

# 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 40<sup>th</sup> 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

Source: “Research Data: Who Will Share What, with Whom, When, and Why?” by Christine L. Borgman, *RatSWD Working Paper* No. 161 (1 October 2010):

[http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1714427](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1714427)

# Helping researchers ‘publish’

“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.

# Social Science Data Archivists Have Research Data Management Expertise

“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*.”

Source: Starting the Conversation: University-wide Research Data Management Policy by Ricky Erway. OCLC Research, 2013: <http://oclc.org/content/dam/research/publications/library/2013/2013-08.pdf>

# 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>



# Assumptions in Article

“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.”

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, p.



# Key Findings

“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

# Key Findings continued

“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: Levels of Service*

*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

Source: *Data Basics: An Introductory Text* by Diane Geraci, Chuck Humphreys and Jim Jacobs, 2012, Chapter 11. Available online: [http://3stages.org/class/2012/pdf/data\\_basics\\_2012.pdf](http://3stages.org/class/2012/pdf/data_basics_2012.pdf)

# Levels of Collection Service

Matrix of Some Collection and Collection Service Alternatives

Level of service	Select	Acquire Access	Acquire data	Organize	Preserve	Service
1 (low)						X
2	X			(X)		
3	X	X				
4	X	X		X		
5	X	X		X		X
6	X	X		X	(X)	X
7 (high)	X	X	X	X	X	X

Source: *Data Basics: An Introductory Text* by Diane Geraci, Chuck Humphreys and Jim Jacobs, 2012, Chapter 11. Available online:

[http://3stages.org/class/2012/pdf/data\\_basics\\_2012.pdf](http://3stages.org/class/2012/pdf/data_basics_2012.pdf)

# Other Ways of Looking at Data Management Levels

## Data Management Layers

Layers	Characteristics	Implication for PI	Implication relative to NSF
<b>Curation</b>	<ul style="list-style-type: none"> <li>• Adding value throughout life-cycle</li> </ul>	<ul style="list-style-type: none"> <li>• Feature Extraction</li> <li>• New query capabilities</li> <li>• Cross-disciplinary</li> </ul>	<ul style="list-style-type: none"> <li>• Competitive advantage</li> <li>• New opportunities</li> </ul>
<b>Preservation</b>	<ul style="list-style-type: none"> <li>• Ensuring that data can be fully used and interpreted</li> </ul>	<ul style="list-style-type: none"> <li>• Ability to use own data in the future (e.g. 5 yrs)</li> <li>• Data sharing</li> </ul>	<ul style="list-style-type: none"> <li>• Satisfies NSF needs across directorates</li> </ul>
<b>Archiving</b>	<ul style="list-style-type: none"> <li>• Data protection including fixity, identifiers</li> </ul>	<ul style="list-style-type: none"> <li>• Provides identifiers for sharing, references, etc.</li> </ul>	<ul style="list-style-type: none"> <li>• Could satisfy most NSF requirements</li> </ul>
<b>Storage</b>	<ul style="list-style-type: none"> <li>• Bits on disk, tape, cloud, etc.</li> <li>• Backup and restore</li> </ul>	<ul style="list-style-type: none"> <li>• Responsible for:               <ul style="list-style-type: none"> <li>• Restore</li> <li>• Sharing</li> <li>• Staffing</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• Could be enough for now but not near-term future</li> </ul>

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

<http://www.dcc.ac.uk/sites/default/files/documents/idcc13posters/Poster192.pdf>

# 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.

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

<http://www.dcc.ac.uk/sites/default/files/documents/idcc13posters/Poster192.pdf>

# Stack...

- **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.

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

<http://www.dcc.ac.uk/sites/default/files/documents/idcc13posters/Poster192.pdf>

# ARL Spec Kit

## **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%).

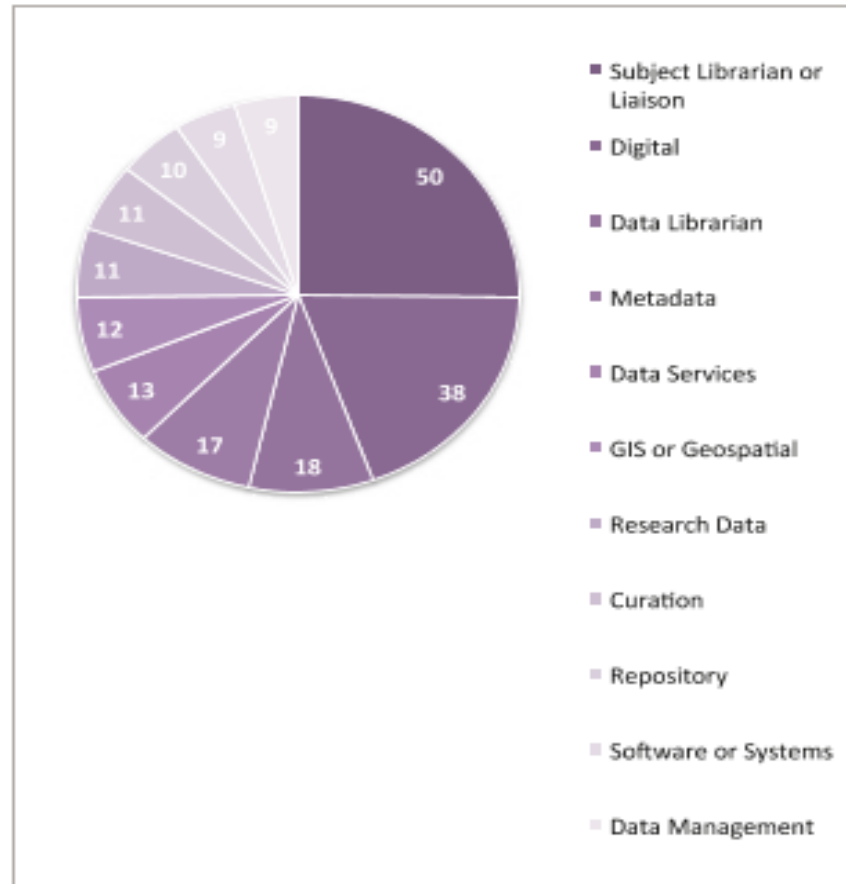
At the libraries where RDM services are provided by staff from different departments, no single department dominates (Q33). About a quarter of the departments provide reference/liaison services, followed by work in collections (19%), digital services (12%), research/instruction (12%), and systems/IT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

Source: ["Research Data Management Services," ARL Spec Kit 334 \(July 2014\):](http://publications.arl.org/Research-Data-Management-Services-SPEC-Kit-334/17)  
["http://publications.arl.org/Research-Data-Management-Services-SPEC-Kit-334/17"](http://publications.arl.org/Research-Data-Management-Services-SPEC-Kit-334/17)





# Variety of Job Titles



Source: "Research Data Management Services," ARL Spec Kit 344 (July 2013), p17:  
<http://publications.arl.org/Research-Data-Management-Services-SPEC-Kit-334/17>

# Research Data Management Services

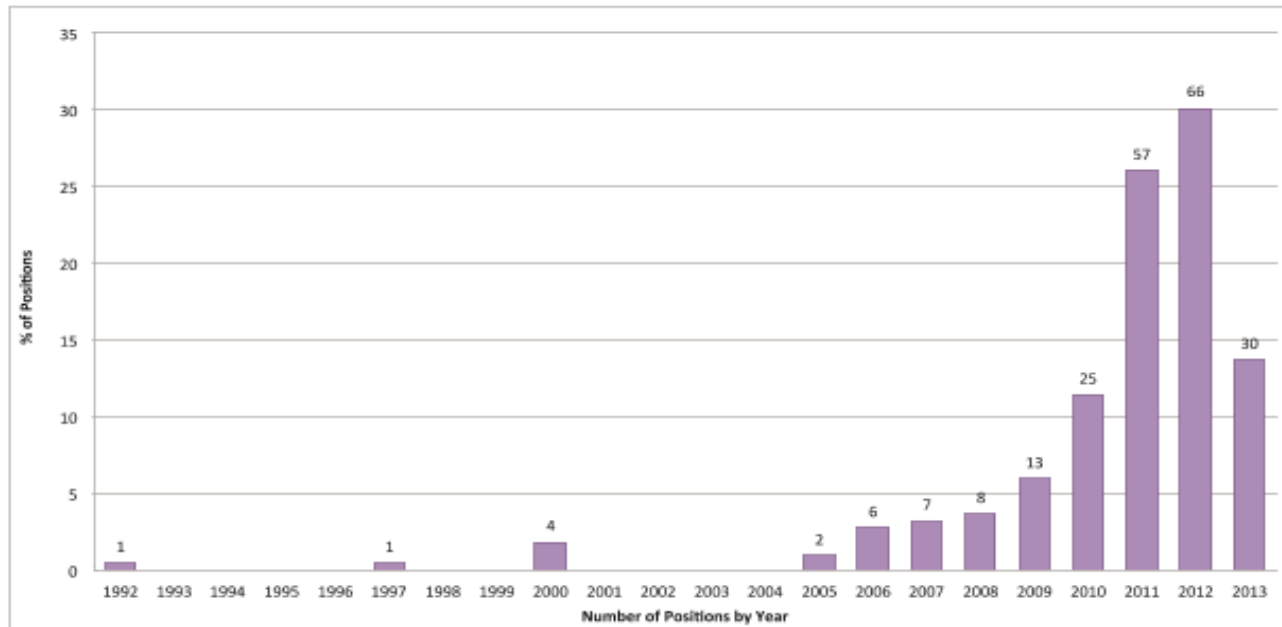
39. For each position, please indicate if it was created specifically for RDM services, substantially redesigned to focus on RDM, or had RDM activities added to existing responsibilities. N=53 respondents, 230 positions

RDM activities added	146	63%
New RDM services position	49	22%
Substantially redesigned	34	15%

Source: "Research Data Management Services," ARL Spec Kit 344 (July 2013), p69:  
<http://publications.arl.org/Research-Data-Management-Services-SPEC-Kit-334/69>

# Years of RDM Services

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



Source: "Research Data Management Services," ARL Spec Kit 344 (July 2013), p69:  
<http://publications.arl.org/Research-Data-Management-Services-SPEC-Kit-334/69>

# Spec Kit Resources on Staffing

## Staffing and Staff Training

---

Auckland, Mary. *Re-skilling for Research*. London: Research Libraries UK, January 2012. <http://www.rluk.ac.uk/content/re-skilling-research>.

Carlson, Jake. "Demystifying the Data Interview: Developing a Foundation for Reference Librarians to Talk with Researchers about Their Data." *Reference Services Review* 40, no. 1 (2012): 7–23. <http://dx.doi.org/10.1108/00907321211203603>.

Cox, Andrew, Eddy Verbaan, and Barbara Sen. "Upskilling Liaison Librarians for Research Data Management." *Ariadne* 70 (2012). <http://www.ariadne.ac.uk/issue70/cox-et-al>.

Creamer, Andrew, Myrna Morales, Donna Kafel, Javier Crespo, and Elaine Martin. "A Sample of Research Data Curation and Management Courses." *Journal of eScience Librarianship* 1, no. 2 (2012). <http://dx.doi.org/10.7191/jeslib.2012.1016>.

Creamer, Andrew, Myrna Morales, Javier Crespo, Donna Kafel, and Elaine Martin. "An Assessment of Needed Competencies to Promote the Data Curation and Management Librarianship of Health Sciences and Science and Technology Librarians in New England." *Journal of eScience Librarianship* 1, no. 1 (2012). <http://dx.doi.org/10.7191/jeslib.2012.1006>.

Digital Curation Centre. "RDM for Librarians." 2013. <http://www.dcc.ac.uk/training/rdm-librarians>.

Source: "Research Data Management Services," ARL Spec Kit 344 (July 2013), p217: <http://publications.arl.org/Research-Data-Management-Services-SPEC-Kit-334/217>

# ARL Spec Kit cont'd

- EDINA and Data Library, University of Edinburgh. *MANTRA: Do-It-Yourself Research Data Management Training Kit for Librarians*. Edinburgh: Digital Curation Centre, 2013. <http://datalib.edina.ac.uk/mantra/libtraining.html>.
- Gabridge, Tracy. "The Last Mile: Liaison Roles in Curating Science and Engineering Research Data." *Research Library Issues: A Bimonthly Report from ARL, CNI, and SPARC*, no. 265 (August 2009): 15–21. <http://publications.arl.org/rli265/16>.
- Garritano, Jeremy R., and Jake R. Carlson. "A Subject Librarian's Guide to Collaborating on e-Science Projects." *Issues in Science & Technology Librarianship* 57 (Spring 2009). <http://www.istl.org/09-spring/refereed2.html>.
- Guy, Marieke. "RDM Training for Librarians." *DCC RDM Services Case Studies*. Edinburgh: Digital Curation Centre, 2013. <http://www.dcc.ac.uk/resources/developing-rdm-services/rdm-training-librarians>.
- Lyon, Liz. "The Informatics Transform: Re-Engineering Libraries for the Data Decade." *International Journal of Digital Curation* 7, no. 1 (2012): 126–38. <http://dx.doi.org/10.2218/ijdc.v7i1.220>.
- Pryor, Graham, and Martin Donnelly. "Skilling Up to Do Data: Whose Role, Whose Responsibility, Whose Career?" *International Journal of Digital Curation* 4, no. 2 (2009): 158–70. <http://dx.doi.org/10.2218/ijdc.v4i2.105>.
- Salo, Dorothea. *Syllabus: LIS 855, Digital Curation*. University of Wisconsin-Madison: School of Library and Information Studies, Spring 2012. <http://dsalo.info/wp-content/uploads/2011/05/855syll2012.pdf>.



# E-Scholarship and Digital Research Services at UCI

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

# Challenges and Tough Questions...

- 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



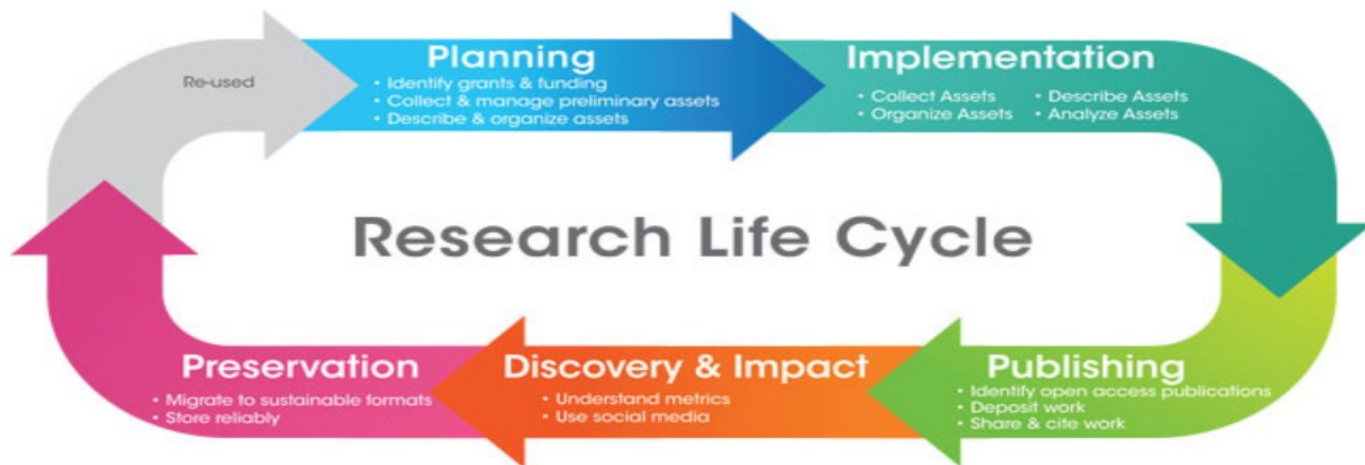
# UCI LIBRARIES

## DIGITAL SCHOLARSHIP SERVICES

- HOME
- ABOUT US
- SERVICES
- RESEARCH TOOLS
- COLLECTIONS
- HOW DO I

- DSS HOME
- PLANNING
- IMPLEMENTATION
- PUBLISHING
- DISCOVERY & IMPACT
- PRESERVATION
- UC OPEN ACCESS POLICY
- COPYRIGHT & FAIR USE
- SCHOLARLY COMMUNICATION
- Langson Library
- Ayala Science Library
- Grunigen Medical Library

One-stop Destination for Help with Research and Beyond



Source: "Digital Scholarship Services," UCI Libraries, 2014: <http://www.lib.uci.edu/dss/>



# Thank You!

Please contact us!

Daniel C. Tsang, Data Librarian

[dtsang@uci.edu](mailto:dtsang@uci.edu)

John P. Renaud, AUL for Research Resources

[jrenaud@uci.edu](mailto:jrenaud@uci.edu)