Journal of eScience Librarianship
putting the pieces together: theory and practice

Full-Length Paper

Differences in the Data Practices, Challenges, and Future Needs of Graduate Students and Faculty Members

Travis Weller and Amalia Monroe-Gulick

University of Kansas, Lawrence, KS, USA

Abstract

Objective: In light of academic libraries expanding the data services they offer, this article hopes to improve the understanding of data practices, challenges and future needs of academic library user groups.

Setting, Design and Method: In the fall of 2013, librarians and campus grant specialist conducted an institution wide web-based survey at the University of Kansas, a major public research institution, with several hundred respondents.

Results: Graduate students and faculty members report differences in their data practices, research challenges and data-related needs.

Conclusions: Academic libraries should target data services to the interests and needs of their distinct user groups.

Introduction

As data-intensive research continues to gain prominence among academic researchers, academic libraries are expanding the data-related services they provide.

Academic libraries have readily embraced increasing research data services (Jones 2014; Tenopir, Birch, and Allard 2012; Keralis et al. 2014). For example, in 2013, the Association of Research Libraries (ARL) released a kit of materials to share data-related policies, job descriptions, and other materials among its member institutions (Fearon et al. 2013). Due to this organizational support for expanded data services, individual academic librarians are increasingly providing data-related services to library patrons. A 2012 survey revealed that two-thirds of academic librarians provide research data services at least occasionally, but for nearly three-quarters of librarians, providing data services is not yet an integral part of the job duties (Tenopir, Sandusky, Allard, and Birch 2013).

Despite interest in expanding research data services, libraries are still in the "early stages" (Fearon et al. 2013). Researchers are calling on academic libraries to "step up" to the challenge of managing and sharing research data (Keil 2014). Influenced by requirements set by federal research funding agencies, libraries have focused, thus far, on assisting researchers as they write data management plans (Fearon et al. 2013; Keralis et al. 2014; Tenopir, Birch, and Allard 2012). Helping with data plans was seen as a way for libraries to engage research-ers on data-related topics (Fearon et al. 2013); however, libraries are quickly moving beyond that. A 2013 survey of ARL libraries revealed that 74% already offer some data services, and another 23% will add those services in the near future[1] (Fearon et al. 2013). The vast majority of those that offer data services have added them in the last four years, and most responding libraries reported that they plan to significantly increase the services they offer over the next two years. Expanded data services may include, among other things, one-on-one consultation on data management issues, archiving of research data, and more detailed online guides. Generally, the potential data roles for librarians can be defined as informational, instructional, infrastructural, cooperative, collaborative, and archival (Keralis et al. 2014). As most academic libraries at major research institutions experiment with new models and services, it is guaranteed that there will be institutional variation in the scope of data services offered.

To ensure that these new data services are responsive to the needs of their patrons, it is essential that academic libraries have a detailed understanding of the data practices and needs of different populations of users. A recent Council on Library and Information Resources (CLIR) report cited an "acute need" for further research that can inform data-related curriculum and training (Keralis et al. 2014).

In a separate paper, we detailed differences in the data needs of academic researchers generally by discipline and research methodology (Weller and Monroe-Gulick 2014). Here, we examine these differences in further detail by comparing the data practices and challenges of graduate students and faculty members.

---

[1] A separate study, conducted in 2012, also indicates a dramatic expansion of data-related services in the near future, particularly by academic libraries at research institutions (Tenopir, Birch and Allard 2012).

Both faculty and graduate students conduct data intensive research. However, they differ in terms of research experience, their familiarity with digital resources, and goals for their work while at a university. As a result, it is inevitable that their interactions with, use and understanding of data will be different, but these differences have, until now, received little attention in the literature. In order to provide effective data services to their users, academic libraries need to ensure they understand the data practices of graduate students and faculty.

Literature Review

A number of studies have examined the data practices of faculty and graduate students, separately.

Graduate students interact with data both to conduct their own research, but also by managing the research data for larger projects led by faculty (Carlson et al. 2011; Adamick, Reznik-Zellen, and Sherdian 2012; Johnston and Jefferyes 2014). Faculty expect graduate students to begin their work with data management skills, needing no formal training (Carlson et al. 2013) As a result, graduate students often learn their skills through trial and error, learning on the job, from each other and by searching online (Carlson et al. 2013; Johnston and Jeffryes 2014).

It is common for graduate students to save data in emails and in commercial cloud storage (Piorun et al. 2012). When formatting, cleaning, and analyzing data, students often struggle with providing appropriate and necessary documentation (Carlson et al. 2013). In general, the local and immediate concerns of an individual project or lab are often prioritized over larger, more generalized learning (Carlson et al. 2013).

In a multi-method study, Carlson and his associated researchers investigated the data practices and literacies of graduate students (Carlson et al. 2011). They conducted interviews with faculty that advised graduate students and assessed graduate students enrolled in a Geoinformatics course. Faculty expressed concerns over graduate students' organizational skills as well as their long-term dedication to the data that they helped to generate, clean, and analyze. The authors concluded that graduate students sometimes struggle with basic information technology issues that complicate their work with data.

A focused case study of structural engineering graduate students concluded that they rarely backup their data and need additional training on data analysis (Johnston and Jeffryes 2014). These graduate students did not see a need for archiving and were unclear about the need for long-term data access.

Many of the data practices of graduate students are also demonstrated by faculty members as well.

A study of faculty with active NSF research grants at Cornell revealed that most relied on their own computer infrastructure for backing up their data, were unsure whether their data was appropriately documented, and wanted to generate the metadata themselves, but a majority also wanted guidance on writing data plans (Steinhart et al. 2012). In a faculty-focused survey at Georgia Tech, nearly three-fourths of respondents wanted additional support for data storage and preservation and half wanted more information on best practices for data management (Wells Parham, Bodnar, and Fuchs 2012).

Akers and Doty conducted a study at Emory on the data practices of faculty (2012). More than half of all tenured and tenure-track faculty members in their survey were familiar with funding agency requirements for data management planning. Data was most commonly stored on computer hard drives (Akers and Doty 2013). Less frequently, faculty used university-based servers for storage, as well as internet-based commercial storage services, but even more rarely than university servers. Only a small percentage of faculty members reported that they preserved their data for the long-term in data repositories, but more than half are interested in beginning to do so in the future. Faculty members reported that the data-related services they were most interested in were workshops on data topics and assistance preparing data management plans.

Recently, a group of librarians at Colorado State University conducted focus groups on data practices and needs with faculty as well as with staff members that conduct research, such as research scientists, but not graduate students (McLure et al. 2014). This study revealed that researchers struggle with all stages of the data lifecycle, and some study participants specifically cited needing additional assistance with data storage and transfer.

A number of other institutions have completed faculty-focused surveys of data practices, for example, Cal Poly, (Scaramozzino et al. 2012), the University of Houston (Peters and Dryden 2011), Georgia Tech (Wells Parham et al. 2012), and the University of Nottingham (Parsons et al. 2014).

While most studies have focused on either faculty or graduate students, a small number have included both groups in their studies. For example, a 2009 study at the University of Minnesota, included both faculty members and graduate students in their survey on various research data topics (Johnston 2014). They found that researchers want data storage to be locally controlled and uncomplicated, they want to share their data with others, and would like to keep their data forever. While faculty, research staff, and students were all included in this survey, the published results did not distinguish between user groups.

Additionally, an in-depth investigation conducted for CLIR into research data practices included graduate students as well as faculty members (Jahnke and Asher 2014). The authors found that the pressures of publication outweighed careful curation of data and long-term data preservation was not valued. No participants in the study had formal data management training; instead, they were learning it in an ad hoc fashion. However, this qualitative study focused exclusively on social scientists and did not draw separate conclusions or comparisons for graduate students and faculty members.

Methods

In this paper, we paper provide information on the data practices, challenges, and needs of graduate students and faculty in a comparative manner. By examining these questions at a single university, we have control for differences in the resources available and in institution-wide policies. This piece is meant to contribute detail to the developing understanding of researcher data management. We hope it can be used by libraries to develop data services that are targeted to specific user groups, and reveal competencies in areas where faculty and graduate students can better support each other. At the same time, the results here are representative of a single institution, so it is possible that some of these differences are due to

unique, institution-specific factors. This underscores the need for larger studies of this type to be conducted across institutions or, if possible, a meta-analysis.

Three librarians and a campus grant-specialist developed a survey to identify current and future practices, research, and support needs, including sections on current data practices and future needs (see Appendix A). An online survey was developed and then pre-tested with representation from all members of the survey target populations: faculty members, research staff, and graduate students in the humanities, social sciences, and sciences.

The survey was distributed during the fall 2013 semester. It was circulated using email distribution lists that reached our three target populations. An email message including a link to the survey was sent from the Dean of Libraries and the Dean of Research & Graduate Studies. One reminder message was sent. The survey was open for one month.

Survey Completion Rate

The University of Kansas (KU) Office of Institutional Research and Planning (OIRP) reports that in fall 2013 KU (Lawrence and Edwards campuses) had 5,691 enrolled graduate students and 1,626 faculty members. The final response rate for each population was 5% of graduate students (n=271) and 9% of faculty members (n=146). Discussed in more detail below, one of the target populations, non-faculty, research staff was excluded from this analysis.

Limitations

Although a goal of this survey was to better understand the data needs and practices of non-faculty and research staff, this population was ultimately excluded from this analysis. In addition to responses about faculty or graduate student status, respondents were also allowed to select "academic staff," "staff," and "postdoctoral researcher." Comments included in the open-text other field indicated confusion from respondents about these options given job definitions within the university not being parallel with the options given in the survey. Also, very few respondents selected these options. Due to the limited number of responses, stated confusion, and the limited nature of the literature regarding non-faculty, non-student research staff, the authors decided to exclude these individuals from this paper.

Results

The KU survey revealed differences between graduate students and faculty members in the methods used to store files of various types. Saving materials to a hard drive or a CD is most common for all file types for both groups. However, graduate students are more likely to rely on cloud-based storage than faculty members (see Chart 1). In contrast, faculty members use university servers more often.

One faculty member who backs up all of his/her data on a hard drive or CD described his/her concern about commercial storage options in space available for open-ended comments:

> *"Don't trust private company to maintain data free forever. Backing up my files fre-quently is more trustworthy. I would consider backing them up on a university server."*

Chart 1: Usage of Cloud and University Server Storage
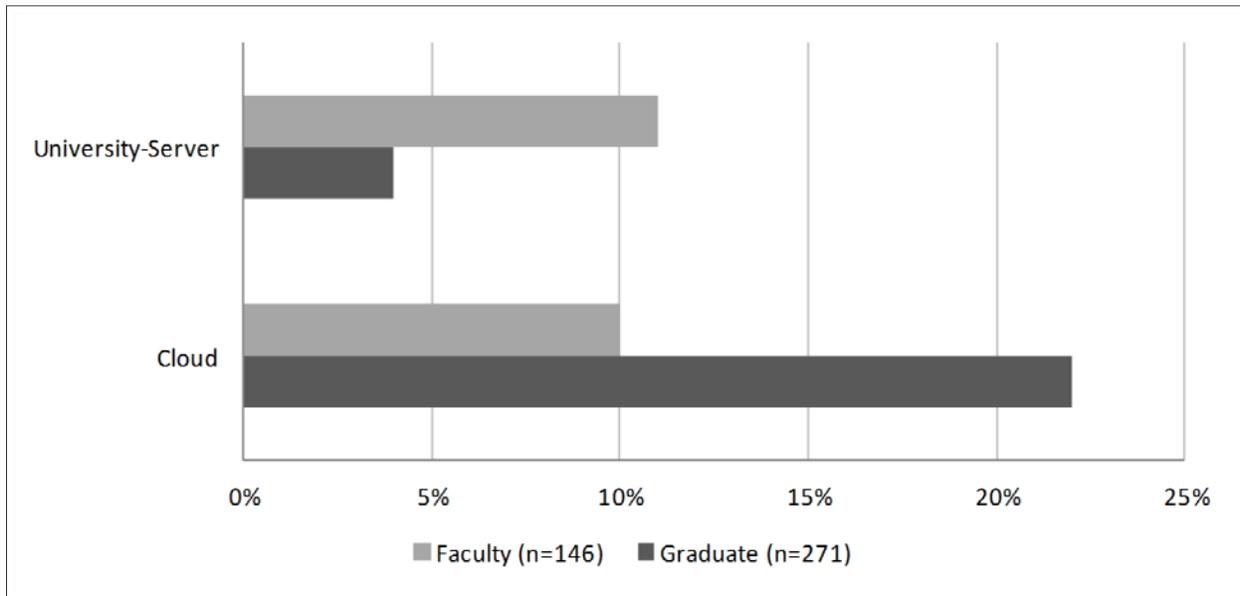
But this faculty member did not explain why s/he did not currently use the university server for storage.

Graduate students and faculty members also reported differences in the factors that they say influence their storage methods.

Faculty members are more influenced by ease-of-storage method than graduate students.[2] The only other factor that faculty cited more frequently than graduate students were grant requirements about data, the least cited concern overall (see Chart 2). For all other factors — backup needs, cost, file size, long-term sustainability, physical space requirements, and privacy and security — graduate students responded that their data storage decisions were more influenced by that factor than faculty.

In this question, we provided respondents with a free-text box where they could insert their own answers as well. Several respondents added "ease of access" in this box, indicating that it is not just whether the method of storage is easy but also whether later attempts to access this data is easy.

From a list of research phases involving data, respondents were asked to select the one that they found most challenging or time-consuming (see Chart 3). The phases were generalized to be applicable to a wide range of research methods. Graduate students and faculty differed in the phase they reported as the most challenging.

Faculty find acquiring access to data significantly more challenging than the graduate students.

---

[2] All percentage responses to questions displayed in charts are calculated based on the number of respondents in the target population for this question.

Chart 2: Factors in Data Storage Decisions
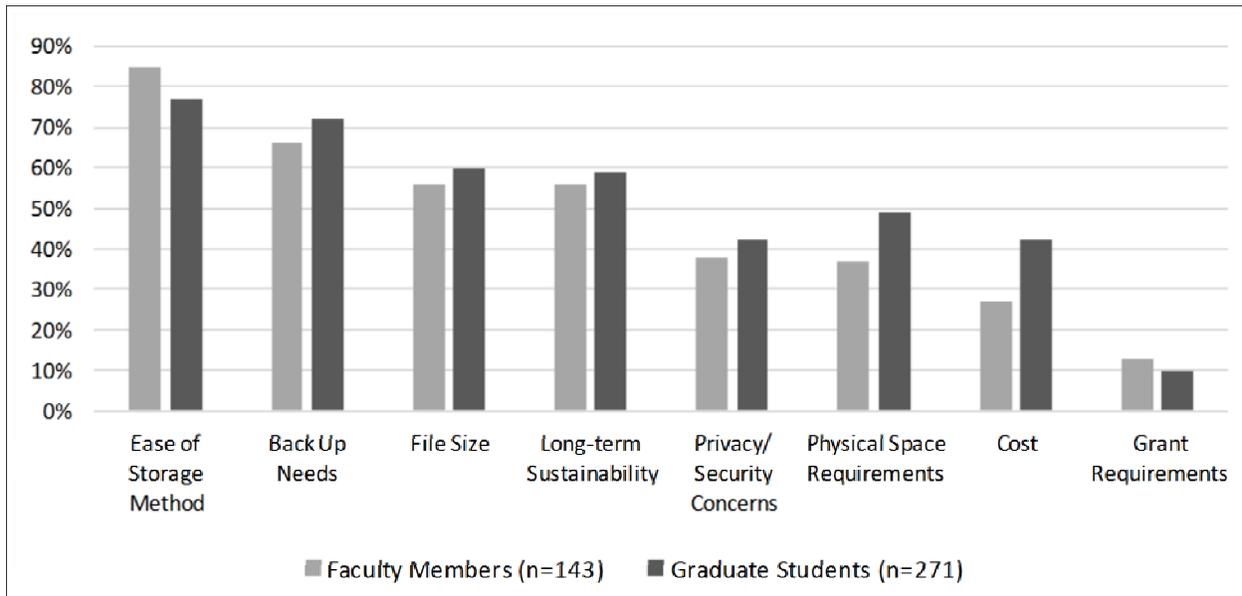
Graduate students have more difficulty identifying relevant data and managing data than faculty. Disseminating research results was the least challenging for both populations, but graduate students did select it slightly more often than faculty. A handful of graduate students added in the open-ended response for this question that they find analyzing the data most difficult. This is consistent with the response from graduate students to the next question about researchers' future needs.

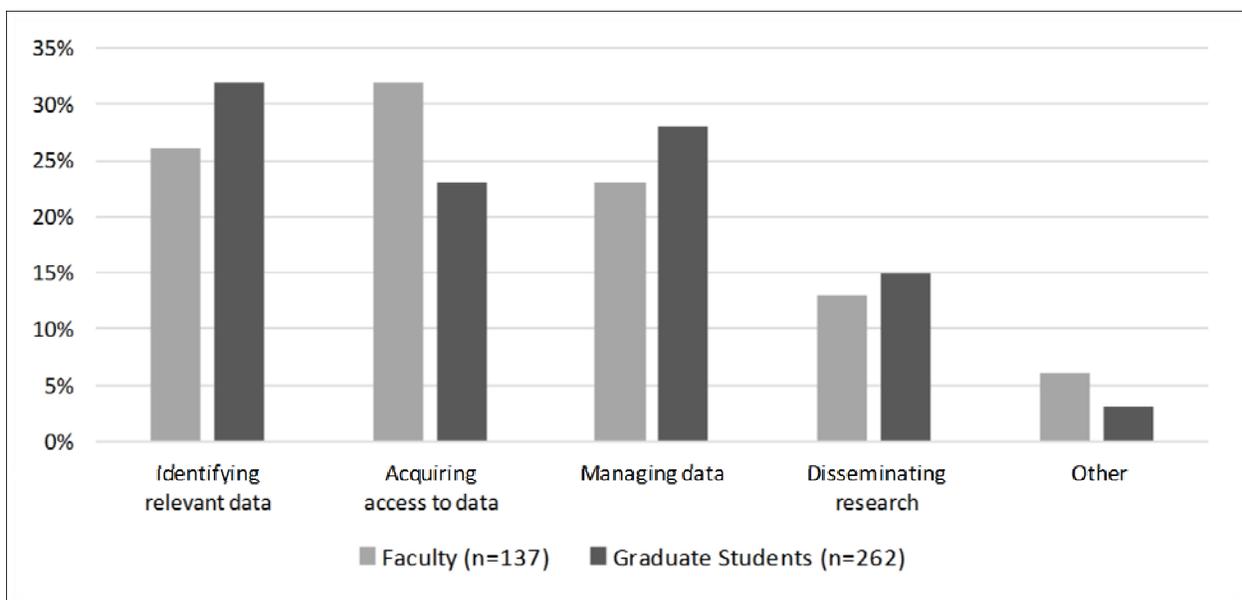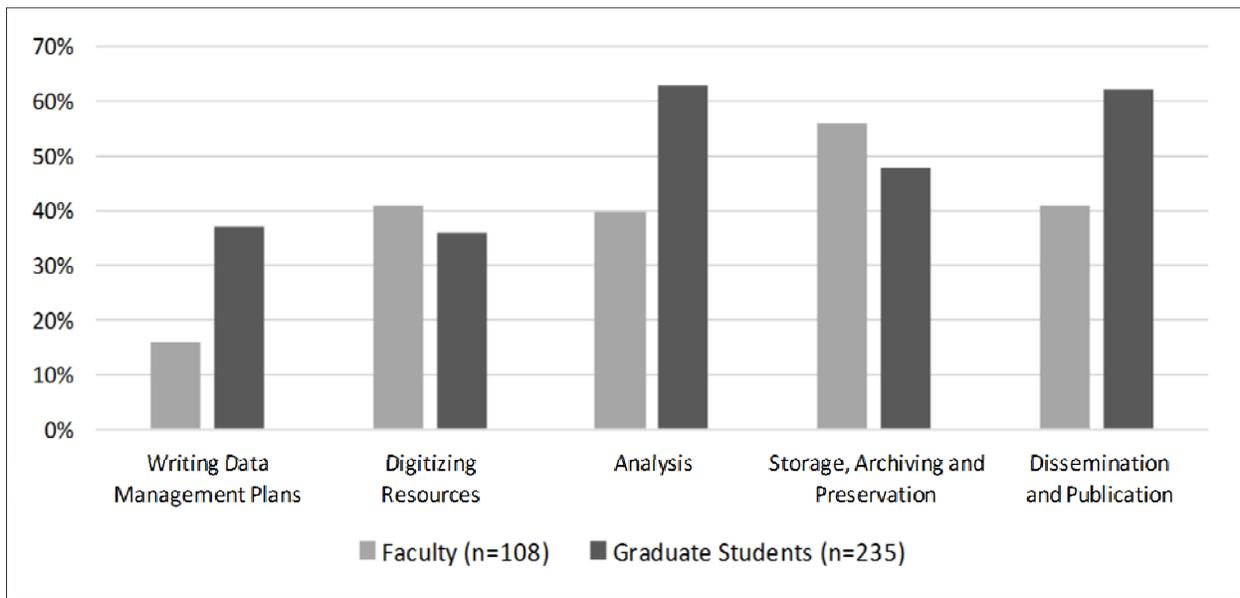**Chart 3**: Most Challenging Phases of the Research Process

Chart 4: Future Needs with Research Data

Finally, respondents were asked to select what needs associated with their research data that they anticipate in the future (see Chart 4). They were free to select as many as were applicable to them.

Analysis and dissemination are two areas where graduate students anticipate needing more assistance than faculty. Faculty, on the other hand, expect that they will need support with digitization and data storage. Interestingly, graduate students are much more likely to want help with drafting data management plans, while only a relatively small number of faculty anticipate needing help with these plans.

Discussion and Conclusions

These results indicate that researchers are likely to welcome academic libraries' shift to provide more expansive data services. As libraries design their services, these results can be useful for creating targeted programming and workshops.

Thus far, driven by research funding agency requirements, libraries have focused their data services on helping with the development of data management plans. Our results indicate that this is the service that researchers overall anticipate needing the least help with in the future. This is underscored by the fact that both graduate students and faculty members reported that meeting grant requirements was the least influential factor in their data storage decisions. However, graduate students expressed a much greater need for assistance in drafting data plans, so workshops on writing data management plans may be a good outlet for connecting with graduate students. Connecting with faculty may require a different approach to make an impact or connection since only 12% expressed a concern with grant requirements in their data storage decision-making; yet, libraries still continue to make this the centerpiece of their data service programs.

An area for further inquiry, especially for librarians or others who provide support for researchers, is the lack of consideration for grant requirements in the decision-making process. Grant requirements, so far, have been focused on pre-project planning and do not add explicit requirements about data practices as long as they are discussed in the proposed plan. As a result, it is not surprising that grant requirements were cited infrequently. However, this will likely change in the near future given the White House Office of Science & Technology Policy memo regarding open data (Holdren 2013).

There is a high level of interest for both populations in topics that libraries have a long track record of service in — long-term storage, preservation and archiving, and dissemination and publication. Rather than looking at data management planning to open the door to library-provided data services, libraries can build on their existing experience and expertise in these areas to provide needed services to researchers and build a relationship with them around data.

The high number of students that need assistance with data analysis is challenging, but this too may be an opening for libraries. This indicates that graduate students may be responsive to workshops around specific data tools or analytical techniques. If libraries can collaborate with faculty advisors teaching methods, these workshops could provide an opportunity for libraries to meet a need for graduate students and demonstrate their ability to faculty at the same time. In addition, libraries may consider partnering with other departments and research centers on campus to identify where analytic expertise already exists so they can connect researchers with resources when the need arises.

There is also a need, particularly among graduate students, for easy-to-use, secure, data storage options. Our results demonstrate that university servers are still relatively underutilized, so additional education about the benefits of university-provided storage and the drawbacks of commercial storage is necessary, whether provided through data management workshops by librarians or through other campus departments, like IT. University-provided options will need to be evaluated to make sure that they are responsive to the factors that researchers cite as important for their research data. As the graduate students of today become the faculty members of tomorrow, and commercial storage options become more prevalent and integrated into the user experience, this need will only become more pressing.

Interestingly, graduate students reported to be much more influenced by privacy and security concerns than faculty members, even though they use commercial cloud storage, generally considered to be less secure, at a higher rate than faculty members. Graduate students likely embrace cloud-based commercial storage because the accounts move with the individual as opposed to university-based servers that belong to the institution. Given that graduate students acknowledge concerns over backing up their data, the size of the data, and physical storage, librarians can find a willing audience among graduate students and change the culture of data management with the next generation.

In contrast, faculty use institutional storage options more often. This is likely due to the fact they have more time at an institution so are more aware of their options. Additional work in this area is necessary though. More than half of the faculty members in our study responded that they anticipated needing assistance with storage, archiving, and preservation. This is an opportunity for librarians, particularly those working in institutional repositories, to engage with

faculty in area where they tend to have knowledge and expertise.

It is clear, also, from these results, that not everything is digital yet. More than a third of both faculty and graduate students anticipate needing assistance with digitizing materials in the future. This is an important reminder that physical format materials are still used regularly and researchers will need assistance manipulating and storing these items digitally.

As research continues to evolve, these responses will also evolve. Therefore, it will be necessary for libraries — and academic institutions in general — to continue to monitor the research data practices and needs of the researchers. Academic libraries can and do contribute to supporting these needs; however, successful collaboration with other academic units may be the most sustainable path for libraries to continue to play a role in supporting research data. Resources are scarce for all academic units, not just libraries, and partnerships will allow for deduplication in efforts and allow for each unit to focus on their unique strengths, which will ultimately enhance services for all researchers. Finally, these enhanced collaborations will allow for more timely evaluations that lead to further agility and adjustments in services as they arise.

Supplemental Content

Appendix A
An online supplement to this article can be found at http://dx.doi.org/10.7191/jeslib.2015.1070 under "Additional Files".

Acknowledgement

Disclosure

The authors report no conflict of interest.

References

Adamick, Jessica, Rebecca C. Reznik-Zellen, and Matt Sheridan. 2012. "Data Management Training for Graduate Students at a Large Research University." *Journal of eScience Librarianship* 1:180–188.
http://dx.doi.org/10.7191/jeslib.2012.1022

Akers, Katherine G. and Jennifer Doty. 2012. "Differences among Faculty Ranks in Views on Research Data Management." *IASSIST Quarterly* Summer 2012:16-20.
http://www.iassistdata.org/downloads/iqvol36_2_doty_0.pdf

Akers, Katherine G. and Jennifer Doty. 2013. "Disciplinary differences in faculty research data management practices and perspectives." *The International Journal of Digital Curation* 8:5-26.
http://dx.doi.org/10.2218/ijdc.v8i2.263

Carlson, Jacob, Michael Fosmire, C.C. Miller, and Megan Sapp Nelson. 2011. "Determining Data Information Literacy Needs: A Study of Students and Research Faculty." *portal: Libraries and the Academy* 11:629-657.
http://dx.doi.org/10.1353/pla.2011.0022

Carlson, Jake, Lisa Johnston, Brian Westra, and Mason Nichols. 2013. "Developing an Approach for Data Management Education: A Report from the Data Information Literacy Project." *International Journal of Digital Curation* 8:204-217. http://dx.doi.org/10.2218/ijdc.v8i1.254

Carlson, Jake and Marianne Stowell-Bracke. 2013. "Data Management and Sharing from the Perspective of Graduate Students: An Examination of the Culture and Practice at the Water Quality Field Station." *portal: Libraries and the Academy* 13:363-384. http://dx.doi.org/10.1353/pla.2013.0034

Fearon, David Jr., Betsy Gunia, Sherry Lake, Barbara E. Pralle, and Andrew L. Sallans. 2013. "SPEC Kit 334: Research Data Management Services." *Association of Research Libraries.* http://publications.arl.org/Research-Data-Management-Services-SPEC-Kit-334/

Holdren, John P. 2013. "Increasing Access to the Results of Federally Funded Scientific Research." *White House Office of Science and Technology Policy*. Last modified February 22. https://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf

Jahnke, Lori and Andrew Asher. 2014. "The Problem of Data." *Council on Library and Information Resources*. Accessed June 26. http://www.clir.org/pubs/reports/pub154

Johnston, Lisa and Jon Jeffryes. 2014. "Data Management Skills Needed by Structural Engineering Students: Case Study at the University of Minnesota," *Journal of Professional Issues in Engineering Education and Practice* 140. http://dx.doi.org/10.1061/(ASCE)EI.1943-5541.0000154

Johnston, Lisa. 2014. "User-needs assessment of the research cyberinfrastructure for the 21st century." *International Association of Scientific and Technological University Libraries, 31st Annual Conference.* Accessed June 26. http://docs.lib.purdue.edu/iatul2010/conf/day1/5

Jones, Elisabeth. 2014. "E-Science Talking Points for ARL Deans and Directors." *Association of Research Libraries*. Accessed June 26. http://www.arl.org/storage/documents/publications/e-science-talking-points.pdf

Keil, Deborah E. 2014. "Research Data Needs from Academic Libraries: The Perspective of a Faculty Researcher." *Journal of Library Administration* 54:233-240. http://dx.doi.org/10.1080/01930826.2014.915168

Keralis, Spencer D.C., Shannon Stark, Martin Halbert, and William E. Moen. 2014. "Research Data Management in Policy and Practice: The DataRes Project." *Research Data Management: Principles, Practice, and Prospects, Council on Library and Information Resources*. Accessed June 26. http://www.clir.org/pubs/reports/pub160/pub160.pdf

McLure, Merinda, Allison V. Level, Catherine L. Cranston, Beth Oehlerts, and Mike Culbertson. 2014. "Data Curation: A Study of Researcher Practices and Needs." *portal: Libraries and the Academy* 14:139-164. http://dx.doi.org/10.1353/pla.2014.0009

Parsons, Thomas, Shirley Grimshaw, and Laurian Williams. 2014. "Research Data Management Survey." *University of Nottingham*. Accessed June 26. http://admire.jiscinvolve.org/wp/files/2013/02/ADMIRe-Survey-Results-and-Analysis-2013.pdf

Peters, Christie and Anita Riley Dryden. 2011. "Assessing the Academic Library's Role in Campus-Wide Research Data Management: A First Step at the University of Houston." *Science & Technology Libraries* 30:387-403. http://dx.doi.org/10.1080/0194262X.2011.626340

Piorun, Mary E., Donna Kafel, Tracey Leger-Hornby, Siamak Najafi, Elaine R. Martin, Paul Colombo, and Nancy R. LaPelle. 2012. "Teaching Research Data Management: An Undergraduate/Graduate Curriculum." *Journal of eScience Librarianship* 1:46-50. http://dx.doi.org/10.7191/jeslib.2012.1003

Scaramozzino, Jeanine Marie, Marisa L. Ramírez, and Karen J. McGaughey. 2012. "A Study of Faculty Data Curation Behaviors and Attitudes at Teaching-Centered University." *College & Research Libraries* 73:349-365. http://crl.acrl.org/content/73/4/349.full.pdf

Shaffer, Christopher J. 2013. "The Role of the Library in the Research Enterprise." *Journal of eScience Librarianship* 2:8-15. http://dx.doi.org/10.7191/jeslib.2013.1043

Steinhart, Gail, Eric Chen, Florio Arguillas, Dianne Dietrich, and Stefan Kramer. 2012. "Prepared to Plan? A Snapshot of Researcher Readiness to Address Data Management Planning Requirements." *Journal of eScience Librarianship* 1:63-78. http://dx.doi.org/10.7191/jeslib.2012.1008

Tenopir, Carol, Ben Birch, and Suzie Allard. 2012. "Academic Libraries and Research Data Services: Current Practices and Plans for the Future, An ACRL White Paper." *Association of College and Research Libraries (ACRL)*. http://www.ala.org/acrl/sites/ala.org.acrl/files/content/publications/whitepapers/Tenopir_Birch_Allard.pdf

Tenopir, Carol, Robert J. Sandusky, Suzie Allard, and Ben Birch. 2013. "Academic librarians and research data services: preparation and attitudes." *IFLA Journal* 39:70-78. http://dx.doi.org/10.1177/0340035212473089

Weller, Travis and Amalia Monroe-Gulick. 2014. "Understanding methodological and disciplinary differences in the data practices of academic researchers." *Library Hi Tech* 32:467-48. http://dx.doi.org/10.1108/LHT-02-2014-0021

Wells Parham, Susan, Jon Bodnar, and Sara Fuchs. 2012. "Supporting tomorrow's research: Assessing faculty data curation needs at Georgia Tech." *College & Research Libraries News* 73:10-13. http://crln.acrl.org/content/73/1/10.full.pdf+html