

Ithaka S+R Agriculture Research Support Services Study University of Illinois at Urbana-Champaign Report

**Sarah C. Williams and Erin E. Kerby
November 2016**

Introduction

The University of Illinois at Urbana-Champaign was one of nineteen institutions that participated in the Agriculture Research Support Services study organized by Ithaka S+R, a not-for-profit research and consulting service for libraries and academia. Part of Ithaka S+R's Research Support Services program, this study investigated the research support needs of scholars in the field of agriculture. Based on the strength of the agriculture programs at the University of Illinois, the University Library was invited to participate in the Agriculture Research Support Services study.

For the purposes of this project, Ithaka S+R was interested in many aspects of agricultural studies, including food, the environment, natural resources, and international issues. The project's definition of agriculture states, "We recognize that the field of agriculture is broad and that researchers fall on a disciplinary spectrum that encompasses the sciences, social sciences, economics and business. These researchers explore a wide variety of topics such as but not limited to: agronomy, molecular biology, informatics, remediation of soils for production on earth, cellulosic research for biofuels, nutrient enhancements in food, and the human dimensions of working with youth, rural and urban populations."

The University of Illinois' College of Agricultural, Consumer and Environmental Sciences (ACES) exemplifies the broad nature of agricultural studies. The College of ACES has seven departments – Agricultural & Biological Engineering, Agricultural & Consumer Economics, Animal Sciences, Crop Sciences, Food Science & Human Nutrition, Human Development & Family Studies, and Natural Resources & Environmental Sciences – and four other academic programs – Agricultural Communications, Agricultural Education, Nutritional Sciences, and a Professional Science Masters Program. Across these departments and programs, the College of ACES has around 180 tenure-system faculty, 2,700 undergraduate students, and 730 graduate students.

This report is the initial research product of the study and findings at the University of Illinois. The research team included the Life Sciences Data Services Librarian and the Veterinary Medicine Librarian. We conducted semi-structured interviews with College of ACES faculty to learn about their research methods and practices and their research support needs. This report includes a description of the methods used for the study and discussion of three overarching themes that emerged from the interviews. It concludes with initial thoughts on how the interview findings can inform current and future library services and initiatives.

Methods

As part of the larger, multi-institutional Ithaka S+R Agriculture Research Support Services project, this study utilized some materials provided centrally. Ithaka S+R provided an Institutional Review Board (IRB) guide that included information and language that we modified and adjusted to complete our IRB application for the University of Illinois at Urbana-Champaign. Ithaka S+R also provided the semi-structured interview questions (Appendix), so our questions were consistent with those asked at the other participating institutions.

This study's target population was agriculture scholars who (1) are tenure-system faculty or have active research programs and (2) conduct research aligned with the project's broad definition of agriculture. Therefore, we focused our attention on the tenure-system faculty in University of Illinois' College of ACES. For each of the seven departments in the college, we reviewed faculty directory webpages and individual faculty webpages to identify scholars with active research programs. We identified 189 potential interviewees and compiled in a spreadsheet their basic information, including name, email address, department, and a very brief description of one to two research areas.

Using the purposive sampling approach, we selected 63 faculty members to invite to participate in the study. Our main selection criteria included faculty with very active research programs, established relationships with librarians, and/or service on the college or campus senate library committees. While the larger Ithaka study did not require a sample with a mix of faculty ranks or gender, we contacted a variety of individuals to try to gather diverse perspectives. Twelve faculty members agreed to participate, which included at least one faculty member from each of the seven departments. This small non-representative sample should not be used to make generalizations about the entire population.

We communicated with each faculty member directly. In April 2016, we emailed the first invitation to participate. The initial response was very low, which was not particularly surprising given the timing near the end of the spring semester. Consequently, we requested support from the Dean of the College of ACES, who may have encouraged some invitees to participate. In early May 2016, we sent a follow-up email to every faculty member who had not responded, which resulted in each of us scheduling six one-on-one interviews. The interviews were conducted in May and June in each faculty member's office or laboratory, and they were recorded to allow transcripts to be produced. An external transcription service provided the twelve transcripts as Microsoft Word documents. Upon receipt, we each reviewed the transcripts for our six interviews to redact any identifiable information (e.g., names, distinctive publication titles) and to fill any gaps.

After anonymizing the transcripts, we read them again and coded them using the comments function in Microsoft Word. We used open codes, rather than pre-defined codes, to allow ideas and themes to reveal themselves organically. Next we grouped and organized the many open codes into categories and sub-categories. For example, within the topic of scholarly dissemination, one category was choosing where to publish, with sub-categories for impact factor or prestige, audience, and publisher turnaround time or production. Some codes were

included in more than one category or sub-category. Only interview number, not interviewee name, linked the codes to the transcripts.

To identify major themes, we individually reviewed the codes in and across all of the categories and sub-categories and then met to discuss our ideas. The codes revealed more themes than could be addressed in this initial, local report, so we focused on major themes that each had significant sub-elements and that were more likely to have a connection to libraries and library services. We decided on three main themes: research data, information sources and management, and communication, which are addressed in the next section.

Themes

Research Data

Despite the scope of research data, three important topics emerged from the interviews. Data from external sources is a critical component of some research projects, but researchers often face challenges in acquiring data from external sources. Obtaining data from individual researchers, even other researchers at the University of Illinois, can be particularly frustrating. Interviewees also described their data management practices, and in some cases, explained how their practices have changed over time, as a result of funding levels, data quantity, or past problems. Data sharing via repositories emerged as another significant topic, although interviewees also discussed other notable data sharing avenues (e.g., supplementary materials, by request).

Some interviewees rely heavily on data from external sources, particularly those who develop and experiment with models. Those sources include federal agencies (e.g., U.S. Geological Survey), the state surveys (e.g., Illinois State Water Survey), and individual researchers. In writing extensive review articles, one interviewee discovered “the literature truly is a vast resource, and it gives countless datasets from field work in the past.” The interviewee makes a point of using datasets, not discussions, from the literature, because the interviewee has found that authors occasionally tell a different story than their data. Other interviewees mentioned weather data in particular as a data type they obtain from external sources, especially the National Oceanic and Atmospheric Administration (NOAA).

Interviewees discussed the challenges of acquiring data from external sources, particularly from individual researchers. Contact people for datasets “are not as responsive as we’d like them to be,” or interviewees are referred from person to person – “we’re going around the circle, and we’re not getting any data.” One interviewee has encountered researchers who are unwilling to share their data because it will take too much time to prepare; the interviewee has even offered to organize the data, if the researchers will just provide it. Another interviewee has experienced awkward situations when requesting data from other University of Illinois researchers who are not eager to share their data. These situations are challenging in large part because data sharing protocols are vague; it is not clear if or when University of Illinois researchers should be expected to share their data. In some cases when local data requests have been unsuccessful, the interviewee has attempted to reproduce the data, which requires additional time and resources, or has broadened a project beyond Illinois, which reduces the local impact of the research. The challenges are so large that when asked how a magic wand could help the research and

publication process, one interviewee asked without hesitation for the wand to find and acquire the data needed for models being run.

Research data management was another data topic addressed in multiple interviews. Interviewees mentioned tools they use for their data, including Microsoft Excel, databases, and REDCap. Most seemed satisfied with their current tools, although some limitations were acknowledged. One interviewee noted that while Excel is easy to use, it is also easy to make mistakes with data in Excel.

Some interviewees' data management practices have changed over time. One interviewee commented that the research group does less database programming, because "there's no sort of continuous funding to maintain [databases]" and "it just has become unreasonable to save everything in databases." An increased quantity of data changed the data practices of another interviewee. In the past, streamflow and rainfall data were measured daily, but now they are measured every five minutes. When files contain data measured every five minutes, they are so large that "you cannot even open that in Excel." Now a big part of the interviewee's research is developing tools to process, visualize and manage all of the data. The data management practices of another interviewee evolved as a result of a problem within the research group. As an early-career faculty member, the interviewee had fewer controls on student data, and "it was a disaster." One student did not understand the research process, manage the data correctly, or apply the right equations; it took months for the interviewee to clean and re-organize the data. Now the interviewee stresses the importance of data management, sets file access permissions for different researcher categories (e.g., undergraduate student, doctoral student, laboratory technician), and reviews the students' data regularly. The interviewee commented, "little by little, we start to put all of those controls in place, and I think that's been crucial."

Interviewees also described their experiences with and thoughts on data sharing via repositories. Several interviewees have deposited data and code into repositories, including National Center for Biotechnology Information (NCBI) databases, Dryad Digital Repository, and IDEALS (University of Illinois' institutional repository). In many cases, their motivation was driven by funding agency or journal publisher requirements, while some interviewees deposited data because it is expected in their field. Some early-career interviewees had not yet deposited data, because they had not yet completed a funded project, but they were aware that they would likely need to deposit data in the future.

Interviewees expressed mixed opinions about repositories. One interviewee said it seemed reasonable to deposit data in NCBI databases, because "we benefit a great deal from other peoples' data in there." Yet the interviewee was not convinced about the benefit of repositories like Dryad because they have never used data in Dryad and are not sure anyone has ever viewed the materials they deposited there. Among interviewees with no experience sharing data via repositories, several responded positively to the idea. In particular, one interviewee was open to the idea but stressed that the repository must be user-friendly and provide user support. Several interviewees voiced concerns about repositories. The longevity of repositories was one concern, with one interviewee wondering "how permanent is IDEALS?" and another discouraged by the transient nature of code repositories. Mirroring concerns well documented in the literature, other

concerns included potential misinterpretation of data, the competitiveness of research, privacy of research subjects, and the time required to learn a new system.

Within these three research data topics, the interviewees identified areas where librarians and the library could be involved. One interviewee commented that finding publicly available datasets is an important skill for students and “an area where the students aren’t using the research librarians enough.” Consequently, the interviewee makes a point of referring students to the library. Another interviewee who struggles with obtaining external data suggested “it would be nice if the library would have the facility that will ask [University of Illinois researchers for their data] and then we can have data that we will need.” As noted earlier, the interviewees had mixed opinions about data sharing via repositories, but one hesitant interviewee “would feel more comfortable” depositing data with the library, as opposed to “these nebulous things outside of our control.” Based on previous experience turning to the library for data management guidance, one interviewee said that the library is also well situated to point researchers to other campus data services and infrastructures.

Information Sources and Management

When it comes to information sources and management, the interviewees described many different starting points and approaches to searching the scholarly literature, accessing and managing that material, and keeping current with trends in their respective fields. For many, these activities are blended together and do not occur independently of each other. Some commented that the Internet has improved all aspects of finding, accessing, and managing information, and that they rarely go to the library in person any longer. Nonetheless, time remains a barrier for the interviewees, both in terms of finding enough time to do these tasks but also because technology has not yet made them seamless.

Approaches to searching for information varied considerably, depending both on the purpose and preferred search tools. Some interviewees are more methodical than others; one creates a bibliography, reads the abstracts, keeps the useful articles and then discards the rest. Others are less strategic and have a messier process, which one interviewee saw as problematic, “My process is terrible. It is really terrible.” Many of the interviewees use keyword searches, whether browsing online or doing a targeted search in a library database. Other approaches included limiting database searches to the most recent year or to specific article types, and mining the references of relevant journal articles.

Most of the interviewees seemed aware of library-provided resources, and some indicated they routinely use databases such as Web of Science, CAB Abstracts, and PubMed. One interviewee commented that they “find everything needed in the library,” however, another said it is very time consuming to go through the library because there are so many clicks. Many interviewees said they use Google Scholar, either as their main search tool or in addition to others. They find it easy to use because it does not require special techniques or knowledge and gets them to the full-text very quickly. Overall, there was no indication that any single search tool is the best for every researcher all the time.

Citation management software came up a number of times during the interviews. EndNote, RefWorks, and RefManager were all mentioned, but use of these types of tools appears to be

very individualized. Some use it to create bibliographies, some use it to collect PDFs, and others use it to share references with others. One interviewee prefers to simply go back to the Internet to find or access an article. In general, this speaks to the individualized nature of the research process, where the librarian's role can be to identify options and help select the best one."

Communication

We had wide-ranging conversations about communication issues, but three significant topics emerged. Given the importance of scholarly publications in communicating research, interviewees explained how they choose the best journal for their research, and while the deciding factors are not particularly surprising, the perspectives of individual researchers were interesting to hear. In addition to scholarly publications, communication with the public emerged from the interviews as a significant topic, even though the interview questions did not specifically address this. Lastly, while external communication to other scholars and the public is important, internal communication among collaborators and within research groups is also critical.

Scholarly publications are the main avenue for agricultural researchers to communicate their research, and the interviewees shared many insights into how they choose where to publish. The three main deciding factors are: impact factor or prestige, audience, and publisher turnaround or production. In some agricultural disciplines, professional society publications have a low impact factor but are prestigious and/or reach a key audience. As some interviewees explained, this means they must strike a balance between these factors. With regards to audience, interviewees consider whether publications reach a specific target, such as a local audience, an audience "that's genuinely going to be interested and hopefully going to change their behavior or ... try and build upon what we've done," or a clinical audience in order to have a clinical impact. Some interviewees argued that agricultural researchers should more aggressively pursue top journals outside of agriculture in order to reach a broader audience. For example, one interviewee commented, "the top-flight econ journals are very much of interest to me because I want to reach an audience that's not necessarily thinking about agriculture but is interested in these environmental things." Regarding publisher turnaround or production, some interviewees spoke highly of traditional academic publishers, citing "quality work in fantastic time" and valuable pre-publication support. Yet one interviewee described a society's publication process as "outdated," and another shared examples of how the review and publication process of many open access publishers has declined.

Communication with the public also emerged as a significant topic from the interviews. For one interviewee in particular, communication with the public is a personal passion; the interviewee spends "a lot of time translating scientific literature for public consumption" and thinks of it as an extension of teaching. Other interviewees emphasized that this is an important role for the University of Illinois – "[the institution] was founded to give a neutral, independent, objective source of information that would benefit the public" – and the College of ACES – "That's got to be our mission in the College of Agriculture to help people understand: What do we know? What don't we know?" To them, this is particularly critical for agriculture, because it is a challenge for farmers to get information they can trust. Extension plays a role in communicating with the public (one interviewee has an Extension appointment and others have collaborators with Extension appointments), but interviewees noted Extension's limitations, such as dwindling

county agent levels, inefficiencies in the county agent model, and corporate influence on county agents. With a different, but related, perspective, one interviewee described the importance of communicating research to the public via the departmental and college outlets, including magazines and social media. To this interviewee, the departmental and college support for this sort of communication is strong.

Interviewees also commented on the importance of communication among collaborators in different disciplines and with students in research groups. With interdisciplinary research, effective communication among collaborators can be a challenge. To overcome this, one interviewee mentioned, “developing our own language, you might say, or setting up semantics so that we can effectively communicate in this interdisciplinary group.” Interviewees also described the significant time commitment required to communicate with students in their research groups and help them develop into skilled and ethical professionals. One interviewee regularly communicates to students about the importance of data management. Another makes a point to work with students as they prepare their research for publication because “before those follow-ups occurred, I think we saw publishable research not getting published on the first try.” With other research products, such as code or models, faculty and students may have different priorities – i.e., faculty want the products prepared to share with or pass down to others, while students are focused on “getting a degree and moving on.” This requires a “negotiation between us the lab directors and the students who are creating those products.”

As illustrated by these three topics, the research process requires effective communication at all stages and in many forms. Interviewees did not mention specific roles for librarians or libraries in any of these cases, but the communication challenges and expectations expressed by the interviewees can certainly help inform future library services.

Conclusion

This report is the initial output of University of Illinois’ Agriculture Research Support Services study, and as such, contains our initial thoughts on how the interview findings can inform current and future library services and initiatives. We will continue to reflect on the findings in these three main themes, as well as other interesting themes that emerged from our interviews. We anticipate that the University of Illinois findings and the cross-institutional report produced by Ithaka S+R will inform our library services and initiatives for the foreseeable future. Related to these three major themes, we identified several areas in which the University of Illinois Library, especially the Funk ACES Library (which serves the College of ACES), could use the findings to inform current and future services and initiatives.

The possibilities were most explicit with data. Interviewees saw a role for librarians in helping them and their students find and acquire data from external sources, whether publicly available or held by individual University of Illinois researchers. While finding and acquiring data is akin to library reference service, data has extra layers of complication as compared to traditional publications. Traditional publications have consistent metadata elements and robust, well-established discovery tools, such as library databases and Google Scholar; but data is very disparate, and the tools for finding data, even publicly available data, are often siloed and not well known. Despite the challenges, University of Illinois librarians and reference staff already

help patrons find publicly available data, but to increase awareness of this service, the Funk ACES Library could certainly work to promote this service specifically. At the same time, we would need to ensure that our librarians and reference staff are well trained to respond to these requests.

Requests to help acquire data from other University of Illinois researchers are even trickier. Without a clear campus or university data sharing policy, librarians would likely be in the same situation as the interviewees – awkward conversations with researchers unwilling to share their data and no data sharing policy to backup the request. The new University of Illinois data repository, the Illinois Data Bank (<https://databank.illinois.edu/>), could help address this issue. At least one interviewee who was hesitant to share data publicly expressed more willingness to deposit data with the library than an external repository. If other University of Illinois researchers feel the same, the Illinois Data Bank could become a platform for data sharing across campus because researchers might be more likely to deposit data in a repository that directly benefits them.

The Library certainly plays a role within the information sources and management theme because it develops or subscribes to many of the information tools. The indication that no single search tool is the best for every researcher all the time was of great interest to us. It suggests that librarians should focus on raising researcher awareness of information tools and helping researchers use them effectively. Based on our interview experiences, librarians should approach these conversations as an advisor and not be judgmental of researchers' current search and management practices. The interviews also revealed the difficulties that researchers face in finding and accessing scholarly literature (e.g., full-text access is not seamless or requires too many clicks). With this in mind, one priority for the Library should be continuous improvement of information tools. While the developers of University of Illinois' Easy Search tool do continuously gather user feedback, study usage and implement enhancements, this is just one of the Library's tools. The Library should make similar efforts with other tools, like our Discover (SFX) tool, and the Funk ACES Library should also strive to continuously improve its website and research guides.

As noted in the communication theme, the interviewees did not mention specific roles for librarians or libraries, but the findings can still help inform future library services. In particular, interviewees described the passion, interest and obligation they have to communicate their research to the public. In the future, the Funk ACES Library could further explore ways to support University of Illinois researchers and Extension in communicating research to the public. One possibility could be to provide library instruction and services to Extension personnel, as some land-grant libraries have done, so that Extension personnel can more effectively find information to meet the needs of the public. All options would need to be carefully considered, but this could lead to exciting opportunities for the Funk ACES Library.

Appendix

Semi-Structured Interview Guide

Research focus

1. Describe your current research focus and how this focus is situated within the broader agriculture discipline and the academy more broadly. [Probe for whether/not they see themselves as located firmly within agriculture as a discipline or located across/between disciplines]

Research methods

2. What research methods do you currently use to conduct your research?
3. What kinds of data does your research typically elicit?
4. How do you locate the primary and/or secondary source materials you use in your research?
5. Think back to a past or ongoing research project where you faced challenges in the process of conducting the research.
 - a. Describe these challenges.
 - b. What could have been done to mitigate these challenges?
6. How do you keep up with trends in your field more broadly?

Dissemination Practices

7. Where do you typically publish your research in terms of the kinds of publications and disciplines? How do your publishing practices relate to those typical to your discipline?
8. Have you ever deposited your data or final research products in a repository?
 - a. If so, which repositories and what has been your motivations for depositing? (i.e. required, for sharing, investment in open access principles)
 - b. If no, why not?

Future and State of the Field

9. What future challenges and opportunities do you see for the broader field of agriculture?
10. If I gave you a magic wand that could help you with your research and publication process – what would you ask it to do?

Follow-up

11. Is there anything else about your experiences as a scholar of agriculture and/or the agriculture discipline that you think it is important for me to know that was not covered in the previous questions?

Research Team Bios

Sarah C. Williams is Interim Head of the Funk ACES Library and the Life Sciences Data Services Librarian. With these dual roles, her responsibilities include management of the Funk ACES Library and outreach to the College of ACES, as well as data management training and consultations for life scientists. As an Associate Professor in the University Library, her research concentrates on data practices in the life sciences and services that can facilitate better data practices. Sarah's education includes a M.S. in Information Systems from Illinois State University, a Master of Library Science from Indiana University, and a B.S. in Soil and Crop Science from Purdue University.

Erin E. Kerby is Head of the Veterinary Medicine Library and an Assistant Professor with the University Library. In addition to managing the Veterinary Medicine Library and its collection, Erin provides reference, instruction, and outreach services to the College of Veterinary Medicine and related programs. Her current research focuses on the flow of information between veterinary researchers, and research data practices and management in particular. Erin received her M.S. in Information from the University of Michigan and worked there at the Taubman Health Sciences Library for several years.