Research Data Management

A View from Penn State University

Patricia Hswe, Digital Collections Curator

EDUCAUSE 2011 | October 20
Panel: Access to Data Tomorrow Means Good Data Management Today
The data management landscape at Penn State (& elsewhere, very likely)

Where is the data? Who has data? What types of data are there?
Penn State’s “enterprise” of response

- **Data Curation Services Working Group (DCSWG)**
  - Collaboration between Libraries, Information Technology Services (ITS), Office of Sponsored Programs, and faculty (research institutes)
  - **Activities:** survey of NSF PIs and pending PIs; future – benchmarking services with peer institutions

- **Research Data Management Services Team (RDMST)**
  - Based in libraries, diverse membership
  - **Activities:** website, consultations on data management planning, information session in spring 2011, brochure for distribution in the Libraries and beyond (i.e., to faculty and students), data management planning workshop for graduate students
Research Data Management Toolkit

This toolkit is intended to help orient teaching and research faculty, as well as undergraduate and graduate students, around the topic of data management planning. Please contact the Research Data Management Services Team for consultations on data plans.

NEW RESOURCE: Guidance for Penn State Researchers on Developing Data Management Plans (PDF)

Data Planning

What the NSF says, questions to ask about your data

Data Management

Metadata (describing data), storage, sustainability

Data Sharing

Making data available and accessible

Contact us!

The Libraries can help you manage your data.

Primary Contact Information

For questions about data management planning, please contact a co-chair of our Research Data Management Services Team:

- Patricia Hswe - phswe@psu.edu
- Marcy Bidney - mma17@psu.edu

http://www.libraries.psu.edu/scholar/datamanagement.html

Our Service Statement

The Research Data Management Services Team (RDMST) offers consultation services for data management planning to researchers applying for grant funding from agencies requiring a data management plan (DMP). We ask for a minimum of two weeks in which to help you develop a DMP. This will allow adequate time to schedule a face-to-face consultation and to discuss the nature of the research data, policies for sharing and access, expectations for use and re-use, and other relevant issues. For requests made less than two weeks before a grant proposal deadline, we encourage you to read through the toolkit carefully and email RDMST with questions and concerns to our address: UL-RDMST@lists.psu.edu.
Data Management

Documenting your data is an initial step in managing them

Managing data is an integral part of the research process. It can be challenging particularly when studies involve several researchers and/or when studies are conducted from multiple locations. How data is managed depends on the types of data, how data is collected and stored, and how it is used through the length of the study.

The outcome of your research depends in part on how well you manage your data. Managing data helps you as a researcher organize research files and data for easier access and analysis. It helps ensure the quality of your research. It supports the published results of your work and, in the long term, helps ensure accountability in data analysis. Good data management starts with comprehensive and consistent data documentation and should be maintained through the life cycle of the data.

- Designate the responsibilities of every individual involved in the study.
- Designate how data will be stored and backed up.
Data Management Planning Guidance

• Faculty sought a template or guidance based on Penn State conditions (e.g., research administration policies/guidelines)

• The Libraries and ITS partnered with the Strategic Interdisciplinary Research Office (within the OVPR office) to develop this guidance
This section is a key reason why the NSF now requires a data management plan as part of the proposal. Here, you should provide a statement about access to the data and any policies related to sharing the data, particularly if the data your project will produce will require restricted access (e.g., because of human subject research), or a period of embargo (e.g., for publisher, patent, or commercial reasons). That is, any issues regarding "privacy, confidentiality, security, intellectual property, or other rights or requirements" need to be stated here, as suggested in the NSF Grant Proposal Guide.

PSU policies that apply to this section and that should be reviewed:

- AD23: Use of Institutional Data: [http://guru.psu.edu/policies/AD23.html](http://guru.psu.edu/policies/AD23.html)
- AD35: University Archives and Records Management: [http://guru.psu.edu/policies/AD35.html](http://guru.psu.edu/policies/AD35.html)
- RA14: The Use of Human Participants in Research: [http://guru.psu.edu/policies/RA14.html](http://guru.psu.edu/policies/RA14.html)
Services (Infrastructure) Development


• **Since summer 2011**:
  - ITS (specifically, Digital Library Technologies) has been working with Hydra community (cf. [http://projecthydra.org/](http://projecthydra.org/)) on a framework for repository-based services
  - ITS is also planning new, tiered storage services for research data, consulting with Research Data Management Services Team

• **Goal**: new suite of publishing and curation services in production in fall 2012
Challenges (Opportunities)

• Engaging more subject librarians in data curation efforts
  – ARL/DLF E-Science Institute
• Taking entrepreneurial approach (following up with new NSF/NIH/NEH PIs)
• “Data literacy” for graduate students
• Meeting other commitments, including but not limited to:
  – Addressing commercial, born-digital data sets
  – Publishing (ingesting) student-based work
  – Developing e-records management program
  – Satisfying faculty requests for help in starting peer-reviewed, online journals
Thank you!

• Questions? Comments?
  – Email me: phswe@psu.edu

Picture credit:
http://www.flickr.com/photos/jmd41280/2975460863/ • CC-BY-ND by jmd41280