the Excel spreadsheet will be saved as a csv file and imported into the program R for statistical analysis. Strauss will be responsible for all data management during and after data collection.
Create DMP template

- Start fresh
- Copy an existing template
Fill-in template requirements

Requirements can be arranged hierarchically
Review DMPs

- Optional or mandatory
- Comments can be passed back and forth between Owners, Co-owners, Reviewers
- Important for administrative oversight
Benefits to scholars

- Better plans, more efficiently
  - Login with institutional credentials
  - Access to up-to-date requirements, guidance, and resources
  - Comprehensive dashboard
  - Plan co-ownership better reflects intra- and inter-institutional collaboration
  - Copy existing plans to leverage prior statements of intention
  - Submit for mandatory, or request optional, review
  - Control visibility
  - Self-service profile and preferences maintenance
Benefits to institutions

- Better plans, better support, better oversight
  - Integration with local IdP for campus SSO
  - Institutional hierarchies for appropriate administrative delegation
  - Self-service customization
    - Institutional branding and profile
    - Plan templates
    - Help, suggested responses, resource links
  - Mandatory or optional plan review
  - API for statistical reports
Engage scholars at the point of greatest need

- Consolidate and leverage existing data management information and services
Next steps

- **Beta evaluation**  
  December 2013

- **Administrator pre-release**  
  January 2014

- **Public release**  
  February 2014

  - All current user accounts, plans, and institutional customizations will be migrated forward

  - Version 1 will continue to be available for a limited time to allow scholars to complete in-process plans

- Start planning for the next version ...  

  - Post-award project management
  - Integration with SHARE and similar initiatives
  - Other features? Please provide your ideas and feedback!
Get involved

- Have a voice, lend a hand
  - **Collaborate**
    - Contribute to maintenance and enhancement efforts
  - **Integrate**
    - Configure institutional IdP
  - **Customize**
    - Institutional branding, guidance, resource links, DMP templates
  - **Advocate**
    - Take advantage of promotional materials
      - Customizable poster/postcard
      - Taking points for administrators, IT groups, students

[bitbucket.org/dmptool/main/wiki/Materials]
Summary

“Sound, reproducible scholarship rests upon a foundation of robust accessible data. For this to be so in practice as well as theory, data must be accorded due importance in the practice of scholarship and in the enduring scholarly record.”

Data management planning is

- Increasingly a prerequisite for funding and publication
- Increasingly a disciplinary and institutional best practice
- The best way to leverage research investment and exploit research outputs
- Good for scholars and good for scholarship
- Streamlined by use of the DMPTool

www.force11.org/datacitation