Local ICPSR Data Curation Workshop Pilot Project

Linda Detterman, Jennifer Doty, Jared Lyle, Amy Pienta, Lizzy Rolando, and Mandy Swygart-Hobaugh

Overview

Researchers are now increasingly encouraged or required to share and archive their data, yet training in good data practices is still lacking. In a 2009-10 survey of data sharing by scientists (Tenopir et al., 2011), nearly two-thirds (59 percent) of respondents reported that their organization or project does not provide training on best practices for data management. Only one-third (35 percent) of respondents said they “are provided with the necessary tools and technical support for long-term data management.”

Libraries are well-positioned to help researchers fulfill data policies and possess the skill sets, longevity, and infrastructure needed to manage, disseminate, preserve, and track usage of data (Heidorn, 2011). Yet, they, too, indicate a desire to train their own staff since many librarians enter the profession with minimal or no data experience. A recent analysis of iSchool and LIS programs, for instance, found less than a quarter offer a course focused on research data management and curation (Creamer et al., 2012).

Domain repositories have long-term expertise in data management and curation, and they are increasingly interested in connecting with and training their user communities, although they have limited opportunities to meet researchers locally.

Objectives

This pilot project teamed a social and behavioral science domain repository, ICPSR, with three local universities, Emory University, Georgia Tech, and Georgia State, to offer two day-long data curation trainings: one for faculty, graduate students, and research staff, and another day for librarians and library staff.

The goals:

• Raise awareness of funder requirements and journal policies to preserve and share data, and resources available to help.
• Educate both researchers and librarians in best practices for documenting, preparing, and curating data for long-term preservation and sharing.
• Provide guidance and support to researchers depositing their data with appropriate domain repositories (e.g., ICPSR, Dryad).
• Offer an opportunity to reach the researchers where they reside.

Agenda

Workshop content was patterned on the ICPSR Summer Program workshop, “Curating and Managing Research Data for Re-use.” Presentations were followed by hands-on exercises and discussion.

Topics covered included:

• Identifying and Finding Data to Archive
• Reviewing Data
• Reviewing Data for Confidentiality Issues
• Cleaning Data
• Describing Data
• Depositing Data
• Disseminating and Publishing Data
• Local curation resources

Feedback

Positive:

• “This was a great workshop and I’m glad that I had the opportunity to attend. It made me want to learn more and provided me with great resources that I can return to and explore.”
• “Got both a broad yet detailed enough view of what ICPSR is, chances to ask my project specific questions, and helpful hands-on sessions.”
• “I learned a lot, and the topics were varied enough to give an overview, but not so in-depth as to be overwhelming.”

Suggestions:

• “I believe I misunderstood what the workshop was about. The description was perhaps too broad.”
• “Some workshops were hard to complete within a given time range. Workshop should be easier to comprehend (what we have to do) and more focused (e.g., fewer tasks/questions).”

Next Steps

For ICPSR:

• Revise agenda to vary the approach between researchers and librarians.
• Condense into one day with two distinct sessions:
  • Researchers: emphasize data management planning, best practices in preparing data, and how to deposit data.
  • Librarians: focus on curation topics and hands-on experiences.
• Plan for additional offerings in other locations.
• Make materials available for anyone to reuse and remix.

For local institutions:

• Identify related training to offer locally.
• Adopt methods to support researchers preparing data for archiving and sharing.
• Explore additional opportunities to partner with domain data archives.

References

Creamer, Andrew T., Morales, Myrna E., Kafel, Donna, Crespo, Javier, and Martin, Elaine R.

http://escholarship.umassmed.edu/cgi/viewcontent.cgi?article=1016&context=jeslib

http://doi.org/10.1080/01930826.2011.601269