Research Data Service Campus Annual Report

2017-2018

Mission Statement

"The Research Data Service (RDS) provides the Illinois research community with the expertise, tools, and infrastructure necessary to manage and steward research data."

2017-2018 in Review

2017-2018 Goals

- Maintain an acceptable base level of service for core areas during periods of staff shortage
- Recruit excellent replacement staff
- Undertake a comprehensive evaluation of all aspects of the RDS to identify strengths, weakness, and provide holistic basis for future directions
- Continue to grow as national leaders who support improvement of the research ecosystem, specifically through data-related work and application of LIS skills

Major Activities and Accomplishments

Now in its fourth year, the RDS firmly established itself as a valuable resource on campus and around the country. This year was comprised of equal parts external growth and internal reflection and development.

At Illinois, awareness continued to grow among researchers that the RDS exists and is staffed by knowledgeable and efficient professionals skilled in managing data throughout its lifecycle. RDS staff were invited to share their knowledge and expertise in multiple conferences and as collaborators in the nation-wide Data Curation Network (DCN). That these forms of recognition came during a year of profound transition for the RDS, when staffing was either drastically reduced or new staff were being onboarded, speaks to the strong reputation for excellence that the service has established among researchers and collaborators. The service was able to largely maintain that excellence during a challenging year due to the collaborative nature of the Library, as many individuals generously covered understaffed services areas.

Staff departures necessarily reduced some activities normally performed on campus, particularly workshop creation and delivery. The RDS took this time to reflect upon and refine the duties carried out by staff members. By December 2017, the Data Curation Specialist role was successfully transitioned to the Research Data Specialist role, and two new Research Data Specialists were hired.

Reflection and refinement activities continued with the complete revision of the data management workshop series, driven by data collected from experienced workshop instructors. The Illinois Data Bank was also reviewed, with regard to both how researchers used it during its first year of availability and how researchers were directed to deposit their data for curation and publication. Suggestions for improvement in both aspects of the Illinois Data Bank were developed and are currently being implemented. These suggestions include clearer language and allowing researchers to make choices about when their datasets are curated.

The goal for all growth and reflection activities remains to make it easier for researchers to get the information they need to create well-managed data for their own projects and those of other researchers now and in the future. Through presentations and publications, RDS staff continue to strengthen and develop the overall network of data management practitioners, from subject specialists in the Library to faculty and staff at Illinois and at other universities across the globe. Through reflection and refinement activities, we affirm Illinois as a crucial node within that network. These two moves—outward expansion and deep establishment—happened simultaneously this year, and both make it possible to provide the best in services like data curation, data management workshops, and data management plan (DMP) reviews to researchers right here at home.

In 2017-2018, the RDS's major accomplishments include:

- **Maintained engagement** (at scaled-back pace) during a period of staff transitions
- **Requested and received an evaluation of services** by an external reviewer (the results were positive)
- **Refined staff titles** to move away from narrower focus on curation to include all aspects of research data management (Data Curation Specialist role became Research Data Specialist role)
- **Hired and onboarded** two Research Data Specialists
- **Reviewed and refined the data management workshop series** based upon feedback provided by the library staff who serve as the instructors of the workshops
- **Reviewed subject categories** in the Illinois Data Bank
- **Reviewed the Illinois Data Bank's first year** and added a new, proactive option (pre-review) for researchers when they are depositing datasets in the Illinois Data Bank
- **Actively participated in campus-wide data-related efforts** like the Illinois Data Science Initiative (iDSI)
- **Published several highly-accessed, peer-reviewed articles** describing the work in the Illinois RDS

Staffing

The RDS provides its employees with highly valuable skills that translate well to a variety of employment environments. As such, during fiscal year 2018, both Data Curation Specialists left their

positions to pursue other opportunities, in March and June 2017, respectively. In the middle of the fiscal year, the RDS successfully filled these two positions. In December 2017, Hoa Luong (MLS '17) and Ashley Hetrick (MLS '18) began their work at the RDS as Research Data Specialists. The Research Data Specialists possess excellent knowledge of both the RDS and of the campus as a whole. Hoa served as an Academic Hourly for RDS from May 2017-December 2017, when she began her new position within the RDS. Ashley worked for Technology Services in AV/IT support and communications before joining the RDS. The Research Data Specialists collaborate well to perform all of the essential functions within the RDS, such as data curation, DMP review, and workshop creation and delivery.

The new Research Data Specialists hit the ground running, offering consultations, creating and refining marketing and communications products and processes, IDB developments, curation, and workshop delivery and evaluation activities almost immediately. By the end of the fiscal year, both had already curated several datasets, reviewed multiple DMPs, designed and delivered several workshops and presentations, reviewed usage over the first year of the Illinois Data Bank and published this review, and participated in RDS-related committees and events within the RDS and the Library. Now that the onboarding process has been completed, the RDS looks forward to greeting 2018-2019 as a fully-staffed unit.

Evaluation and review

In July 2017, at the behest of the Director of the RDS, all aspects of the RDS were reviewed by an external reviewer. Dr. Marta Teperek, Research Data Facility Manager at the University of Cambridge, conducted the review. The results of the review were very positive, with Dr. Teperek stating that she was "overall extremely impressed with the work done by the RDS at the University of Illinois. The achievements of the RDS team are world-leading." Several aspects of the RDS with which the reviewer was particularly impressed include, but are not limited to: value for money; library innovation; and user-centered design. The reviewer also provided recommendations for improvement within the following areas: HR processes (particularly concerning the extended timeline for hiring staff); better defining scope of the RDS; and more effectively measuring the effectiveness of RDS services on assisting researchers with better managing their research data.

RDS also conducted an internal evaluation or review of two major components of its service offerings: the data management workshops in the Savvy Researcher workshop series and use of the Illinois Data Bank during its first full year of availability (2016-2017). In February 2018, RDS staff evaluated the perceived effectiveness of the data management workshop series from the perspectives of the instructors who taught and observed the workshops. From these data, RDS staff were able to determine that, while instructors valued the opportunity to teach the workshops, enjoyed interacting with participants, and believe in the importance of instruction in data management principles and practices, they believe several changes would make the workshops even more effective and enjoyable to attend and to teach. This feedback was collected, analyzed,

presented to instructors, and then applied to both the structure and content of the data management workshop series. Major changes include providing learning objectives and goals for all workshops, standardizing the structure of workshops across the series, and creating an Activity Bank to which instructors are invited to contribute and from which they are invited to borrow to better customize their workshops to their teaching styles.

Outreach

Data continues to be a profound interest of many members of the Illinois research community. Interdisciplinary collaborations like the iDSI have emerged to support and refine this interest. Along with being a "new" phenomenon, data remains a familiar part of research across the disciplines. Data management thus remains an integral part of all research reliant upon data.

As such, RDS staff have worked to help researchers situate data and data management in their research processes, and to make ease of use (in data management, documentation, and publishing) a central pillar of outreach activities.

The RDS was significantly under-staffed (operating, as the external reviewer noted, at approximately 50%) this year, and outreach efforts were necessarily limited. For instance, the number of outreach and training events offered by RDS staff dropped 65% between fiscal years 2017 and 2018. Even with limited staff, however, the RDS presented or co-presented over 30 workshops, seminars, and presentations across campus that were either wholly or partially dedicated to data management. Through these workshops, seminars, and presentations, RDS staff were able to engage with 466 members of the campus community.

Some kinds of RDS engagement experienced an uptick during this time; these include data sharing consultations (86% increase) and Illinois Data Bank datasets curation requests (700% increase). The increase in these categories suggest that researchers understand that we offer data publication and curation services, and that they value these services as either new or returning users.

Additionally, the RDS added a new category of engagement related to data publication and curation: pre-review. Please see more information on pre-review, and changes made to the Illinois Data Bank, in the Illinois Data Bank section.

See Table 1 below for the breakdown of coverage of campus research units (not including Savvy Researcher workshops) and Figure 1 for a breakdown by service type.
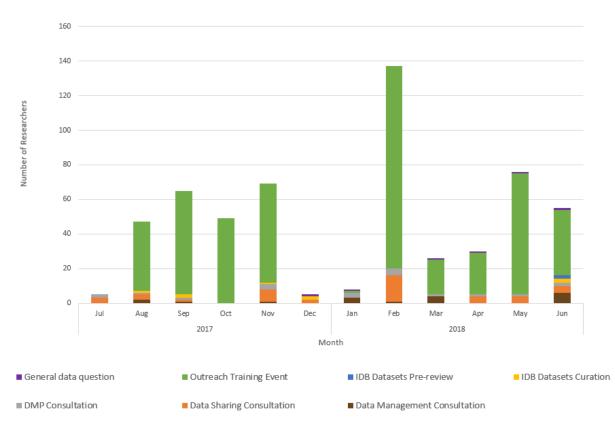
RESEARCH
DATA SERVICE
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

	Data Management Consultation	Data Sharing Consultation	DMP Consultation	General question*	IDB Datasets Curation	IDB Datasets Pre-review*	Outreach Training Event
ACES		✓	✓		✓	✓	
AHS	✓	✓	✓		✓		✓
Beckman		✓					✓
Blue Waters		✓					
Business	✓						
Education			✓				✓
Engineering	✓	✓	✓		✓		
Family Resiliency Center *	✓						
Grad college							✓
iSchool		✓		✓			
ISEE*			✓				
LAS	✓	✓	✓	✓	✓		✓
Law*		✓					
LER*			✓				
Library IT							✓
NCSA		✓	✓			✓	✓
OPD			✓				
PRI	✓	✓			✓		✓
University IT*							✓
University Library			✓				✓
Non-Illinois - External*				✓			
IFSI	✓			✓			

* new this year

Table 1. Coverage by outreach type for colleges (blue) and research institutes (green). Abbreviations: ACES, Agriculture, Consumer, and Environmental Sciences; AHS, Applied Health Sciences; LAS, Liberal Arts and Sciences; LER, School of Labor and Employment Relations; NCSA, National Center for Supercomputing Applications; PRI, Prairie Research Institute. Additional engagement was also carried out with the Illinois Fire Service Institute (IFSI), the Office of Sponsored Programs (OPD), and Institute for Sustainability, Energy, and Environment (ISEE), Family Resiliency Center (FRC) and people visited us from different universities.

RESEARCH DATA SERVICE
UNIVERSITY OF ILLINIOS AT URBANA-CHAMPAIGN

Figure 1. Breakdown of direct interactions by service type for each month in the FY2017-2018 year. Indirect interactions, such as DMPTool or Grainger DMP template uses, are not included.

Illinois Data Bank

In June 2018, the Illinois Data Bank passed an important symbolic milestone: it received its 100th dataset for ingest. RDS staff organized a celebration in order to recognize the people who made such a milestone possible: the researchers who contributed their datasets, the subject specialists who consulted on dataset development and nurtured relationships with researchers, and the IT support staff who designed and implemented the technologies upon which we all rely to create, share, and manage data. Several dozen people attended the event, including one of the authors of the second most downloaded dataset in the Illinois Data Bank.

Use of the Illinois Data Bank continues to grow. In the first year the Illinois Data Bank was available, over fifty researchers relied upon the Illinois Data Bank and accompanying services to curate and publish their datasets. Through the Illinois Data Bank, datasets produced by Illinois researchers are already enjoying more exposure. Preliminary analysis of download data from August 29, 2017 to

February 15, 2018 shows that download tallies of datasets published in the Illinois Data Bank are already at 6,646. In just the first half of its second year, the Illinois Data Bank is already reporting almost as many downloads as in its entire first year (74%).

RDS staff continue to refine the Illinois Data Bank—both the workflow that underlies deposit and curation and the interface through which researchers interact with the system—to make it even easier for researchers to deposit their best work. Initially, the workflow for depositing data in the Illinois Data Bank was *reactive*; researchers would deposit a finished dataset, curators would review and curate the data, curators would offer feedback to researchers, researchers would implement feedback (thereby necessitating a new version of the dataset) or not (the version would remain unchanged, and any errors would remain visible). To make this process more *proactive* (and thereby reduce the need to version), RDS staff worked together to meet the technical and communications requirements to offer researchers the option of having their datasets reviewed before the finished version was published in the Illinois Data Bank. At the time of this writing, two researchers have already selected pre-review.

See the *Illinois Data Bank First Year in Review* for more information about this service: https://www.ideals.illinois.edu/handle/2142/99650.

Contribution to Campus Data Initiatives

The RDS continued to both lead and support various campus-wide data-related initiatives in 2017-2018.

Director Heidi Imker has long contributed to a vision for a cohesive data science presence at Illinois, and that activity has increased this year. Director Imker has actively participated in the iDSI Steering Committee and Strategic Planning task force by collaborating with other stakeholders on the Initiative's core values, campus-level strategy, and campus-wide communications. Imker continues to work with the committee to define a vision which includes "data science for all" and what role libraries in particular can play in that vision. It is anticipated that the vision will be available for implementation next fiscal year.

Research Data Specialists Hoa Luong and Ashley Hetrick participated in several events and projects that served to further data management on and beyond the Urbana-Champaign campus. After attending the IT Professional's Forum (ITPF) in June, Hoa developed a relationship with IT staff in the College of ACES who have requested assistance in curating Morrow Plots data. That collaboration is on-going and is shaping up to be beneficial to researchers of the Morrow Plots for years to come. Ashley contributed to the Library's first HackCulture hack-a-thon as a mentor and partnered with the Mortenson Center and Illinois librarians to offer customized workshops on how data management works at Illinois and managing textual data to participating libraries in the IMLS-funded project

"Digging Deeper, Reaching Further: Librarians Empowering Users to Mine the HathiTrust Digital Library", respectively.

Ashley has also been working with Megan Ozeran, Data Visualization Librarian, to create and launch a three-part workshop series on data visualization. The series aims to help participants better understand and practice data visualization, including managing the data that make their visualizations possible. Skills taught include exploratory and explanatory approaches to data, visualization literacy, and Excel and Python. The new workshop series is scheduled to run in Fall 2018.

Feedback from 2017-2018 Services

DMP Consultation

From the Department of Food Science and Human Nutrition:
"Thank you for your quick and helpful feedback."

From the Social and Behavioral Sciences Research Initiative (SBSRI):
"Thanks so much for your helpful review, Ashley!"

From the Department of Civil and Environmental Engineering:
"This is super helpful. I greatly appreciate your fast and helpful response!!"

Data Management Training and Consultation

From the Mortenson Center for International Library Programs:
"We want to thank you for having been part of our 2018 Mortensen Center Associates Program. Our group of international librarians enjoyed your session and were inspired by your presentation to introduce new perspectives and practices at their organizations. We hope to continue to work with you, and to have the opportunity of collaborating on future Mortenson Center activities."

From the Department of Educational Psychology:
"The presentation on data management to the counseling psychology program faculty and students was OUTSTANDING. Personally, I was able to make immediate changes in how I store and organize my various files. I know a number of the students have implemented changes based on the information presented as well. The workshop was practical, informative, and extremely timely. Plus, we appreciated the handouts and goodies."

Illinois Data Bank

From the College of ACES:
"Thanks for going through things in so much detail and finding those errors!"

From the Department of Natural Resources and Environmental Sciences:
"Thanks a lot for your suggestions! They are very helpful and will make the dataset more easily understood."

From the Department of Water Resources Engineering and Science:
"Thank you very much. I didn't realize that someone would help me make my data more accessible; you've been a great help."

See the *Illinois Data Bank First Year in Review* for more feedback: https://www.ideals.illinois.edu/handle/2142/99650.

2017-2018 Leadership in the Profession

External Presentations

- Ilieve, Polina, Anderson, Bethany, **Imker, Heidi,** and Aimee Mederios "Collecting and Preserving Contemporary Science in the 21st Century" Panel, SAA, Portland, OR, July 28, 2017.
- Ringersma, Jacquelijn, **Imker, Heidi** "Policy Realities in Research Data Management" OCLC Works in Progress Webinar Series, July 25, 2017.
- **Imker, Heidi "**RDM Services in ARL Libraries: Research Data Service at the University of Illinois at Urbana-Champaign as a case study" National Institute of Informatics, Tokyo, Japan, September 25, 2017.
  Imker, Heidi "Research Data Management at the University of Illinois at Urbana-Champaign" iPres2017-WDS Collaborative Session, Kyoto University, Kyoto, Japan, September 28, 2017.
- **Imker, Heidi** "Cross-Campus Collaborations Enable More Trust-Worthy Data Sharing at the University of Illinois at Urbana-Champaign" World Data System Conference, Kyoto University, Kyoto, Japan, September 29, 2017.
- Malpas, Constance & **Heidi Imker** "The Realities of Research Data Management," CNI Spring Membership Meeting, San Diego, CA, April 13, 2018.

Written Works

- Johnston, Lisa R., Jacob Carlson, Cynthia Hudson-Vitale, **Heidi Imker,** Wendy Kozlowski, Robert Olendorf, Claire Stewart. "How Important is Data Curation? Gaps and Opportunities for Academic Libraries." *Journal of Librarianship and Scholarly Communication*, 6.1 (2018): eP2198.
- **Imker, H.J.** 2018. "25 Years of Molecular Biology Databases: A Study of Proliferation, Impact, and Maintenance." *Frontiers in Research Metrics and Analytics* 3. https://doi.org/10.3389/frma.2018.00018.
- **Hetrick, Ashley**, **Hoa Luong**, and **Heidi Imker**. "Illinois Data Bank First Year in Review." Urbana, IL: University of Illinois at Urbana-Champaign, 2018. http://hdl.handle.net/2142/99650.

2017-2018 for the RDS

Core Staffing

- Heidi Imker, Director, May 2014—Present
- Colleen Fallaw, Research Programmer, April 2015—Present
- Hoa Luong, Research Data Specialist, December 2017—Present
- Ashley Hetrick, Research Data Specialist, December 2017—Present
- Qian Zhang, CLIR Postdoc Fellow (w/ iSchool), August 2015—August 2017
- Dena Strong, ongoing AP development contribution, May 2016—Present
- Daria Orlowska, GA, August 2017—Present

Goals for 2018-2019

- Continue to contribute to leadership for the iDSI and thereby represent the RDS, the Library, and the diverse needs and groups encountered therein
- Continue to develop Illinois' reputation for and skills in data curation through co-leadership of the nation-wide Data Curation Network (DCN)
- Implement a research management training program for PIs and their graduate students
- Develop a professional relationship with the Center for Innovation in Teaching and Learning (CITL) by creating and delivering data management workshop(s) for the Grad Academy, a campus-wide orientation for graduate teaching assistants
- Collaborate with International Student and Scholar Services (ISSS) to conduct targeted outreach to international graduate students
- Evaluate Illinois Data Bank, including policies on deposit, preservation, and versioning