

Research Data Service Annual Report

2019-2020

“The Research Data Service (RDS) is a campus-wide program that provides the Illinois research community with the expertise, tools, and infrastructure necessary to manage and steward research data.”

2019-2020 in Review

2019-2020 Goals

- Goal 1: Refine and expand promotion around the Illinois Data Bank (SD1)
- Goal 2: Implement mission-critical feature developments for the Illinois Data Bank (SD1)
- Goal 3: Develop goals for and formalize participation in the Data Curation Network (DCN) (SD3)
- Goal 4: Collaborate with partners in the DCN and at Illinois to expand the network of subject specialists who curate datasets (SD1)
- Goal 5: Develop strategic relationships with groups and programs on campus dedicated to training their cohorts in data-oriented skills (including Training in Digital Methods for Humanists, the College of Business’ efforts in data science and analytics, the iSchool’s MS/IM program, and the Office of Proposal Development) (SD2)
- Goal 6: Create and offer instructional events that embed data management principles and best practices in data science and analytics tools and methods (SD2)
- Goal 7: Create and implement onboarding and training programs for unit GA and ARL fellow interested in data science and analytics (SD4)

Major Activities and Accomplishments

Overview

FY20 was an active and successful year. The RDS met last year’s goals, which centered on updating key service offerings and collaborating with others, both within and outside of Illinois, to promote data management principles and best practices on campus and at partner institutions. Major activities and accomplishments included:

- **Continued** broad marketing for the Illinois Data Bank and RDS (Goal 1)
- **Implemented** mission-critical feature developments for the Illinois Data Bank, including access to the draft dataset for co-author who has Illinois netID, and Globus, to continue supporting research at Illinois at the highest levels (Goal 2)
- **Strengthened** Illinois’ position as a leader in data management by formalizing expectations with the DCN, as well as providing examples of policy-, technical-, and content-development that serve as models for other institutions (Goal 3)

- **Expanded** collaborations with library subject specialists in engineering, humanities, and in the life and social sciences to provide domain knowledge during dataset curation (Goal 4)
- **Continued** contributions to data science and analytics initiatives on campus by offering several workshops that centered on the importance of data management principles and best practices to data-intensive projects (Goal 5)
- **Continued** collaborating with ISSS to conduct targeted outreach to international graduate students (Goal 6)
- **Celebrated** the one-year anniversary of the Data Interest Group, which is a monthly meeting that features researchers and support staff from within the University Library and across campus who share their projects and create space for collaboration (Goal 6)
- **Maintained** a professional relationship with CITL by being invited back to deliver a new iteration of a data management workshop for the Grad Academy, a campus-wide orientation for graduate teaching assistants (Goal 6)
- **Created** thorough and well-received training plans for unit graduate assistants and the ARL fellow (Goal 7)

Illinois Data Bank

Since going live in 2016, the Illinois Data Bank has been at the center of RDS’ service offerings. As Illinois’ institutional data repository, the Illinois Data Bank provides researchers with a free, stable, and convenient place to share finished research data with the world. To date there are **294** published datasets, including versions, which have amassed almost **140,000** downloads. Despite uncertainties caused by the pandemic as it escalated in the US, the Illinois Data Bank has maintained remarkably steady use in recent months (Figure 1 below).

As service providers who exist to support researchers, our services should and must change in response to these needs and solutions. One such need was large-file (GB and TB) transfer, and one such tool was Globus. As such, we are pleased to share that Globus has been introduced into the Illinois Data Bank environment, for researchers to use at their convenience. The feature has been adopted by several users.

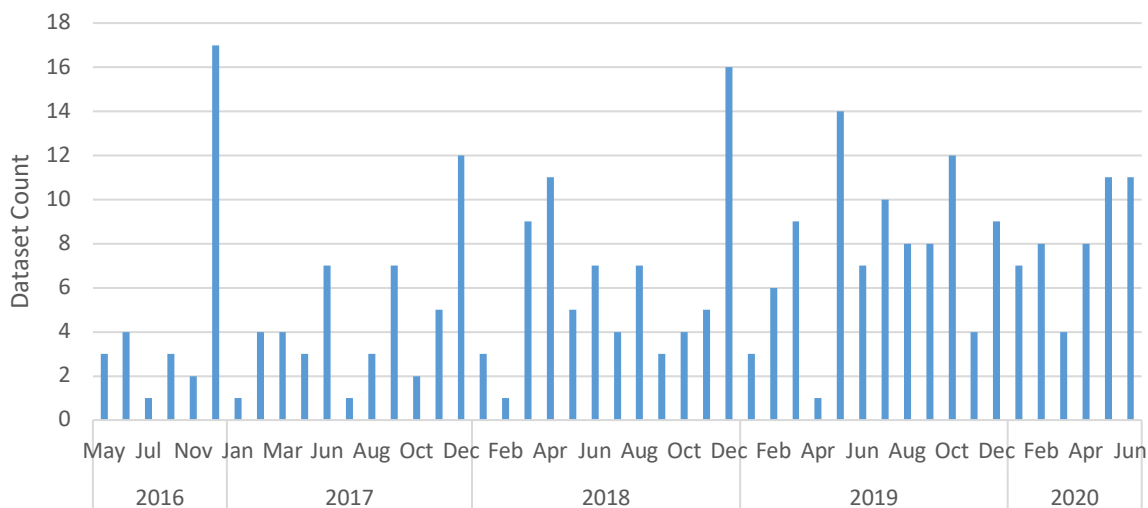


Figure 1: Number of published datasets in the Illinois Data Bank from May 2016 to June 2020

Data Curation Network (DCN)

FY20 could be said to be the year of data curation at the RDS. Along with making data sharing even more possible and convenient with the introduction of Globus into the Illinois Data Bank, RDS staff worked with members of the DCN to successfully develop goals for and formalize Illinois' participation in the DCN. The DCN is a cross-institutional data curation service currently comprised of ten (10) partner institutions across the nation. The Illinois RDS is a founding partner and has been active in the network from inception. The opportunity arose this year for Illinois to further lead by example by formalizing the outputs and timeframe that can be expected from RDS staff. Goals included submitting a certain number of datasets to the network for curation on a monthly basis, as well as curating a certain number of datasets submitted to the network from partner institutions. This formalized exchange of datasets among curators from varying institutions, backgrounds, and areas of domain expertise both improves the quality of research data overall and provides exposure to and training in data curation for a wider network of subject specialists. Additionally, RDS members participated in multiple DCN interest groups such as institutional outreach, marketing communication group, and sustainability planning.

5 Year Review of the Research Data Service

FY20 marked the fifth anniversary of the RDS. This milestone provided an opportune time for evaluation, and the RDS staff designed and executed a five-year review of the unit to gauge how effectively the RDS has met the needs of Illinois researchers. The review consisted of a survey and a series of in-depth focus group interviews. The survey was made available to the Urbana campus via campus-wide communication channels and was also circulated to subject specialists and other collaborators for additional distribution. Focus groups were comprised of researchers and collaborators across disciplines and campus departments, who volunteered to participate.

Results from survey and focus group interviews suggest that the RDS is an important resource on campus. The Illinois Data Bank is the most useful service among the suite of services offered by the RDS, following consultation, data management review, and workshop/training.

The following are research needs that RDS met:

- Data Management Plan review – fast, effective by enhancing the DMP
- Staff interactions with other partners and researchers – prompt and professional through email and in person
- Illinois Data Bank – user friendly interface, easy to deposit and use data. Data curation activities are considered to be helpful and an effective in way to help researchers enhance their dataset

By contrast, the following areas were identified as needs that have yet to be met:

- More educational workshops
- Present at different outreach events to increase the campus awareness
- More examples/templates for DMP, exemplar datasets and quality documentation would be handy to guide users

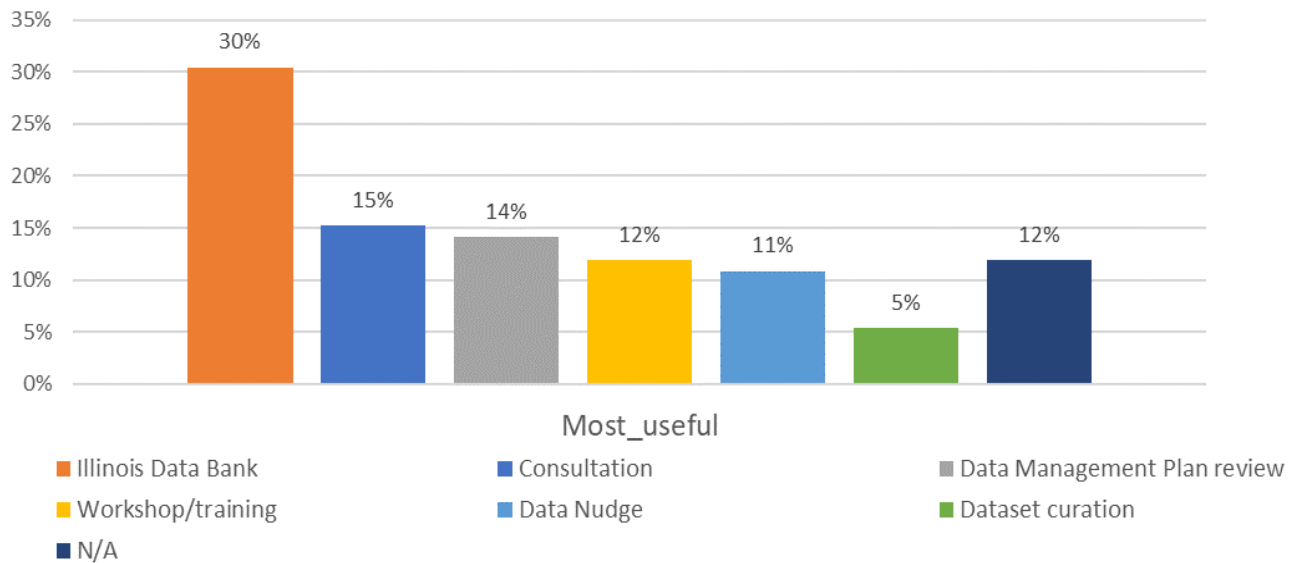


Figure 2: Most useful service offerings by RDS through the 5-year review survey

Education and Outreach

FY20 saw RDS staff again serving both individual users and groups of users at over **two hundred (200)** activities, including consultations, dataset curation, data management review, and workshops (Figure 3). Due to staff changing at the end of February and COVID-19, the number of users that RDS staff directly served dropped slightly for the year from over fifteen hundred (1500) users in FY19 to over **fourteen hundred (1400)** users in FY20.

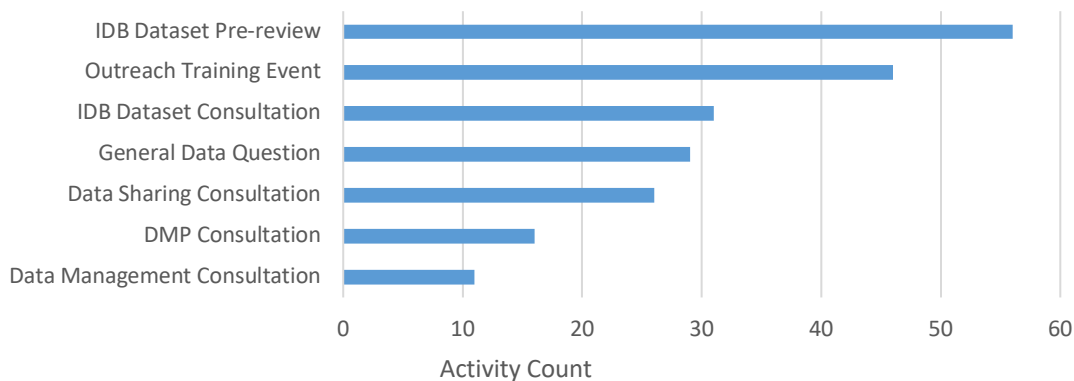


Figure 3: Breakdown of activities in FY20

In March, with the stay-at-home order from the Governor, all the services were offered virtually, mainly through Zoom. Outreach training events scheduled for in-person delivery in-person were postponed, which lead to a decrease in number of participants for the latter part of the spring semester (Figure 4).

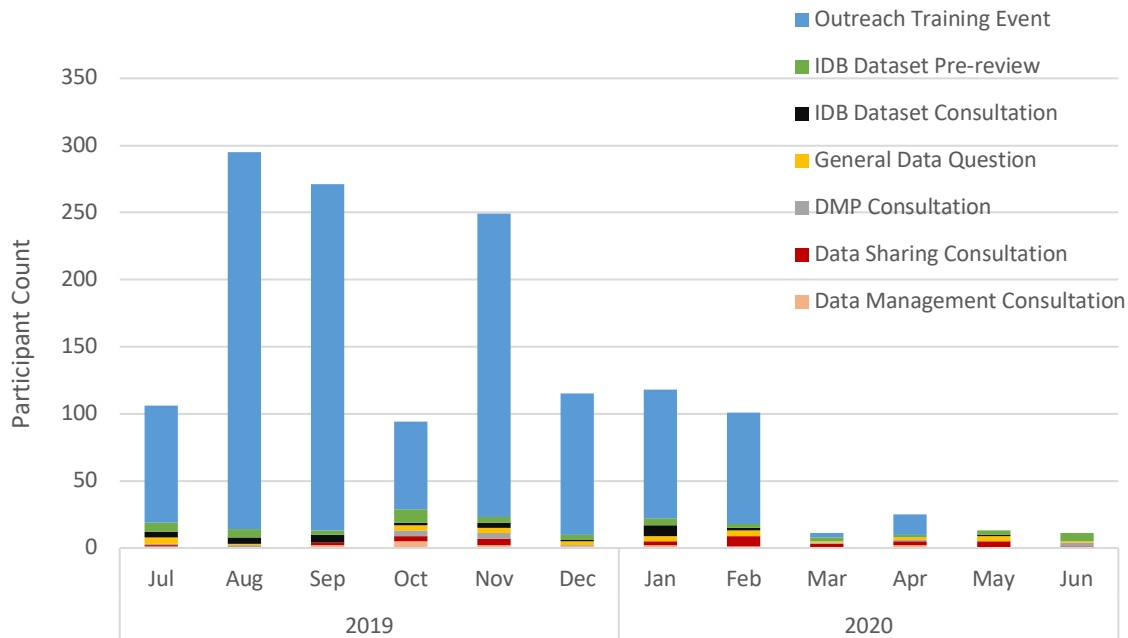


Figure 4: Monthly breakdown of user engagement in FY20

Training continued as a central activity for the RDS in FY20. Along with working with the DCN to help train subject specialists and new data curators in data curation principles and practices, the RDS broadened training opportunities provided at Illinois. RDS staff created and delivered several instructional events for a variety of specialized audiences, including a new cohort of teaching assistants (through the Center for Innovation in Teaching and Learning) and IT professionals interested in data management practices (through the IT Professional Forum). RDS staff also designed and implemented formal training programs for our graduate assistant and a practicum student, as well as an undergraduate student selected as an Academic Research Library (ARL) fellow. Both training plans share a focus on professional development through utilizing the resources and expertise available at the University Library and around campus.

Feedback from 2019-2020 Services

From the department of **Geography**:

“If there’s any evaluative forms or responses that I can fill out to show my appreciation for your service please send along and I will write about how helpful you’ve been!”

From the **Cline Center**:

“Thank you once again for your dedicated and detailed analysis of our data. This was very helpful as you caught some details we missed. We went ahead and made the revisions you suggested and then published the data. [...] we want to thank you very much for your help this past week by answering our questions and reviewing our data. Your insights were invaluable.”

From the **US Army Public Health Center**:

“I have gone through your Data Nudges and I love them. I have been tasked as the lead of a data initiative within the entire [department] to improve the current data infrastructure. I would love to hear more about you thought processes...”

From **NCSA**:

"Thank you both so much for your comments and edits. This is my first time writing a DMP for a project proposal, so I especially appreciate you sharing your knowledge and your time."

From the department of **Psychology**:

"Wow, great feedback [for DMP], thanks so much for making this so much less stressful."

From the department of **Biology**:

"Thanks for the super-fast reply and on a Sunday morning even! I'll get on the Lynda training and make notes on questions. Thank you so much!"

Major challenges

There were two main challenges that the RDS faced in FY20. First, staffing changed in February with the departure of the Assistant Director of Engagement and Education, Ashley Hetrick. Second, due to the COVID-19 pandemic, number of participants in outreach events dropped dramatically (Figure 4). A number of events were either canceled altogether or the course delivery method changed to asynchronous instruction. For example, previously scheduled in-person workshops morphed into providing the presentation materials for the students to review at their convenience. That this transition coincided with the departure of Ashley, who lead nearly all outreach and instructional sessions, turned out to be fortuitous. To address the shortfall in other areas that Ashley lead, Hoa Luong agreed to take on additional responsibilities and was promoted Associate Director of the RDS. This was especially critical given Heidi Imker's time continued to be constrained by her split RDS/Associate Dean position.

2020-2021 for the RDS


Core Staffing

- Heidi Imker, Director (1 FTE)
- Hoa Luong, Associate Director (1 FTE)
- *open* (1 FTE)
- Colleen Fallaw, Research Programmer (1 FTE)
- Lauren Phegley, Graduate Assistant (0.5 FTE)
- Dena Strong, Senior Information Design Specialist (0.1 FTE volunteer time from Tech Services)

Annual Goals for FY21

- Goal 1 - Expand on the 5 Year Review and implement key recommendations that surface
- Goal 2 - Articulate replacement position and complete hiring process
- Goal 3 - Engage with and build on existing diversity and inclusions efforts within unit, Library, and on campus
- Goal 4 - Implement mission-critical feature developments for the Illinois Data Bank
- Goal 5 - Continue to engage with organizations and communities of practice dedicated to advancing data stewardship, including the Data Curation Network
- Goal 6 - Strengthen existing and develop new outreach mechanisms to promote the tools and services provided by the RDS
- Goal 7 - Review and update service policies
- Goal 8 - Begin 5 year review of early dataset deposits

Appendix



RESEARCH DATA SERVICE
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Winter Snapshot

July 1 - December 31, 2019

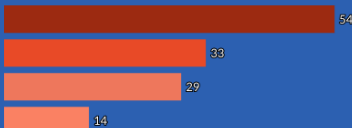
1130

Researchers Served

98

New Data Nudge Subscribers

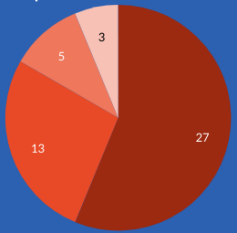
Service Breakdown



- Illinois Data Bank Curation (41.5%)
- Outreach Training Event (25.4%)
- Consultation (22.3%)
- General Data Question (10.8%)

"Thanks -- this was an illuminating introduction to what I need to learn."

Illinois Data Bank Discipline Breakdown




- Life Sciences (56.3%)
- Social Sciences (27.1%)
- Technology and Engineering (10.4%)
- Physical Sciences (6.3%)


22,250

Dataset Downloads


Top New Datasets Downloaded



509
Althaus
2019-08-30




443
Jackson
2019-09-01



317
Smith
2019-08-15

"Thank you once again for your dedicated and detailed analysis of our data."

powered by

RESEARCH DATA SERVICE
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Summer Snapshot

January 1 - June 30, 2020

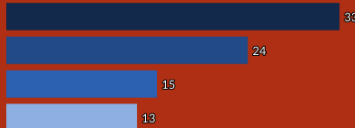
279

Researchers Served

30

New Data Nudge Subscribers

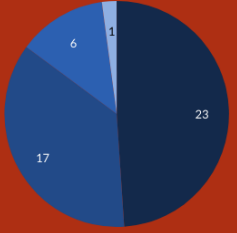
Service Breakdown



- Illinois Data Bank Consultation (38.8%)
- Consultation (28.2%)
- General Data Question (17.6%)
- Outreach Training Event (15.3%)

"Great feedback. Thanks for much this is so much less stressful."

Illinois Data Bank Discipline Breakdown




- Life Sciences (48.9%)
- Social Sciences (36.2%)
- Physical Sciences (12.8%)
- Technology and Engineering (2.1%)


28,672

Dataset Downloads


Top New Datasets Downloaded



253
Eick
2020-05-12



212
Cisneros
2020-01-27



183
Mishra
2020-05-15

"[...]it was great to have your support and it very vindicates our decision to use the Data Bank rather than off-campus data warehousing."

powered by
