



PROJECT MUSE®

Piloting the Purchase of Research Data Sets as Collections:
Navigating the Unknowns

Anita K. Foster, Amanda K. Rinehart, Gene R. Springs

portal: Libraries and the Academy, Volume 19, Number 2, April 2019, pp.
315-328 (Article)

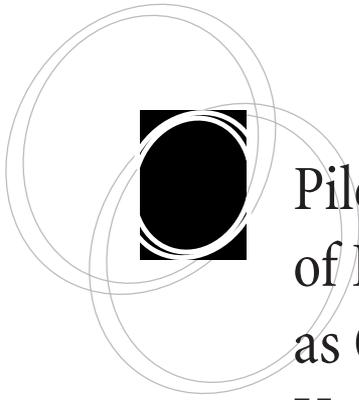
Published by Johns Hopkins University Press

DOI: <https://doi.org/10.1353/pla.2019.0018>



➔ *For additional information about this article*

<https://muse.jhu.edu/article/721427>



Piloting the Purchase of Research Data Sets as Collections: Navigating the Unknowns

Anita K. Foster, Amanda K. Rinehart, and
Gene R. Springs

abstract: In fiscal year 2017, The Ohio State University Libraries in Columbus piloted the purchase of research data sets to explore how to integrate this format into the standard workflows of the collections strategist and electronic resources officer. The pilot project had few restrictions except that one-time money must be used and purchases must be available to the entire campus community. This pilot enabled the libraries to purchase several large-scale data sets while exploring issues around licensing, technical deployment, support, and assessment in an organization without a data librarian. This article details the pilot process and outlines the resulting considerations for future purchases of data sets.

Introduction

The research data landscape continues to change rapidly. In 2017, such initiatives as the Data Refuge movement, a group based at Penn State University seeking to establish best practices for data rescue and preservation, dramatically heightened awareness of the fragile and complex nature of digital data.¹ The topic cuts a wide swath across academic libraries, as evidenced by the plethora of recent books about research data for and by librarians.² Libraries get many requests to locate data sets; to assist with federally required data management plans—that is, explanations of how researchers will handle data both during and after a research project; and to acquire complex data sets for high-level research endeavors. The demand for data services in libraries requires new resource allocations, technical infrastructure, policies, and strategic models.³

portal: Libraries and the Academy, Vol. 19, No. 2 (2019), pp. 315–328.

Copyright © 2019 by Johns Hopkins University Press, Baltimore, MD 21218.

Within this changing landscape, the evolution of research data services at The Ohio State University Libraries (OSUL) in Columbus has been demand-driven yet sporadic. OSUL librarians have explored and written about research data over the years, but the

The demand for data services in libraries requires new resource allocations, technical infrastructure, policies, and strategic models.

institution only recently committed to specific research data services. In 2007, interested librarians formed a Dataset Working Group to explore this emerging topic. This group led to OSUL's participation in 2011 in the Association of Research Libraries (ARL) E-Science Institute, a program to help libraries develop strategies to support the use of information technology in existing and new forms of research. That effort, in turn, led to the hiring of a data management services librarian at OSUL in 2014. At the same time, various librarians undertook independent investigations into the data practices of agricultural scientists, research data in university archives, and data trends in the health sciences.⁴ These and similar activities gave impetus to a new library initiative, the Research Commons. The Ohio State University Libraries established the Research Commons to advance high-level research at the university by providing services and space to help researchers explore collaborative, interdisciplinary, and emerging research methods and to connect with experts.

The Ohio State University Libraries established the Research Commons to advance high-level research at the university by providing services and space to help researchers explore collaborative, interdisciplinary, and emerging research methods and to connect with experts.

and emerging research methods and to connect with experts. The creation of the Research Commons led to the hiring of a half-time geospatial information specialist, responsible for the libraries' Geographical Information Systems (GIS) efforts.⁵

An initial task of the data management services librarian was to perform an environmental scan, a systematic survey of the organization's internal and external environment gathering information about things that might affect future operations, such as legal or technological changes. Interviews with teaching faculty and librarians revealed emerging needs for data services, such as data visualization, the location and use of publicly available data sets, access to local or institutional data sets, and the coordination of data set purchases. The environmental scan identified several recommendations: (1) continue to develop library faculty and staff interest in data management, (2) sponsor the creation of a campus-wide discussion group, (3) use the Research Commons as a sandbox, or an experimental space where researchers could discover new data services, and (4) clarify and coordinate existing data-related services in the library.⁶ The library administration approved the first and fourth recommendations, and two efforts were initiated, an in-reach program by the data management services librarian and the creation of a group called the Research Data in the Library Task Force.



The Research Data in the Library Task Force was charged with conducting an internal audit of current data-related services, articulating the challenges and complexities that prevented the offering of more support, and recommending solutions and next steps for delivering a more robust suite of services. The task force consisted of the data management services librarian; the digital humanities librarian; the Research Commons manager/GIS specialist; one representative each from University Archives, Publishing and Repository Services, and the Health Sciences Library; two subject librarians; and the electronic resources officer. This group had animated discussions that resulted in a detailed report of current services, the level at which the services were offered, where the assistance impacted the research life cycle, and which service areas to consider expanding. The task force also determined primary service owners and supporting units, which served to improve communication among task force members and began to sort out both authority and responsibility.

Finally, the task force categorized recommendations into short-, medium-, and long-term endeavors. In the short term, the task force recommended developing a workflow for purchasing data sets, listing purchased data in a central location, developing a team approach around data topics, and increasing campus awareness of open data tools—tools that can be freely used, shared, and built upon by anyone, anywhere, for any purpose. In the medium term, the task force recommended developing data mining guidelines for subscription resources and managing legacy data sets, primarily those ingested as part of other collections in the archives. For the long term, the task force suggested inventorying campus data services and centralizing information regarding purchased data sets and applicable licenses (regardless of whether the library had acquired them). The OSUL administration approved the task force recommendations, but by the time of publication, they had not yet been implemented.

After the task force identified the need for more clearly articulated data set purchasing processes, the libraries administration designated the collections strategist and the electronic resources officer to lead a pilot project to do just that. This pilot project represented a shift to a more adventurous and forward-looking outlook, which many experts consider necessary to become more innovative and experimental.⁷

As an academic research institution, OSUL has acquired data in various formats across a range of subjects and disciplines through a distributed engagement model. In this model, subject librarians instead of a single data librarian identify needs for data resources. Decision-making is then centralized with the collections strategist. Workflows may differ for institutions that charge a single data librarian with identifying or purchasing data set resources. This is the first report of a pilot project for data set purchases at an institution that operates on a distributed engagement model, coupled with central decision-making. The case study that follows includes a review of the literature that informed the purchasing program piloted in 2017, details the program's implementation, and outlines the lessons learned. It may provide a guide for other institutions navigating this complex landscape and making decisions about deploying their own resources.

Literature Review

The literature about purchasing research data sets in the academic library is scant but growing. Since 2016, authors from libraries at three different institutions of the Big Ten Academic Alliance, a consortium of the universities in the Big Ten athletic conference plus the University of Chicago, have written or presented case studies about programs developed specifically to purchase data sets. Assessing the library-funded Data Purchase Program that the University of Illinois at Urbana-Champaign implemented in 2010, Beth Sheehan and Karen Hogenboom describe both the process for faculty and students to request the purchase of small data sets and report findings from interviews with some of the applicants.⁸ The authors noted several challenges in acquiring data sets, especially the long wait from initial request to availability of the data set, if purchased. However, feedback from participants demonstrated appreciation for a program that addressed a gap in funding for researchers on campus. Librarians from the University of Michigan in Ann Arbor modeled their “one-time purchase” data grant program on the Data Purchase Program at the University of Illinois. They found similar complications related to the lengthy period required for purchase, noting that some data vendors were not accustomed to working with libraries. Among the benefits, they highlighted providing a central location for data licenses and preventing duplicate purchases of data sets.⁹ Anita Foster and Gene Springs detailed outcomes from a data purchase pilot program at The Ohio State University Libraries covering a single fiscal year, including modifications to acquisition workflows for data sets. Foster and Springs highlighted challenges in licensing these resources for the entire campus community.¹⁰

Karen Hogenboom and Michele Hayslett undertook the most comprehensive study to date in their survey and interviews of data librarians to learn how academic libraries

collect and manage small data sets. The authors organized their findings into three broad categories: (1) collection development, including roles, license provisions, policies, budget, and scope; (2) access, including mediated or unmediated access, notification of license requirements, and technology; and (3) assessment, detailing the local

Notable items from the “wish lists” of data librarians included making data sets more discoverable, developing more standardized workflows, and checking the quality of data from vendors.

practices of respondents and challenges inherent to collecting data sets in an academic library. Notable items from the “wish lists” of data librarians included making data sets more discoverable, developing more standardized workflows, and checking the quality of data from vendors.¹¹

The Pilot: Purchasing Research Data

OSUL launched a strategic collections initiative created by the vice provost and director to explore the acquisition of data sets for potential research use across campus during fiscal year 2017—that is, July 1, 2016, through June 30, 2017. OSUL administration as-



signed the initiative, designed as a pilot project, to the collections strategist and electronic resources officer. All data sets had to be purchased within that fiscal year, and the entire campus community must be allowed to use any purchased data sets. With these parameters, the collections strategist and electronic resources officer devised an exploratory and highly experimental plan.

As a first step, the collections strategist asked all OSUL collections managers (subject librarians and curators with collections responsibilities) for recommendations of data sets or data products for which they had received requests. This call yielded nine suggestions. Next, the collections strategist and electronic resources officer considered three data sets that had been previously investigated for potential acquisition. Additionally, a faculty member in the College of Education and Human Ecology (with 12 cosignatories from various academic departments and colleges across the university) contacted the collections strategist to propose licensing a data resource and purchasing the microdata associated with the resource. These means of identifying potential data sets for purchase yielded a total of 13 resources to explore, with seven eventually purchased within fiscal year 2017.

The purchased data sets represented a variety of disciplines, file sizes, formats, hosting availability, license requirements, discoverability, and cost. Four data sets were hosted via a publisher's platform; thus, they could be managed and made discoverable, much like traditional electronic resources.

However, three data sets—the largest and most expensive ones purchased—required hosting on locally managed servers and networks (see Table 1). To explore the feasibility of meeting the local server hosting requirements for the data sets, the collections strategist and electronic resources officer met with staff members from the OSUL Research Commons and Information Technology units to determine the availability of server space, technical support, and service support. The buy-in from staff in these areas became a critical component in moving forward with the purchase of the three data sets that required local hosting. Traditional, vendor-provided

Four data sets were hosted via a publisher's platform; thus, they could be managed and made discoverable, much like traditional electronic resources. However, three data sets—the largest and most expensive ones purchased—required hosting on locally managed servers and networks . . .

usage data are not collected to assess locally hosted data sets, so it was critical that the staff provide user support for, and local access to, these resources.

Since the purchase of these data sets, some vendors have changed the delivery or hosting options for their resources. These alterations illustrate the nascent nature of data set resources and the volatility of hosting and delivery. For example, the 2018 Collaborative Archive Data Research Environment (CADRE) project funded by the federal Institute of Museum and Library Services (IMLS) and led by Indiana University in Bloomington will develop a shared cloud-based infrastructure for cross-institution purchases of large data sets.¹² One of the first data sets intended for this system was the Web of Science XML data purchased on behalf of the Big Ten Academic Alliance—the same data set that

Table 1.
 Hosting requirements* of purchased data sets

Web hosted	Locally hosted
Gallup Analytics	Web of Science XML
BioCyc	Infogroup Business USA
Trucost	Gallup microdata
ICSD (Inorganic Crystal Structure Database)	

*Web hosting means the vendor provides a Web interface to access the data set. Locally hosted means that the data must be managed on local servers and networks.

OSUL, a member of the alliance, had bought during the pilot project before the alliance acquired it. Knowledge gained during the exploration and acquisition of data resources through the pilot program allowed the collections strategist to develop general criteria to aid in future data set purchase decisions (see Figure 1).

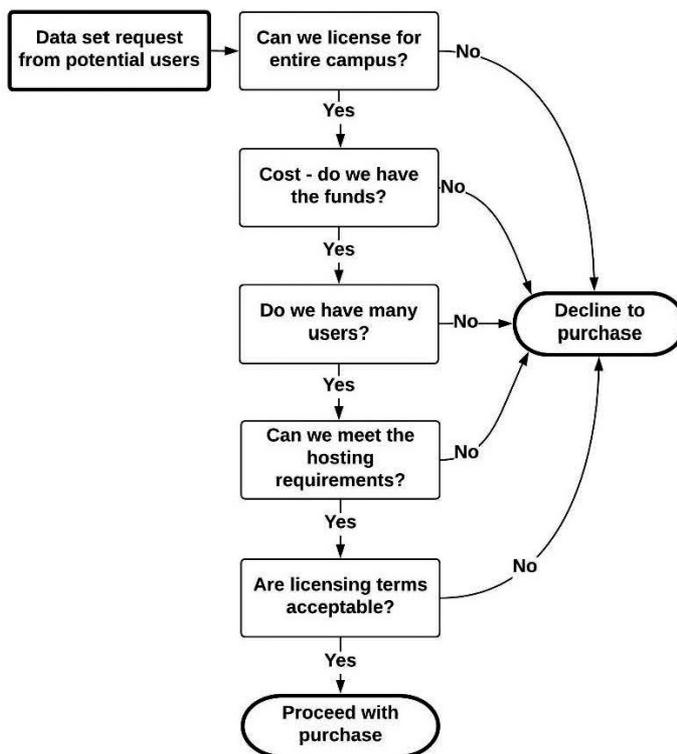


Figure 1. Decision tree for the purchase of data sets developed by The Ohio State University Libraries.



Although the pilot established no formal program for data set acquisitions, OSUL continues to explore user requests on a case-by-case basis using the criteria listed earlier. Because the time to thoroughly explore each factor varies, communication with the requestor is important. It is critical that requestors understand that, if the library pursues the potential acquisition of a data set, the library takes ownership of following and completing the process. If researchers have tight deadlines, the timelines for the library and the researchers may not always align.

It is critical that requestors understand that, if the library pursues the potential acquisition of a data set, the library takes ownership of following and completing the process.

Discussion

Licensing Data Sets

At first glance, licensing research data sets may seem to fit into standard processes used for licensing other types of resources. There are differences, however. How complicated the differences become depends on vendor familiarity with academic library procedures for licensing resources and how a data set might be used.

The licenses required for the data sets purchased during this pilot had unfamiliar and unexpected requirements, such as auditing and export control. Some of the agreements were also missing provisions that had to be added. The electronic resources officer consulted model licenses, which provide language identified as ideals and which can be adapted to local needs. Two model licenses consulted came from LIBLICENSE, a collaborative project hosted at the Center for Research Libraries, and from the California Digital Library, which purchases and manages electronic resources for the statewide University of California system.¹³ While the models consulted had provisions for text and data mining, none included guidelines for licensing research data or for auditing or export control. As academic libraries start to make such purchases more frequently, future updates to model licenses should begin to address these areas.

The licenses required for the data sets purchased during this pilot had unfamiliar and unexpected requirements, such as auditing and export control.

Auditing has many meanings for licensing resources and for business processes within libraries. Auditing, in this context, is a formal examination of records, typically by a representative from the vendor. There are three typical types of auditing language in licenses, two that apply to networks and one that pertains to financial records. The network auditing can include deletion of files when a subscription or partnership ends or investigations of possible breaches of license terms. While the ideal for academic libraries is to have no auditing language in a license, this language cannot always be

removed from the licenses for research data sets due to the unique nature of the material and how it may be used. Understanding the rules and expectations for auditing informs how best to proceed with a license negotiation. Auditing concerns and how they are managed may vary from institution to institution. There are several questions to consider:

- Do institutional regulations prohibit or allow auditing?
- What can be audited (financial records, other systems, servers, and the like)?
- What are the preferred methods for auditing, if allowed?
- When would auditing happen—at the end of the contract, at regular intervals throughout, only when misbehavior is suspected, or at other times?
- If outside people perform the audit, what sort of background check, if any, might be needed?
- Can the library provide documentation to serve the same purpose as auditing?
- What happens if the auditing has an unsatisfactory outcome?

There may be additional questions to consider, depending on institutional rules and regulations. Appropriate local resources (legal counsel, library administrators, purchasing departments, or other experts) should be consulted if a vendor insists on an auditing provision. It may be necessary to seek advice from people within a library organization, people external to it, or people at other institutions who previously licensed the resource.

Export control is another area that can be difficult to manage. More often found with federal government activities, export controls regulate the shipment or transfer of controlled items, software, technology, or services out of the United States.¹⁴ License terms regulating export control have arisen occasionally for other resources, more often to provide information about vendor activities with their products in other countries. For the data set purchasing pilot, one license had an export control clause that suspended negotiations for months while everyone tried to determine how to move forward. The institution could not take responsibility for the activities as the language was written but was satisfied if the vendor could monitor activity from its side. In the end, both parties agreed upon suitable language, and the negotiation progressed. When negotiating a license with export control language, questions like those for auditing may need to be answered. Familiarity with how an institution expects to use data with potential export control concerns—both vendor-created and user-created—will help the process.

While auditing was a common issue with licensing the various research data set resources and the export control issue took the most time, other unanticipated issues with licensing language arose. Libraries expect that they will be allowed to inform users of resources they have acquired, and licenses seldom include language limiting such communication. However, in multiple data set resources, a literal reading of the license terms suggested such announcements were not allowed. Why would that be? For some of the research data sets purchased, the type of data may have been a factor; for others, an unfamiliarity with academic libraries may have led to the difficulty. During the licensing process, OSUL added language to licenses allowing inclusion of the resources in standard library information pathways, such as a databases list, LibGuide, or library catalog. As OSUL had yet to determine how to market the data acquired, explicit arrangements for marketing campaigns were not included but would be managed as needed with vendors.



Some vendors were new to working with academic libraries or had not worked with academic libraries with this kind of resource. The lack of familiarity did not impact the ability to purchase the research data, but it did affect language around authentication and access, especially for users wanting to access a resource from off-campus. As the electronic resources officer described to vendors the methods used to provide off-campus access to appropriate, authorized users, suggested language was incorporated into the licenses as necessary. Additionally, The Ohio State University has a decentralized technology network on campus, which led to discussions with more than one vendor about how the libraries could manage expectations around data removal at the end of an agreement. It is common with restricted-use data, such as that from the U.S. Department of Education, to require deletion of data sets “and confirm that all restricted-use data have been wiped from the computer, [and] that all backup copies and any restricted-use data printouts have been destroyed.”¹⁵

In addition to the concerns about license terms discussed earlier, librarians negotiating licenses should be attentive and include details regarding the data delivery itself, either within the licenses themselves or as an amendment or exhibit. Delivery information should include:

- what will be contained in the data set delivery, especially when buying data that will be hosted locally (for example, raw data, data dictionaries, and teaching or training materials);
- frequency of delivery;
- how the data set will be delivered (for example, on an external hard drive or via File Transfer Protocol, a procedure for transfer of files over a computer network);
- the file sizes and formats of what will be delivered;
- contact information for the vendor department or person responsible for delivering the data to the library;
- the library personnel to whom the data will be delivered, with language that accommodates changing contacts as people switch jobs for both the vendor and the library;
- what happens if a vendor is purchased by another company, can no longer deliver the data in the agreed-upon manner, or either party needs to cancel the agreement.

Negotiating licenses for research data sets for this project required much learning by both library and vendor staff. Since negotiating licenses can take time, it is best to be patient: it is better to be clear on the terms under discussion than to rush and end up with language unfavorable for the library.

Access and Support

Access and support are two additional areas that warrant consideration when purchasing data sets. Unlike electronic journal, book, or database platforms, how users access the data sets may vary widely. Most Web-based data sets come packaged with relatively user-friendly interfaces and can include built-in tutorials, but data sets that require local hosting may consist solely of several hundred text or data files with no front-end user interface, leaving the purchasing libraries to develop their own. To provide access to the data set, the library may need to supply dedicated server space, develop authentication processes, manage user and usage tracking, and provide technical customization—all

of which may not be within the scope of the library or the campus information technology staff. Strategic discussions of the myriad potential needs for local development

Most Web-based data sets come packaged with relatively user-friendly interfaces and can include built-in tutorials, but data sets that require local hosting may consist solely of several hundred text or data files with no front-end user interface, leaving the purchasing libraries to develop their own.

and support are vital to provide optimal user access to, and discovery of, a locally hosted data set.

for Web-based data sets; however, none of these use standard assessment tools such as Project COUNTER (Counting Online Usage of Networked Electronic Resources), a nonprofit organization that develops uniform methods and reports for measuring the use of electronic resources. The locally hosted data sets required more creative approaches to assessment. Though staff can record transactions related to accessing these data (x number of requests to access a data set y number of times in z time interval), the size, scope, and complex file structures of the locally hosted data sets make additional quantitative analysis challenging.

Using qualitative methods to aid in assessment would be beneficial for all the purchased data sets, although the evaluations will be particularly cogent for the large data sets that require local hosting. Since post-purchase access has evolved for some of the

Options to evaluate purchased data sets may include interviews, focus groups, and surveys of library staff and campus researchers to better understand how these data impact research.

locally hosted data sets, available usage information may not be stable enough for a consistent, single mode of assessment. Options to evaluate purchased data sets may include interviews, focus groups, and surveys of library staff and campus researchers to better understand how these data impact research. A longer-term approach may include examination of the published literature by university researchers to analyze citations or track acknowledgements for the data sets. Determining the impact of these purchased data sets on research—using both quantitative and qualitative methods—is key for assessment.

An additional perspective included in overall assessment is the fiscal efficiency that the purchase of a data set may create, with the library's investment resulting in campus-wide access. Since most data set purchases by researchers limit access to an

Assessment

Demonstrating the value of investments made in library collections is increasingly essential. For this pilot project, determining strategic methods of assessment was not possible until late in the process due to the shifting understanding of requirements related to licensing, access, and support. Traditional quantitative assessment based on usage metrics is possible

locally hosted data sets, available usage information may not be stable enough for a consistent, single mode of assessment. Options to evaluate purchased data sets may include interviews, focus groups, and surveys of library staff and campus researchers to better understand how these data impact research. A longer-term approach may include examination of the published literature by university researchers to analyze



individual or small group, researchers may buy the same or similar data that others on campus have already purchased. When the library can provide access to a data set for the entire campus, it reduces potentially redundant purchases. In this pilot project, the library established an understanding with each vendor to refer inquiries from campus-affiliated researchers to the library.

Growing Pains

As libraries continue to invest in resources that support emerging fields, learning must occur both vertically and horizontally throughout the library organization. First, the lack of a specific directive to manage data requests creates disincentives—otherwise known as the “management gap,” the inability of library managers to understand and value the benefits of managing data requests and to provide organizational and managerial support.¹⁶ As a result, there may be confusion as to who is responsible for fulfilling a request, and some requests may fall between the cracks. Second, some librarians may see data as a fad, may suffer from change or “future” fatigue, or may be reluctant to develop services in this area.¹⁷ If the maxim of “doing more with less” has been applied, librarians may have to choose between traditional activities and emerging services. Third, librarians may lack the knowledge or confidence needed to facilitate the use of data sets.

Librarians may need support and training in this area, particularly if they lack a background in the research methods used in their discipline. Unlike text, which we learn to read at an early age, data sets may be numbers, shape files, or other unfamiliar materials. Both users and librarians can benefit from greater data literacy, an increased ability to derive meaningful information from data sets. Programs that support librarians in developing data literacy include Data Information Literacy, focused on data management and curation; the Institute for Research Design in Librarianship, which assists librarians in developing and completing their own research projects; and Data Carpentry, a community that teaches the use of software to accomplish data tasks.¹⁸ The suggestions in the IMLS-sponsored 2018 report “Shifting to Data Savvy: The Future of Data Science in Libraries” may also prove helpful.¹⁹

Conclusion

When buying research data sets, partnering with other interested parties—such as vendors, nonlibrary entities, or consortia—may be an effective route to success. Challenges exist at many levels, including purchasing, providing access, and supporting usage. Most hurdles, however, may be overcome (or significantly reduced) though partnering, as evidenced by the Big Ten Academic Alliance CADRE project. Since the conclusion of CADRE, the Center for Research Libraries has released a draft document detailing Global Data Licensing Terms and Specifications that will assist in future license negotiations.²⁰

When buying research data sets, partnering with other interested parties—such as vendors, nonlibrary entities, or consortia—may be an effective route to success.

As libraries incorporate research data services into their existing practices, it is necessary to provide venues for additional communication, so that these services can co-

Libraries should investigate a variety of user needs, work with institutional representatives on licensing, explore technical hosting arrangements, and prepare public services or subject librarians for new interactions.

evolve at similar rates across the library. Libraries should investigate a variety of user needs, work with institutional representatives on licensing, explore technical hosting arrangements, and prepare public services or subject librarians for new interactions. Additionally, librarians can create explicit, unified messaging that clarifies, both within the library and to users, what services are offered. For example, even though

the library may purchase data sets, the user may have to wait significantly longer than expected. Revealing these facts up front will reduce anxiety and confusion.

Since high-level research is often fraught with complex logistics, legal barriers, and administrative procedures, researchers may already understand the challenges inherent in this work. In fact, articulating that the initial forays into this area are pilot projects or experiments can help researchers empathize with the process and encourage them to provide relevant feedback. In most pioneering research areas, failure is understood to be a “necessary role in experimentation and key role in learning.”²¹ Being explicit about barriers, challenges, successes, and failures allows researchers to both relate to and gain a deeper understanding of library efforts.

This pilot project has resulted in successful access to several data sets, a deeper understanding of the challenges of purchasing them and facilitating their use, and development of both criteria and a framework for considering these requests in the future. As a result, the OSUL collections strategist has successfully fielded data set requests since the pilot period. Additionally, the pilot project facilitated a shift from conservative library procedures to a more risk-taking, innovative process. Overall, OSUL is better prepared to explore and support work with research data thanks to this pilot project.

Acknowledgment

We would like to gratefully acknowledge Leigh Bonds, the digital humanities librarian at The Ohio State Universities, for reviewing and editing our manuscript.

Anita K. Foster is the electronic resources officer and an assistant professor at The Ohio State University Libraries in Columbus; she may be reached by e-mail at: foster.1037@osu.edu.

Amanda K. Rinehart is the life sciences librarian and an assistant professor at The Ohio State University Libraries in Columbus; she may be reached by e-mail at: rinehart.64@osu.edu.

Gene R. Springs is the collections strategist and an assistant professor at The Ohio State University Libraries in Columbus; he may be reached by e-mail at: springs.14@osu.edu.



Notes

1. Margaret Janz, "Maintaining Access to Public Data Lessons from Data Refuge," *Against the Grain* 29, 6 (2017): 30.
2. Kristin Briney, *Data Management for Researchers: Organize, Maintain and Share Your Data for Research Success* (Exeter, U.K.: Pelagic, 2015).
3. Lisa R. Johnston, Jake R. Carlson, Patricia Hswe, Cynthia Hudson-Vitale, Heidi Imker, Wendy Kozlowski, Robert K. Olendorf, and Claire Stewart, "Data Curation Network: How Do We Compare? A Snapshot of Six Academic Library Institutions' Data Repository and Curation Services," *Journal of eScience Librarianship* 6, 1 (2017): e1102.
4. Florian Diekmann, "Data Practices of Agricultural Scientists: Results from an Exploratory Study," *Journal of Agricultural & Food Information* 13, 1 (2012): 14–34.
5. Craig Gibson and Meris Mandernach, "Reference Service at an Inflection Point: Transformations in Academic Libraries," in *Imagine, Innovate, Inspire: The Proceedings of the ACRL [Association of College and Research Libraries] 2013 Conference in Indianapolis, Indiana April 10–13, 2013*, ed. Dawn M. Mueller (Chicago: ACRL, 2013).
6. Juleah Swanson and Amanda K. Rinehart, "Data in Context: Using Case Studies to Generate a Common Understanding of Data in Academic Libraries," *Journal of Academic Librarianship* 42, 1 (2016): 97–101.
7. Gerald Beasley, Marwin Britto, Holly Ann Burt, Samuel "Scott" Hall, Amanda Rinehart, Lorelei Rutledge, Lis Pankl, and Catherine Soehner, "Vignettes on the Value of Failure," *Research Library Issues* 292 (2017): 10–15, <https://doi.org/10.29242/rli.292.3>.
8. Beth Sheehan and Karen Hogenboom, "Assessing a Patron-Driven, Library-Funded Data Purchase Program," *Journal of Academic Librarianship* 43, 1 (2017): 49–56.
9. Ramune K. Kubilius, "Best Practices for Building Data Collections," *Against the Grain* 29, 6 (2018): 63, <http://dx.doi.org/10.1016/j.acalib.2016.10.001>.
10. Anita K. Foster and Gene R. Springs, "Data Stories: Using Narratives to Reflect on a Data Purchase Pilot Program," *Serials Librarian* 74, 1–4 (2018), <https://dx.doi.org/10.1080/0361526X.2018.1427991>.
11. Karen Hogenboom and Michele Hayslett, "Pioneers in the Wild West: Managing Data Collections," *portal: Libraries and the Academy* 17, 2 (2017): 295–319, <https://doi.org/10.1353/pla.2017.0018>.
12. Indiana University Network Science Institute, "CADRE: Collaborative Archive Data Research Environment," 2017, <http://iuni.iu.edu/resources/cadre>.
13. Center for Research Libraries (CRL), *Liblicense Model License Agreement with Commentary*, <http://liblicense.crl.edu/wp-content/uploads/2015/05/modellicenseneu2014revmay2015.pdf>.
14. Office of Research Compliance, The Ohio State University, "Export Control," <http://orc.osu.edu/regulations-policies/exportcontrol/>.
15. National Center for Education Statistics, U.S. Department of Education, *Restricted-Use Data Procedures Manual* (Washington, DC: U.S. Department of Education, 2011).
16. Matt Burton, Liz Lyon, Chris Erdmann, and Bonnie Tijerina, "Shifting to Data Savvy: The Future of Data Science in Libraries," *D-Scholarship @ Pitt*, Institutional Repository of the University of Pittsburgh, 2018, <http://d-scholarship.pitt.edu/33891/>.
17. Meredith Schwartz, "Future Fatigue," *Library Journal* 141, 15 (2016): 44.
18. Institute of Museum and Library Services, "Data Information Literacy," <http://www.datainfoit.org/>, a joint project of the Purdue University Libraries, University of Minnesota Libraries, University of Oregon Libraries, and Cornell University Library; IRDL, Institute for Research Design in Librarianship, <http://irdlonline.org/>, a partnership among the William H. Hannon Library at Loyola Marymount University, San José State University School of Information, and the Statewide California Electronic Library Consortium; Data Carpentry, "Building Communities Teaching Universal Data Literacy," <http://www.datacarpentry.org/>.



19. Burton, Lyon, Erdmann, and Tijerina, "Shifting to Data Savvy."
20. CRL, "Global Data Licensing Terms and Specifications," 2018, <https://www.crl.edu/electronic-resources/tools-resources>.
21. Kaylyn Groves, "Editor's Note," *Research Library Issues* 292 (2017): 4-5, <https://doi.org/10.29242/rli.292.1>.