Learning by Doing: Cases of Librarians Working with Faculty Research Data for the First Time

IASSIST 2010

Jake Carlson
Data Research Scientist

Michael Witt
Interdisciplinary Research Librarian

Purdue University Libraries
The Roles for Libraries

“What role do the research and academic libraries envision for themselves and do scientists envision for librarians in a digital data framework…?”

- To Stand the Test of Time…(2006) ARL. p.24
Research traditionally moves in this direction

```
| "published" data/datasets | unpublished research traditional/non | "published" research non-traditional | published research traditional | secondary tertiary resources |
```

Library service traditionally moves in this direction

Given the changing nature of research and scholarly communication (i.e., e-research), how do librarians change to adapt to new ideas regarding what they collect to support research & education and how they provide access to it?

E-data Task Force

Goals:

- To move the development of a data repository and data curation services from research to operational status in the Libraries.
- To inform development and formalization of the activities and infrastructure necessary for supporting a data repository and data services.
- To engage a wider audience of librarians in exploring issues related to data curation and developing data collections.
Membership

- Michael Witt, Interdisciplinary Research Librarian
- Jake Carlson, Data Research Scientist
- Mark Newton, Digital Collections Coordinator
- Chris Miller, GIS Librarian
- Marianne Bracke, Agriculture Information Specialist
Activities

- Complete a “Data Repository Prospectus”
- Obtain a diverse array of sample datasets and ingest them into a prototype repository.
- Draft a simple metadata template for each dataset.
- Draft policies for the selection, submission, use, and preservation of these datasets.
- Summarize what was learned from this experience and develop recommendations.
Data Sets

- edata 1 - Controversial Facilities in Japan (Witt)
- edata 2 - Vehicle Signatures (Newton)
- edata 3 - Materials Thermal Properties (Witt)
- edata 4 - Professional Development for Health Care Professionals (Carlson)
- edata 5 - Surface Water Provider Districts of Indiana (Miller)
- edata 6 - Phosphorus and Potassium Influence on Alfalfa Nutrition (Bracke)
Subject Librarians

- Bert Chapman - Political Science
- Megan Nelson - Civil Engineering
- Amy van Epps – Materials Engineering
- Jean-Paul Herubel - Communications
- Maribeth Slobodnik – Medical Sciences
- Chris Miller and Marianne Bracke
Output

Report

- Services before repository.
- Service definition – determine the roles, infrastructure, and cost model from this definition.
- Resources – an ongoing commitment of resources (money and personnel).
- Roadmap – need a plan of action to move from research to operation.
- Urgency.
Application of Librarian’s Skills

- Subject Knowledge
  - Knowing the major resources
  - Knowing how others in the field use data

- Collection Development
  - Value and Justification

- Connections to Faculty
  - Trusted Relationships

- Librarian’s Perspective
  - How others would make use of the data - needs
Librarian’s Skills: Marianne
Challenges

- Lack of Knowledge
  - Metadata
  - Terminology / Data Issues
  - Technologies
  - Formats

- Time
  - Of Librarian
  - Of Data Author
Designing a Training Program

- Strengthening liaison relationships
  - Expose librarians to how researchers use CI and tools
- Understanding of audience needs
- Being able to discuss options for data
- Explanation of how data interacts traditional librarianship
- Metadata
- Technologies
Training Program: Chris
What was learned

- “Building a data repository is even harder than I thought”
- Different kinds of data that faculty work with (scope, format, use)
- Diversity of data issues (problems, requirements, expectations)
- Better able to conduct a data interview
- Better knowledge of what needs to be learned
What was learned: Megan
Roles

Potential Roles
- Education / Outreach – “having conversations”
- Broker between data curator and researcher
- Metadata
- Collection Development Type Activities

“Data” Work vs. “Library” Work
Roles: Megan
Roles: Chris