

2-2019

Iowa State University Library Research Data Services Gap Analysis

Curtis Brundy

Iowa State University, cbrundy@iastate.edu

Megan N. O'Donnell

Iowa State University, mno@iastate.edu

Follow this and additional works at: https://lib.dr.iastate.edu/cos_reports



Part of the [Scholarly Communication Commons](#)

Recommended Citation

Brundy, Curtis and O'Donnell, Megan N., "Iowa State University Library Research Data Services Gap Analysis" (2019). *Collections and Open Strategies Principles, Policies, and Reports*. 2.
https://lib.dr.iastate.edu/cos_reports/2

This Report is brought to you for free and open access by the Collections and Open Strategies at Iowa State University Digital Repository. It has been accepted for inclusion in Collections and Open Strategies Principles, Policies, and Reports by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

Iowa State University

University Library: Research Data Services Gap Analysis

February 2019

Curtis Brundy, Associate University Librarian for Scholarly Communication and Collections
Megan O'Donnell, Data Services Librarian

Where are we now?

The University Library has made progress in establishing itself on campus as a hub and leader for research data services, especially in the areas of education and training. Since the Fall of 2013, the library has worked with campus partners in order to address and direct questions related to data management plans, data storage, and data security to local experts. For the past three years, the library has held regular workshops on data management and data management plans. Liaison librarians have also visited graduate level classes to speak on the same topics.

The Library's Data Management Plan Guide (<https://instr.iastate.libguides.com/dmp>) provides the ISU community with access to a writing tool (DMPTool, <https://dmptool.org/>), summaries and links to funder requirements, and a short, self-paced learning module. This guide has seen significant growth in the number of pageviews over the past three years: from 2,493 views in 2016 to 7,144 views in 2018.

With the launch of DataShare in August 2018, the University Library expanded its services to include data curation through the process of publishing publicly accessible data produced by Iowa State University researchers. All data published on DataShare first goes through a data set disclosure process, which checks to make sure that there are no contractual, legal, or compliance reasons that would prohibit data sharing. This review process includes the Office of Intellectual Property and Technology Transfer, Office for Responsible Research, Office of Sponsored Programs Administration, and Ames Laboratory (when applicable). The data services librarian monitors the process to ensure that submissions are resolved in a timely manner. To date, DataShare has 16 datasets from ISU researchers, and submissions are growing with expanded awareness of the service. The DataShare website (<https://instr.iastate.libguides.com/datashare>) is also seeing a rise in page views since it was launched in the summer of 2018.

The Library is currently involved with a cross-campus group working to advance data sharing. The Data Sharing Task Force includes representatives from the Office of the CIO, Office of the VPR, University Library, and Faculty. The AUL for Scholarly Communications and Collections is a co-chair. The task force has created draft data guidelines, a data sharing compliance review, and helped launch DataShare. This is a productive collaboration that includes representatives from many key campus data stakeholder groups.

The Library does not currently offer dedicated space or equipment to support research data work by Iowa State researchers and students.

Both services, education and DataShare, are currently supported by only a handful of librarians, all of whom have other commitments. The data services librarian currently leads both data service areas while also covering additional liaison librarian responsibilities.

Where do we wish we were?

A recent survey of ISU researcher data use and data sharing found the least satisfaction in Iowa State's support and services in the areas of cataloging, long-term storage, and sharing. Addressing these needs will require greater support for data management education, with more frequent offerings at more advanced skill levels. Iowa State's results are supported by survey findings from member libraries of the Greater Western Library Association (GWLA) that revealed all members provide data management instruction (i.e. consults, workshops, and/or course related instruction) but that the level and scale of instruction varies greatly by institution. The most invested libraries offer from 3 -10 different types of data related workshops per year, often on very specific topics or tools. This is similar to what other U.S. library leaders in research data services offer.¹ Workshops offered by GWLA member libraries include,

- Working with the Command Line
- Text-Mining (Python)
- How to Manage Your Data
- Research Data Management Better Practices
- Organizing and Sharing Data
- Writing an Effective Data Management Plan
- Deposit your dataset in ScholarsArchive@OSU
- Storing, Backing Up and Archiving Data
- Reproducibility and Using the OSF
- Using Spreadsheets to Manage and Analyze Data

The survey also revealed a large spectrum of readiness and services related to data curation and the operation of a data repository. Curation strategies were about evenly split between mediated and unmediated deposit models. Data preservation strategies rarely addressed much beyond bit-level preservation and commitments to cross deposit and backup in systems such as LOCKSS, Rosetta, and Amazon Web Services. These results align with the OCLC's *The Realities of Research Data Management* report series² which noted that "it would be a mistake to imagine that there is a single, best model of RDM service capacity, or a simple roadmap to acquiring it." And yet, libraries know from experience with other digital projects that curation and preservation are critical components to any digital initiative.

The staffing levels and organizational structures of libraries that lead in providing research data services reflect a high level of institutional commitment. Data Management Services at John Hopkins Libraries, for example, is a department reporting directly to the library Dean and it includes approximately 5 FTE³.

Research data services are cross-disciplinary and interdepartmental endeavors, existing within and beyond the library. However, many libraries offer a dedicated space for data services that provides a crucial, centralized place for researchers to seek and receive help. Such spaces often include reservable

¹ i.e. [MIT](#) and [Johns Hopkins](#).

² <https://www.oclc.org/research/publications/2017/oclcresearch-research-data-management.html>

³ https://www.library.jhu.edu/wp-content/uploads/sites/3/2018/12/Org-Chart_10-2018.pdf;
<https://dms.data.jhu.edu/about-us/history-and-the-team/>

workstations with data processing software and peripherals. The libraries at MIT and UC San Diego both offer these types of dedicated spaces⁴.

How do we close the gap?

The University Library's research data services are in a position to grow. While the Library has successfully established basic services in the areas of education and training, curation, and preservation, a significant gap exists between where we are today and where we would like to be. Several areas of improvement and investment have been identified around staffing, collaboration, and facilities.

Staffing

Training and education on specific data tools, programming languages, and techniques should be added to the library's workshop line-up. Additionally, training tailored for specific groups is needed in order to onboard different communities into data management/sharing best practices and resources. Increased training and education will reaffirm ISU's commitment to data sharing, which should lead to more business for DataShare and increased demand for robust data curation services. OCLC (2018) stated "There are incentives for institutions to develop Education and Expertise services locally..." and that "provision of a minimal level of local services, particularly Expertise, is necessary for institutions to establish credibility - both with local researchers and also within the larger RDM scholarly communications community."⁵ Expanded data services and support in the areas of education and curation will require additional library staff and data expertise.

The Library proposes creating two new data services positions. The first, a Data Education Specialist, will focus on strategically expanding the Library's data management education and training program. This position will add expertise in teaching command line data management through computing languages, such as R and Python, as well as open science publishing tools, such as [Jupyter Notebook](#) and [R Shiny](#). The second hire, a Data Curation Specialist, will support the growth of DataShare, which will allow the Library to improve the quantity, quality, and robustness of the materials published through DataShare.

Reflecting the importance that the Library places on meeting researcher needs around data services, funding for the Data Education Specialist has already been allocated and a search for this position will begin this spring. Funding for the second position, the Data Curation Specialist, will require additional campus support. Ideally, this position would be filled in early 2020. Planning is also underway to transition the existing Data Services Librarian position to a full-time data services and support role. Together, these three data positions will comprise a new Research Data Services unit in the Library. Liaison librarian support will continue in a vital, but reduced, capacity under this new model.

Collaboration

The challenge ahead in meeting the research data support and service needs at Iowa State is large, spilling across the boundaries of any one campus unit, including the Library. Partnerships and

⁴ UC San Diego Library Data and GIS Lab <https://ucsd.libguides.com/c.php?g=383319&p=2597303>

MIT Library GIS and Data Lab <http://news.mit.edu/2018/mit-gis-and-data-lab-now-open-rotch-library-1113>

⁵ Bryant, Rebecca, Brian Lavoie, and Constance Malpas. 2018. *Sourcing and Scaling University RDM Services*. The Realities of Research Data Management, Part 4. Dublin, OH: OCLC Research. <https://doi.org/10.25333/C3QW7M>.

collaborations, therefore, will continue to be an important part of the Library's approach. Two specific areas of future collaboration that should be explored, prioritized, and developed include:

- Support and services for Data Management Plans. The Library will work with the OVPR and Grants Hub to identify DMP support roles that would produce the best level of service and support to Iowa State researchers.
- Education and training efforts would benefit from Library and faculty collaborations that would allow the offering of more advanced and discipline specific learning opportunities. The Library brings planning and marketing expertise, learning spaces, and a user focused service approach to such collaborations. Such partnerships with faculty would provide a healthy cross-fertilization of data skills and interests.

In addition, the Library must continue existing collaborations while also developing new collaborations with other campus data stakeholders in the areas of technology (Office of the Chief Information Officer, Biology Information and Technology, Research IT, High Performance Computing), methodology support (Statistical Consulting, Center for Survey Statistics and Methodology), compliance (Office for Research Integrity, Office for Responsible Research, Office for Sponsored Programs Administration), legal (Office of University Counsel), data projects (Iowa State faculty, staff, and students; Ames Laboratory), and intellectual property (Office of Intellectual Property and Technology Transfer).

One example of work under a current collaboration that should continue is the streamlining and refinement of the data set disclosure workflow. Clear documentation of which office is responsible for each part of the form should be created as well as a re-evaluation of the use of Smartsheet for tracking and notifications. Special projects and the integration of DataShare into university workflows such as dissertation submissions must also be developed and supported in partnership with appropriate campus stakeholders.

Facilities and Technology

A dedicated space within the library, the Data Lab, will house the Library's new Research Data Services unit, including both staff and dedicated workstations with specialized software. This space will serve as a campus "hub" for research data services. This public facing space will also offer semi-private space to facilitate researcher and librarian data management consultations, and staff offices.