Title
Challenges in Developing A New Library Infrastructure for Research Data Services

Permalink
https://escholarship.org/uc/item/8x36m8sv

Authors
Tsang, Daniel C
Renaud, John P

Publication Date
2014-06-06
Challenges in Developing A New Library Infrastructure for Research Data Services

Daniel C. Tsang
John P. Renaud
University of California, Irvine, Libraries
IASSIST Conference 2014
Social Science Data Librarians

- Existed for over half a century
- ICPSR began in 1962
- IASSIST is celebrating its 40th year
- Work in many academic libraries and research centers
- Let’s not forget the pioneering role
Roles Have Changed

Librarians are moving from just helping patrons find material such as books, journal articles and even data to partnering with faculty in the life cycle of a research project, especially helping the faculty meet national mandates for archiving research data.

In other words, the need for expertise in research data management has imposed new demands on staffing and the possibilities for new configurations of organizational structure.
“The role of libraries in research institutions is evolving from a focus on reader services to a focus on author services (an insight first voiced by Kimberly Douglas of Caltech).” – Christine L. Borgman, 2012

“The processes associated with gathering or producing research data are a form of authorship, whether or not the researchers accept that view…” – ibid.

Essentially Borgman is suggesting that we can support and even partner with researchers in their research process, one views data along a life cycle.
“Some research institutions already maintain data archives and their curators provide valued input on how to prepare, acquire, and curate data during the research data life cycle. See, for example, UCLA’s Social Science Data Archive Policy on Acquisitions and Archiving.”

Data Librarian Jobs Changing

“Competencies and Responsibilities of Social Science Data Librarians: An Analysis of Job Descriptions” by Jingfeng Xia and Minglu Wang, College & Research Libraries, July 2014 (forthcoming):

http://crl.acrl.org/content/early/2013/02/06/crl13-435.full.pdf
“Recognizing the importance of data management in social sciences was grounded on the development of scientific data management and was, therefore, a little behind in time. This availability of job advertisements through IASSIST’s website permits an examination of the short history of data librarianship in social sciences from the very beginning in the mid-2000s. The fact that job announcements first appeared on IASSIST’s job portal in 2005 may signify the start of the profession as well, which can serve perfectly as the baseline year of our analysis, and allow us to explore the chronological changes of the profession with regard to its responsibilities and competency requirements as well as preferences that are stated in the job advertisements.”

“In the early years of this time sequence, the responsibility requirements focused on various forms of instructions, from delivering formal classroom teaching to arranging specialized workshops, and on different types of library services, from providing general references to offering specific data services… This instruction-and service-orientation in the job responsibilities has changed to an increasing focus on data management from 2008 on.”

Source: “Competencies and Responsibilities of Social Science Data Librarians: An Analysis of Job Descriptions” by Jingfeng Xia and Minglu Wang, College & Research Libraries, July 2014
“It becomes apparent that social sciences data professionals are still performing traditional primary services in the stages of data discovery, data collection and data analysis… At the same time, support for data preservation as a relatively new task of research data services has already been taken up by the profession quickly, probably due to the professional spirit of performing data stewardship…”
Data Basics: An Introductory Text delineates three different levels of service by function:

Levels of Collection Service
Levels of Reference Service
Levels of Computing Service

# Levels of Collection Service

## Matrix of Some Collection and Collection Service Alternatives

<table>
<thead>
<tr>
<th>Level of service</th>
<th>Select</th>
<th>Acquire Access</th>
<th>Acquire data</th>
<th>Organize</th>
<th>Preserve</th>
<th>Service</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (low)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(x)</td>
<td>X</td>
</tr>
<tr>
<td>2</td>
<td>X</td>
<td></td>
<td></td>
<td>(x)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>6</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>(x)</td>
<td>X</td>
</tr>
<tr>
<td>7 (high)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

# Other Ways of Looking at Data Management Levels

## Data Management Layers

<table>
<thead>
<tr>
<th>Layers</th>
<th>Characteristics</th>
<th>Implication for PI</th>
<th>Implication relative to NSF</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Curation</strong></td>
<td>• Adding value throughout lifecycle</td>
<td>• Feature Extraction</td>
<td>• Competitive advantage</td>
</tr>
<tr>
<td></td>
<td>• New query capabilities</td>
<td>• New capabilities</td>
<td>• New opportunities</td>
</tr>
<tr>
<td></td>
<td>• Cross-disciplinary</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Preservation</strong></td>
<td>• Ensuring that data can be fully used and interpreted</td>
<td>• Ability to use own data in the future (e.g. 5 yrs)</td>
<td>• Satisfies NSF needs across directorates</td>
</tr>
<tr>
<td></td>
<td>• Data sharing</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Archiving</strong></td>
<td>• Data protection including fixity, identifiers</td>
<td>• Provides identifiers for sharing, references, etc.</td>
<td>• Could satisfy most NSF requirements</td>
</tr>
<tr>
<td></td>
<td>• Backup and restore</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Storage</strong></td>
<td>• Bits on disk, tape, cloud, etc.</td>
<td>• Responsible for:</td>
<td>• Could be enough for now but not near-term future</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Restore</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Sharing</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Staffing</td>
<td></td>
</tr>
</tbody>
</table>

Source: “Levels of Service and Curation for High Functioning Data” by G. Sayeed Choudhury *et al.*

The Stack

Increasing layers of support and functionality; each level depends on the level below....

- **Storage**: lowest service; basic physical storage with backup and restore services.
- **Archive**: following BRTF, “activities that enable long-term retention of digital materials”; DC focus on data protection through replication, fixity, and identifiers.

• **Preservation**: providing enough representation information, context, metadata, fixity, etc. to support use and interpretation by agents other than the original data producer.

• **Curation**: processes that add value to foster discovery and reuse.

The curation level identifies a range of services, enabling use for purposes not necessarily envisioned by the data producers.

RDM Service Staffing

The depth and range of RDM services that libraries offer are, of course, directly proportional to staffing, both in the number of positions and the amount of time given to RDM activities when a position has other responsibilities. Currently, the most prevalent organizational structure for providing RDM services at the 53 responding libraries is a committee of staff from departments within the library (27, or 51%)(Q32). Less common organizational structures include a committee/group comprised of staff from across the university, including the library (9, or 17%), a single position within the library (8, or 15%), and a single department within the library (6, or 11%).

At the libraries where RDM services are provided by staff from different departments, no single department dominates (Q33). About a quarter of the departments provide reference/liaison services, followed by work in collections (19%), digital services (12%), research/instruction (12%), and systems/IT (11%).

Variety of Job Titles

39. For each position, please indicate if it was created specifically for RDM services, substantially redesigned to focus on RDM, or had RDM activities added to existing responsibilities. N=53 respondents, 230 positions

<table>
<thead>
<tr>
<th>RDM activities added</th>
<th>146</th>
<th>63%</th>
</tr>
</thead>
<tbody>
<tr>
<td>New RDM services position</td>
<td>49</td>
<td>22%</td>
</tr>
<tr>
<td>Substantially redesigned</td>
<td>34</td>
<td>15%</td>
</tr>
</tbody>
</table>

40. Please enter the year that the person in each position began providing RDM services. N=52 respondents, 220 positions

Spec Kit Resources on Staffing

Staffing and Staff Training


The current layout:
Cross divisional teams that oversee the e-research and digital scholarship

- The Digital Scholarship/Scholarly Communication Strategic Council (DSSC) is charged with developing high-level strategy for the UCI Libraries’ digital scholarship and data curation and for the work of the Digital Services Operations Team (DSOT)
E-Scholarship and Digital Research Services at UCI…

• DSOT has as its charge the management all the services and projects in the areas of digital services, data and e-research on an interim basis, pending the creation of a permanent structure. DSOT is to support the UC Irvine community in providing persistent access to unique digital research assets.

• The Emerging Technologies & Service Innovation unit of the Reference Department, within the Public Services Division. The Unit’s role is to help the Libraries in determining the best tools and methods to use as the Libraries move more content and services online to augment or transform in-person services with digital asynchronous services.
Data Librarian

Within the Libraries:
The UCI Data Librarian

• Selects datasets for adding to the collection
• Maintains Libguides on Social Science Data & on Research Data Management
• Helps users find data
• Manages the UCI Social Science Data Archive with online analysis for selected datasets via SDA
Data Librarian…

Campus-wide:
• Serves as ICPSR Representative
• Sits on an advisory committee overseeing UCI’s Research Data Network
• Acts as the Liaison to the UCI Office of Research
• Administers California Digital Library’s DMPTool & Web Archiving Service
A New Librarian Position?

• In order to meet better position ourselves for the near future, we envision as potentially recruiting for the position of E-Research and Digital Scholarship Services Librarian; it is one of several expansion options

• Several organizational structure issues had to be addressed before we could post the position for recruitment
Challenges and Tough Questions

- Overall organizational structure
- Politics – many folks want to be included
- Understanding what the developments in e-research and digital scholarship services might mean for the Libraries as a whole.
Challenges and Tough Questions...

- Staffing is paramount in the decision-making process.
- The Collection Development Department is the home unit to the Data Librarian as well as librarians with skills and interest in the areas of bioinformatics and digital humanities.
- The personnel with the technical skills to initially support work in these areas are located in the Library IT and Cataloging and Metadata Services Departments.
- The least disruptive course is to potentially move selected IT and Cataloging and Metadata Services personnel to a new unit within the larger department.
• The question was raised as to whether a unit [rather than a department] provided sufficiently high-level representation in the administrative structure.

• The highly collaborative nature of the organization and of the collection development department made this less of an issue in UCI’s case.
Stability and Change

- Digital Scholarship/Scholarly Communication Strategic Council (DSSC) will potentially continue, with likely revised representation.
- Much of the work of the Digital Services Operations Team (DSOT) will be potentially transferred to a new unit within the Collection Development department.
Moving Forward

• Many discussions across the Libraries, at the leadership, departmental, and divisional levels
• Plan being discussed as part of Library-wide Town Hall Meeting and a follow-up survey soliciting feedback
• Potentially posting a position for an E-Research and Digital Scholarship Services Librarian
Digital Scholarship Services at UCI Libraries

One-stop Destination for Help with Research and Beyond

Thank You!

Please contact us!

Daniel C. Tsang, Data Librarian
dtsang@uci.edu

John P. Renaud, AUL for Research Resources
jrenaud@uci.edu