April 9, 2014

Describing Data Repositories

Jennifer Eustis, *University of Massachusetts Amherst*
Carolyn Mills, *University of Connecticut - Storrs*

University of Massachusetts and New England Area
Librarian e-Science Symposium

2014 e-Science Symposium

Apr 9th, 12:00 AM

Describing Data Repositories

Jennifer M. Eustis
University of Connecticut - Storrs

Carolyn Mills
University of Connecticut - Storrs

Follow this and additional works at: https://escholarship.umassmed.edu/escience_symposium

Part of the Scholarly Communication Commons

Repository Citation

DESCRIBING DATA REPOSITORIES
At The University Of Connecticut Libraries

With the rise of eScience, subject liaisons must become familiar with disciplinary data repositories to better serve their clientele. Research data can often be deposited in one or more repositories. For researchers who are not well informed or work in fields that have yet to develop a data repository existing lists such as DataBib, Registry of Research Data Repositories or OpenDOAR provide a combined list of up to 2000 data repositories but little information about each one. At the University of Connecticut Libraries, to support our subject liaisons and our clientele, we created the "Describe Your Data Repository Survey".

The Survey

The survey was designed to retrieve as much information as possible about data repositories. As a result, not all may be applicable.

Section 1: Broad Information

This section asks to provide a general description of the data repository: name, url, subjects covered, governing body, associated fees, policies, licenses.

Section 2: Access To Data

This section asks to provide information about searching (basic/advanced), restrictions on accessing data, data citation.

Section 3: Depositing Data

This section asks to provide information about the submission process (acceptable formats, embargo, policies, restrictions, etc.)

Section 4: Managing Data

This section asks to provide information about how the submitting interacts with the data after deposit.

Section 5: Metadata and/or Additional Documentation

This section asks to provide information about standards and/or vocabularies used to describe data .

Section 6: Other Comments

Distribution

It was important to get investment from our subject librarians for this survey. We faced several challenges:

- *The survey was long consisting of 5 major sections that asked for descriptions and detailed information.*
- *Not all the information was applicable. Each data repository came to be seen as unique.*
- *Familiarity with data repositories and issues related to eScience varied among subject librarians.*
- *Subject librarian's hesitation to filling out the survey because of the time required to complete it (approximately 2 hours per data repository).*

To respond to these challenges, we took the following initiatives:

- *Added the task of filling out the survey to the list of annual goals*
- *Provided hands-on workshops*
- *Sent out multiple announcements and spoke at subject liaison meetings*

Observations

NOT all data repositories are equal!

- *Problem of Open vs. Closed data repositories*
- *Problem of lack of information*
- *Problem of limited functionality and difficult user interfaces*

NOT all disciplines have a data repository.

- *Absence of standards, norms and guidelines for the discipline*

NOT all data sets are alike.

- *The case of the digital humanities*

Some repositories cover too many disciplines.

- *ICPSR*

Next Steps

Continue to solicit input from subject libraries.

Make the results accessible.
- *Create records in our catalog that can be also accessed through our discovery layer.*

Integrate results in our eScience workshops to subject liaisons & our clientele.

Advertise these results on our data management LibGuide.

Track usage of our results through Google Analytics and url tracking.

Jennifer Eustis & Carolyn Mills, eScience Team, University of Connecticut Libraries

UCONN | **University Libraries**