Research Data Services: A New Focus for Librarians

Carol Tenopir
University of Tennessee
ctenopir@utk.edu
Today, I will talk about…

• 3 motivations for thinking about research data services
• A bit about the DataONE project
• Why libraries fit with research data services
• Some definitions
• Current status and future plans for data services of academic libraries and librarians in North America

But first…
A bit about me...

Information Sciences Professor, College of Communication and Information Director of Research, and Director of the Center for Information and Communication Studies
Tennessee
Knoxville
University of Tennessee, Knoxville
Motivation 1: Science is more collaborative, interdisciplinary, computational, and data-intensive
Motivation 2: Open Science Movement…

…governments and funding agencies are requiring data management

**United States**

“To that end, I have issued a memorandum today to Federal agencies that directs those with more than $100 million in research and development expenditures to develop plans to make the results of federally-funded research publically available free of charge within 12 months after original publication. …the memorandum requires that agencies start to address the need to improve upon the management and sharing of scientific data produced with Federal funding.” – U.S. Office of the President

**Australia**

*RCUK Policy on Open Access and Guidance* – immediate, unrestricted, online access to peer-reviewed and published research papers from RCUK funded projects

*Australian Code for the Responsible Conduct of Research* - basic data management mandate for ARC and NHMRC funding
Motivation 3: Data loss potential

- Natural disaster
- Facilities infrastructure failure
- Storage failure
- Server hardware/software failure
- Application software failure
- External dependencies
- Format obsolescence
- Legal encumbrance
- Human error
- Malicious attack by human or automated agents
- Loss of staffing competencies
- Loss of institutional commitment
- Loss of financial stability
- Changes in user expectations and requirements
NSF Sustainable Digital Data Preservation and Access Network Partners (DataNet)

Will create exemplar partners to address “...one of the major challenges of this scientific generation: how to develop the new methods, management structures and technologies to manage the diversity, size, and complexity of current and future data sets and data streams”.
The DataONE Vision and Approach:

Providing universal access to data about life on earth and the environment that sustains it, as well as the tools needed by researchers.

1. Building community
2. Developing sustainable data discovery and interoperability solutions
3. Enabling science through tools and services
Need new types of organizations that will...

- integrate library and archival sciences, cyberinfrastructure, computer & information sciences, and domain science expertise to:
  - provide reliable digital preservation, access, integration, and analysis capabilities for science and/or engineering data over a decades-long timeline
Libraries…

• facilitate interdisciplinary work, data access, and data knowledge through collections and services

• Can take a leadership role in a variety of research data services
Research Data Services are…

• …services that address the full data lifecycle, including the data management plan, digital curation (selection, preservation, maintenance, and archiving), and metadata creation and conversion.
Technical/Hands-On RDS

- Providing tech support for RDS systems or data repositories
- Identifying data for deposit in repositories
- Preparing data for deposit
- Creating or transforming metadata
- Deselecting datasets for removal from a repository
Informational/Consulting RDS

- Consulting on Data Management Plans and standards
- Outreach/collaboration with others
- Reference for finding and citing data
- Creating web guides and finding aids
- Directly participating on a project
- Training co-workers on data issues
Assessment of Stakeholders

Data Managers

Scientists

Students & Teachers

Public Officials

Citizen-scientists

Publishers

Libraries & Librarians

Center for Information and Communication Studies
Results of scientists survey shows opportunities for libraries…
What metadata do you currently use to describe your data?

- none: 676
- My Lab: 266
- ISO: 97
- Open GIS: 96
- FGDC: 95
- EML: 95
- DC: 26
- DwC: 21
- DIF: 12
Gap between Willingness to Share and Accessibility

• Although 75% of scientists agree “I share my data”

• Only 36% agree “Others can access my data easily”
Reasons for Not Making Data Electronically Available

- Insufficient time: 45%
- Lack of funding: 34%
- No place to put data: 20%
- Do not have rights to make data...: 20%
- Lack of standards: 17%
- Sponsor does not require: 15%
- Do not need data: 13%
- Should not be available: 12%
- Other reasons for data not available: 12%
Academic Librarians and Academic Libraries

**Librarians**
- Opinions of individuals
- 302 librarians at 111 ARL libraries

**Libraries**
- What the libraries are actually doing or planning to do re: RDS
- 223 ACRL libraries
librarians (in academic research libraries)…

The University of Tennessee
Knoxville
Center for Information and Communication Studies
Do you interact with faculty, students, or staff in support of their research data services (RDS) as part of your regular job responsibilities? (n = 223)

- Integral to job (27.9%)
- Occasionally (40.5%)
- No (31.5%)
ARL librarians strongly or somewhat agree…

- RDS are as important as other services (82% of integral; 68% of occasional; 32% of no)

- RDS are a priority at my library (67% integral; 40% of occasional; 19% of no)
Librarians’ Skills, Knowledge, and Training Necessary to Provide RDS

I have the skills, knowledge, and training necessary to provide RDS. (n = 159)

- Integral
- Occasional
- No

I have sufficient subject expertise to provide RDS to my patrons. (n = 159)

- Integral
- Occasional
- No
Librarians’ Skills, Knowledge, and Training Necessary to Provide RDS

My library provides opportunities to develop skills related to RDS. (n = 159)

- Integral
- Occasional
- No

My library supports me to attend conferences/workshops on RDS. (n = 157)
For Those Not Currently Involved in RDS, Potential Motivations for

<table>
<thead>
<tr>
<th>Rank</th>
<th>Motivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>If my patrons request RDS</td>
</tr>
<tr>
<td>#2</td>
<td>If RDS became a responsibility in my job</td>
</tr>
<tr>
<td>#3</td>
<td>If my institution becomes more involved with RDS</td>
</tr>
<tr>
<td>#4</td>
<td>If my institution develops an IR that accepts data</td>
</tr>
<tr>
<td>#5</td>
<td>If external funding agencies require RDS</td>
</tr>
<tr>
<td>#6</td>
<td>If RDS becomes important to subject disciplines I support</td>
</tr>
<tr>
<td>#7</td>
<td>If I learn more about RDS</td>
</tr>
</tbody>
</table>
Next *libraries* (in all types of academic libraries)…
Informational / Consulting RDS
Currently Offered or Planned

- Yes, currently offered
- No, and no plans
- No, but > 24 months
- No, but 13-24 months
- No, but < 12 months

Center for Information and Communication Studies
Informational / Consulting RDS
Currently Offered or Planned

- **Train co-workers**: 100%
- **Discuss regularly**: 90%
- **Project Participant**: 80%
- **Create guides**: 70%

Legend:
- No, and no plans
- No, but > 24 months
- No, but 13-24 months
- No, but < 12 months
- Yes, currently offered
Technical / Hands-On On RDS Currently Offered or Planned

- No, and no plans
- No, but > 24 months
- No, but 13-24 months
- No, but < 12 months
- Yes, currently offered
Reference support services, by type of institution…

Associates Colleges

- Yes: 40%
- No: 53%
- Planned: 7%

Baccalaureate Colleges

- Yes: 40%
- No: 41%
- Planned: 19%

Research/Doctoral

- Yes: 49%
- No: 18%
- Planned: 33%

Does your library provide reference support for finding and citing data and data sets? (n = 216)
Does your library provide technical support for research data services systems? (n = 216)
Staff development, by type of institution...

Has your library provided opportunities for staff to develop skills related to RDS? (n = 215)
Conclusions

• RDS are important and consistent with library mission and role
• There are many levels of RDS involvement
• Libraries are at an early point in transition to RDS—requiring resetting of priorities, realignment of responsibilities, and opportunities to develop skills
• Librarians need opportunities to develop RDS skills
Data management services are part of transforming research libraries and librarians…

This is next-generation librarianship. The curation of research data is an activity that has gained traction in the wake of library and information science programs offering concentrations in data curation and institutes in digital curation, promising a cohort of librarians qualified to meet the challenges of managing data.

http://www.arl.org/rtl/eresearch/escien/NSF/leadershiproles.shtml
Authors: Patricia Hswe and Ann Holt. Part of ARL’s transforming research libraries series.

• Academic Librarians and Research Data Services: Preparation and Attitudes, IFLA Journal 2013.  
http://ifl.sagepub.com/content/39/1/70.abstract

http://www.plosone.org/article/info:doi/10.1371/journal.pone.0021101
Thank you!

Carol Tenopir
University of Tennessee
ctenopir@utk.edu