DATA MANAGEMENT
101

ASERL Summertime Summit
pre-conference webinar
August 1, 2013

Jen Doty, Emory University
Wendy Mann, George Mason University
Outline

• What is Data Management?
• Motivation for providing services/support
• Service models: challenges for resources and staffing
• Preparing/training subject librarians
• Methods to raise awareness on your campus
• Institutional partnerships
• Assessments and analyses
What is Data?

Wide variety across domains:

• Physical and life sciences—data are gathered or produced by researchers, such as by observations, experiments, or models.

• Social sciences—researchers may gather or produce their own data, or they may obtain data from other sources such as public records of economic activity.

• Humanities—data most often are drawn from records of human culture, whether archival materials, published documents, or artifacts.

What is Data Management?

“Data management covers all aspects of handling, organising, documenting and enhancing research data, and enabling their sustainability and sharing.” (UK Data Archive)
What is Data Management?

“In the context of research and scholarship, Data Management refers to the storage, access and preservation of data produced from a given investigation.” (Texas A&M guide)
What is Data Curation?

“Data curation is critically important for a research institution because it provides two vital services: (1) data are not merely stored, but are preserved in order to overcome technical obsolescence inherent in any storage system; and (2) data are documented in such a way that they can be referenced in, and linked from, scientific publications and meet the requirements of funding agencies.” (UCSD – Research Cyberinfrastructure, Data Curation)
Motivation

• Funding agency requirements for data management plans (DMP) and data sharing considerations

• Library as repository for scholarly output
What is a Data Management Plan (DMP)?

- Outlines what the researcher will do with the data **during** and **after** the proposed research project, to ensure it is preserved and useful both now and in the future.
Why Prepare a DMP?

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

  - **NSF directorates and NEH Office of Digital Humanities:** 1-2 page data management plan
  - **NIH** (grants over $500K): 1 paragraph data sharing plan
Service models

Large research institution examples:

• Purdue, Cornell, UNC—coordinated and distributed services, working with research offices, central IT, data archives

• JHU, UCSD—including charge-back cost models for users of data management and curation services
Service models

Medium and smaller levels of service:

- Library team approaches—Georgia State, Trinity University, James Madison

- Dedicated RDM position—Emory, Georgia Tech
Preparing/training librarians

In-house training models:

• UVa - “re-tooling subject librarians” through brown-bag discussions and data interview participation

• Rutgers – internal course for librarians

• CU Boulder – one-day workshop for subject librarians
Preparing/training librarians

External training options:

• ICPSR summer course – “Curating and Managing Research Data for Re-Use”

• online LIS credit courses (Wisconsin, SJSU)

• emerging data curation certification programs (UNC, Illinois, UNT)
Resources and Links (1)

Background and guidance:
• DataONE Best Practices Primer - http://www.dataone.org/best-practices
• UK Data Archive - http://data-archive.ac.uk/create-manage
• ICPSR guide - http://www.icpsr.umich.edu/files/deposit/dataprep.pdf

Tutorials and Training Guides:
• DataONE education modules - http://www.dataone.org/education
• UKDA training resources - http://data-archive.ac.uk/create-manage/training-resources
• MANTRA online training for PhD students - http://datalib.edina.ac.uk/mantra/
• University of Minnesota data management course for engineering graduate students - http://z.umn.edu/datamgmt
Resources and Links (2)

Service models:

Large
- Cornell University, RDM Service Group - https://confluence.cornell.edu/display/rdmsgweb/About
- Johns Hopkins University, Data Management Services - http://dmp.data.jhu.edu/
- Purdue University, Distributed Data Curation Center - http://d2c2.lib.purdue.edu/
- UCSD, Research Cyberinfrastructure - http://rci.ucsd.edu/data-curation/
- University of Virginia, Data Management Consulting Group - http://dmconsult.library.virginia.edu/

Medium & Small
- Georgia State - http://research.library.gsu.edu/datamgmt
- James Madison - http://guides.lib.jmu.edu/data (see also article - DOI10.1080/19322909.2012.729394)
- Trinity - http://libguides.trinity.edu/DataManage (see also poster at http://digitalcommons.trinity.edu/lib_faculty/27/)
Resources and Links (3)

Training models:

In-house
- UVa - http://www.slideshare.net/shlake/re-tooling-for-data-managementsupport

External
- ICPSR summer course in data curation - http://www.icpsr.umich.edu/icpsrweb/sumprog/courses/0149
- online LIS credit courses
  - Wisconsin - http://www.slis.wisc.edu/continueed-DataMgmt.htm
  - SJSU -
- emerging data curation certification programs
  - Illinois - http://www.lis.illinois.edu/academics/programs/specializations/data_curation
  - UNT - http://icamp.unt.edu/icamp/content/project-abstract
Part Two: Communication and Assessment

- Outreach and Raising Awareness
- Institutional Partnerships
- Assessment and Analysis
Outreach: Identifying Contacts

Who is involved in managing research?

- Office of Research, Vice President for Research
- Sponsored Programs
- Office of Research Integrity, Institutional Review Board (IRB)
- Information technology, research computing
- Others

Existing expertise: who is already involved in managing research data?
Outreach: Toolkits & Resources

Association of Research Libraries (ARL)

• E-Research. Outreach.
  http://www.arl.org/focus-areas/e-research

• Data Access, Management and Sharing
  http://www.arl.org/focus-areas/e-research/data-access-management-and-sharing

Association of College and Research Libraries (ACRL) report on research data services.

EDUCAUSE report on developing a research data management plan service.
Outreach: Instruction

Holding information sessions or workshops advertises your services to the wider university community.

Information sessions on DMPs, DMPTool, repository services, etc.

Workshops on data management
Outreach: Instruction Examples

Cornell DMP Info Sessions


UNC at Chapel Hill Data Management Workshops

- Odum Institute + UNC Libraries
  http://www.irss.unc.edu/odum/contentSubpage.jsp?nodeid=667
  http://guides.lib.unc.edu/content.php?pid=294213&sid=2874509

University of Virginia Libraries

http://dmconsult.library.virginia.edu/training-sessions/
Outreach: Instruction Examples

University of Oregon

- LIB 407/507: Data Management, 1-credit course, spring term 2013 offered at the University of Oregon.  
  [http://library.uoregon.edu/node/3393](http://library.uoregon.edu/node/3393)

Outreach: Institutional Buy-in

Once you identify stakeholders and start your outreach efforts buy-in should follow.

In addition to coordinating with relevant offices, knowledge of research policies at your institution is essential.

Buy-in fosters collaboration.
Institutional Partnerships

Collaboration between the library and relevant offices is important for offering research data management support.

Are there other partners?

- Selected academic departments, research centers or data archives.
- “Low hanging fruit” or natural partners
  - Cornell’s Research Data Management Services Group = Library + CISER + Advanced Computing
Institutional Partnerships

University Wide Partnership
• Research Data Stewardship at UNC
  [URL](http://sils.unc.edu/sites/default/files/general/research/UNC_Research_Data_Stewardship_Report.pdf)

Regional Partnership
Assessment and Analysis

When to assess and how

Consider an environmental scan part of the initial assessment process.

• Where have you identified gaps or partners?
• Anecdotal evidence

Formal assessment

• Survey instruments and interviews
Formal Assessment: Examples

Research Data Assessment

- Georgia Tech
  [http://www.library.gatech.edu/research-data/data-assessment](http://www.library.gatech.edu/research-data/data-assessment)

Data Interview

- University of Virginia Data Interview Initiative
- Emory Research Data Interviews.
  [http://guides.main.library.emory.edu/rdm-interviews](http://guides.main.library.emory.edu/rdm-interviews)
Formal Assessment: Examples

Survey

- Cornell
  [http://escholarship.umassmed.edu/jeslib/vol1/iss2/1/](http://escholarship.umassmed.edu/jeslib/vol1/iss2/1/)

Content Analysis

- Georgia Tech analysis of submitted DMPs:
  [https://smartech.gatech.edu/bitstream/handle/1853/44391/NSF_DMP_Content_Analysis.pdf](https://smartech.gatech.edu/bitstream/handle/1853/44391/NSF_DMP_Content_Analysis.pdf)
Assessment and Analysis

Resources to guide you

• Data Curation Profiles
  http://datacurationprofiles.org/

• Participating in the E-Science Institute
  http://duraspace.org/e-science-institute
  • Time and financial commitment

What will you do with your results?
Additional Resources and Links


EDUCAUSE. *Developing a Research Data Management Plan Service.*
http://www.educause.edu/library/resources/developing-institutional-research-data-management-plan-service

Sherry Lake. Conducting an Environmental Scan: Who’s Important on Your Campus. DMPTool Webinar Series 4.
http://www.slideshare.net/UC3/dmptool-webinar-4

http://crln.acrl.org/content/73/1/10.full?sid=89c2612a-89d3-43c3-88a0-a9c23a056eff

Thank You!

• Contact information:

• Jen Doty
  Emory University
  jennifer.doty@emory.edu

• Wendy Mann
  George Mason University
  wmann@gmu.edu