

Developing a Data Management Consultation Service for Faculty Researchers: A Case Study from a Large Midwestern Public University

Virginia A. Dressler
Kent State University

Kristin Yeager
Kent State University

Elizabeth A. Richardson
Kent State University

Abstract

To inform the development of data management services, a library research team at Kent State University conducted a survey of all tenured, tenure-track, and non-tenure track faculty about their data management practices and perceptions. The methodology and results will be presented in the article, as well as how this information was used to inform future work in the library's internal working group. Recommendations will be presented that other academic libraries could model in order to develop similar services at their institutions. Personal anecdotes are included that help ascertain current practices and sentiments around research data from the perspective of the researcher. The article addresses the particular needs of a large Midwestern U.S. academic campus, which are not currently reflected in literature on the topic.

Received 16 November 2018 ~ *Accepted* 17 January 2019

Correspondence should be addressed to Virginia A Dressler, 1125 Risman Dr, Kent, Ohio 44242. Email: vdressle@kent.edu

The *International Journal of Digital Curation* is an international journal committed to scholarly excellence and dedicated to the advancement of digital curation across a wide range of sectors. The IJDC is published by the University of Edinburgh on behalf of the Digital Curation Centre. ISSN: 1746-8256. URL: <http://www.ijdc.net/>

Copyright rests with the authors. This work is released under a Creative Commons Attribution Licence, version 4.0. For details please see <https://creativecommons.org/licenses/by/4.0/>



Introduction

Researchers are increasingly required to develop plans for long-term data management and sharing, oftentimes in the U.S. by mandate of grant directives such as NEH, NSF and others. Yet data management and related activities are skill sets that are not always inherent ones to many researchers. Libraries have track records of providing support and outreach in a number of their traditional service points (such as reference and instruction) (Kong, Fosmire and Dewayne Branch, 2017). Certain librarian positions within digital library and institutional repository initiatives, for example, require skill sets in digital media management and overarching digital preservation knowledge. We believe that the library is well positioned to be leaders on this topic and provide relevant consultation and services to support these endeavors.

In early 2017, Kent State University Libraries formally addressed expanding consultation services to include data management through a new internal working group. The working group included members of reference, instruction, institutional repository, technical services, and digital projects. A team of three librarians from the working group conducted a survey during the Fall of 2017 to investigate research data management issues and practices at the institution.

The initial goal was to learn about data management services at other institutions and to understand current data management practices among Kent State University faculty. The team also wanted to know what services or programs would be useful to faculty researchers managing or sharing research data. Finally, we wanted to identify other departments outside of the library for collaboration on the provision of data management services. This article will highlight the top takeaways from the survey that were used to define and implement new services.

Literature Review

A literature review was conducted to examine research articles addressing implementation of data management services within the library, and also includes the review of some higher quality resources and toolkits to assist librarians working within research data topics. The search strategy for the literature review included looking at more recently published scholarly publications, primarily published within the last ten years, focused on data management and academic libraries.

Data Management Services within the Academic Library

The literature emphasizes the importance of research data management and mentions several obstacles to data management and sharing. Patel (2016) states the important role of data in research projects and the benefits of sharing research data along with some challenges in research data management, including copyright, data licensing, erroneous interpretation of data, security, privacy, and a mind-set that prevents some researchers from sharing their data. Whitmire, Boock, and Sutton (2015) highlight the need for standardized metadata as an additional challenge and suggest the library can develop services and training to assist researchers with data management. Funari, in a 2014 article about research data with a focus on a European context and research data in the

humanities, summarizes categories of obstacles to research data access and re-use, including legal, financial, and technical. Funari advises that “research data” may be defined differently by different organizations and may differ greatly in quantity and typology, “and also for the degree of necessity and practice in their sharing and re-using” (Funari, 2014) between natural sciences and humanities.

Faniel and Silipigni Connaway (2018) performed a qualitative research study to examine the experience of academic librarians with research data management programs to support researchers. Qualitative data was collected through hour-long interviews with thirty six library professionals. The authors highlighted five factors of influence which include technical resources, human resources, researchers’ perceptions about the library, leadership support and finally communication, coordination and collaboration. These factors were ultimately found to either act as facilitators for academic librarians to support these initiatives, or as constraints. Further, they call for more subject and technical expertise to address the complex needs of research data. An interesting element that was present throughout the interviews was the fact that some services were still either in planning stages, or in very early stages of implementation, which points to the fact that many of these services are still in their infancy. Core services of these programs centered around writing data management plans, depositing data and/or managing data. The preference for this work by library staff to take place at an early point of the research cycle, an aspect that would entail communication of services, close collaborators with other entities on campus who work regularly with researchers at different points in the project (Office of Research, Information Services/Information Technology, and others).

Chen and Zhang (2017) examined job descriptions and required and preferred qualifications for library job announcements that included the word “data” in the job title to determine what knowledge and skills successful candidates for these positions should have. While the study’s aim was to inform the curriculum of LIS programs, their findings indicate the types of data services that libraries offer, or expect to offer, to meet the needs of their institution’s researchers. Most of the positions examined required the candidate to be able to assist faculty and students with data collection, data management, and data analysis. Some postings mentioned specific software or tools, but these varied.

Frank and Pharo (2016) used a modified Delphi method with two rounds in their study of meteorology students and associated stakeholders at the University of Oslo to assess perceptions of data information literacy and attitudes about its instruction for meteorology graduate students. They formed a panel of experts composed of meteorology professors at the University of Oslo’s Department of Geosciences’ Meteorology Section, researchers from the Meteorological Institute (MET), PhD students from the university’s Department of Geosciences’ Meteorology Section, and academic librarians from the University of Oslo’s Science Library. All panelists agreed that data information literacy skills are important for graduate students in meteorology. There was less consensus about the role of librarians – even among the three librarians on the panel – as stakeholders regarding data information literacy of meteorology graduate students, but all three librarians identified future roles for librarians. Several obstacles to library involvement in data information literacy training were identified.

In November of 2014, a survey was sent to Teaching and Research (T&R) faculty and Research faculty at Virginia Polytechnic Institute and State University to learn about faculty researchers’ existing practices to organize, describe, and preserve data, and their needs for services and education (Shen, 2016). A “lack of systematic planning and preservation activities” along with limited storage options and “sporadic and informal

documentation practices” were noted. The study also identified the need for technical support, application of metadata standards, and education. One interesting finding that Shen notes is that some faculty researchers seemed to be under the impression that addressing issues mentioned in IRB policies, such as confidentiality and sensitivity of data, is akin to data management planning.

Goben and Nelson (2018) outline a new initiative from the Association of College and Research Libraries (ACRL), a full day workshop, “ACRL RoadShow, Building Your Research Data Management Toolkit: Integrating RDM into Your Liaison Work”. The development and design of the module was a backwards design, in that the desired results were used as the starting point of the course. As Goben and Nelson point out, “As academics across disciplines face increasing need for data management skills, librarians have an opportunity to apply their expertise in this additional realm.” This article highlight the importance of integrating data management skill sets into the work of the liaison, who often come across many opportunities for outreach and education through regular job duties.

Whitmire, Boock, and Sutton (2015) presented a case study of using survey results to inform the development of data management services at the Oregon State University Library. Their survey asked faculty at that institution about the type and volume of data generated; who performs the tasks associated with research data management in their research teams; and their current practices of metadata creation. A major finding from their survey was the differences in where faculty stored data; most notably, that over half of the faculty in the colleges of Engineering, Science, and Veterinary Medicine reported storing data on servers that they themselves maintained.

Resources and Toolkits

At University of California Berkeley, a training approach was built to address current knowledge gaps around data management skill sets for all subject liaisons (Wittenberg, Sackmann and Jaffe, 2018). The authors reported that the success of this approach was a higher rate than previous unit-wide efforts to train libraries on the topic. Identified services were shared in the article, as well as reflections on the successfulness of these initiatives; “... the success of their efforts is equally dependent on the process by which they develop these new capabilities” (Wittenberg, Sackmann and Jaffe, 2018). The Librarian Training Program at Berkeley was deemed to be very successful in regard to bringing more awareness and knowledge around research data management to all subject liaisons, regardless of their speciality or expertise. Work will continue to address subject specific needs, but it is admirable that an all-encompassing training initiative was instilled to educate many of the staff members who work most directly with researchers.

The Data Curation Profile Toolkit from Purdue (Carlson, 2010) is a useful place to begin for an institution seeking guidelines and recommendations in the identification and assessment processes of how researchers are currently managing or curating data. The toolkit includes four components; The User Guide, The Interviewer’s Manual, The Interview Worksheet, and the Data Curation Profile Template. The manual, worksheet, and template provide some useful frameworks for practitioners to more consistently and methodically capture information around research needs.

In June 2018, the non-profit Joint Information Systems Committee (JISC) released a research data management toolkit, which features options for three different user types to interact with the included resources (researcher, research support, IT specialist). The toolkit compiles information about research data management, policy planning, infrastructure, associated costs, storage/backup, and more categories. There are also

related resources for courses, videos, and other guides to assist in many aspects under research data management.

Materials and Methods

Kent State is a public university, with eight campuses throughout northeast Ohio, with a total student enrollment of over 36,000 students (including both undergraduate and graduate). Kent State University is designated an R2 Carnegie Classification, and employs approximately 2,700 academic faculty. Kent State University Libraries currently offers consultation services for researchers on the topics of literature reviews, copyright, affordable course materials, and data analysis software.

The questions the team sought to answer were:

- **RQ1:** What questions should libraries ask of faculty when developing data management services? What sort of data management services would be the most appropriate for our faculty?
- **RQ2:** What are the knowledge gaps among faculty about data management that the library could help fill?
- **RQ3:** How can the library identify collaborators on campus to participate in the development of data management services?
- **RQ4:** Are there any differences in data management practices and attitudes between disciplines that libraries should take into account when developing data management services?

To investigate these research questions, a Qualtrics survey was distributed via email to all tenured, tenure-track, non-tenure track, and adjunct faculty at all Kent State University campuses (n=2749). The survey was distributed on October 19, 2017, and closed on November 20, 2017, with two reminder emails sent in the interim. In-progress responses were closed after one week. The survey instrument was adapted from Whitmire, Boock, and Sutton (2015), with minor changes incorporated to address our specific institution.

The data was analyzed using R, an open source statistical software package (R Core Team, 2018). Multiple choice questions were analyzed using frequencies and proportions. Qualitative, open-ended questions were content-analyzed by all contributing authors independently before coming to consensus about final coding. Extended response answers were manually coded by the researchers for mentions of aspects such as software, common practices, and other facets.

Of the initial distribution, 287 people started the survey, for a response rate of 10.4%. Twenty-eight respondents were filtered out after answering “Never” to the first question, which ended the survey for these respondents. In all, there were 259 responses with usable data for some, or all, of the survey questions. There were 180 who completed the survey to the end, for a completion rate¹ of 69.5% (Figure 1).

¹ $(208-28)/(287-28) = 69.5\%$

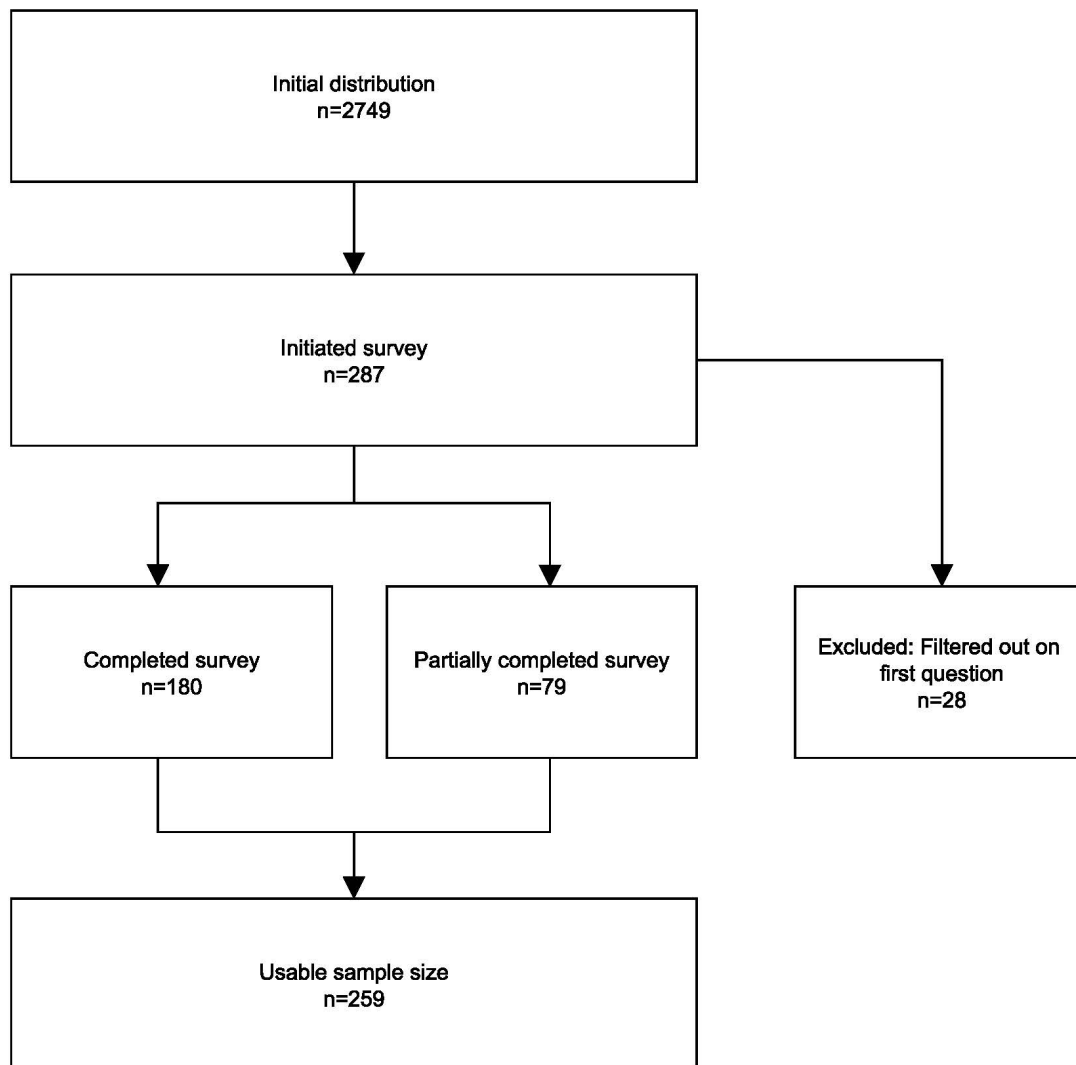


Figure 1. Diagram of distribution and sample sizes.

Results

Demographics

Respondents were asked to self-report their faculty status and college affiliation. The majority of respondents were tenured faculty (see Table 1). Note that the ‘not applicable’ faculty status responses may represent adjuncts who do not consider themselves non-tenure-track.

While our research team does want to create inclusive service points around research data, we particularly want to be sure that tenure-track faculty are receiving the support needed to acquire tenure. Future surveys may look to isolate this group for further study to get more information from this particular base of researchers.

Table 1. Self-reported demographics of survey sample.

	n	%
College (n=177)		
College of the Arts and Sciences	71	40.1
College of Education, Health and Human Services	40	22.6
College of Communication and Information	25	14.1
College of the Arts	14	7.9
College of Nursing	10	5.6
College of Business Administration	7	4
College of Public Health	4	2.3
College of Podiatric Medicine	1	0.6
University Libraries	1	0.6
Other/Interdisciplinary	4	2.3
Faculty status (n=177)		
Tenured	76	42.9
Non-tenure Track	46	26
Tenure Track	43	24.3
Not applicable	12	6.8
Library services used (n=130)		
Reference Desk	60	46.2
Consultation with Subject Librarian	55	42.3
Digital Commons	46	35.4
Statistical Consulting	33	25.4
Special Collections & Archives	24	18.5
Student Multimedia Studio	8	6.2
Spark	3	2.3
Map It	0	0.0
Other	14	10.8

Survey respondents were asked what other library services they have used in the past, and could select as many as were applicable to them. The largest proportion reported assistance from the reference desk or direct consultation with a subject librarian. The subject librarian answer does support the notions addressed in both Wittenberg, Sackmann and Jaffe (2018) and Goben and Nelson (2018) in providing additional education to subject librarians on this topic. If many researchers are currently in active consultation with subject librarians, this approach would be most practical to address research data needs through the existing, regular service points.

Volumes and Types of Data Generated

In our sample, we found that 57% (163) of faculty indicated they generate data always or most of the time at our institution, and 75% (159) generating between one to five datasets annually (see Table 2). The majority of surveyed faculty indicated that they generate less than 1TB of data per year, with 40% of these indicating less than 1GB.

Table 2. Frequency and volume of datasets generated in a year.

	n	%
How often do you generate data? (n=286)		
Never	28	9.8
Rarely	39	13.6
Sometimes	56	19.6
Most of the time	74	25.9
Always	89	31.1
Number of datasets generated in a year (n=212)		
0	28	13.2
1	73	34.4
2-5	86	40.6
6 or more	25	11.8
Typical size of data generated in a year (n=209)		
Don't know/Not sure	44	21.1
Less than 1 GB	83	39.7
1 GB – 1 TB	74	35.4
More than 1 TB	8	3.8

Table 3 provides a view of what kinds of data the researchers at our institution are most commonly dealing with through their work, with quantitative data having the highest representation at 70.9%. One point that our team found interesting is that audio was rather high on the list, accounting for 38% of reported data types from the surveyed faculty. Framing the types of data most commonly produced in research is helpful in designing services to complement these particular data types. For example, our library staff may decide to focus on audio as a place to address further education for library faculty and staff, such as identifying local transcription services or exploring automated transcription services.

Table 3. Data type prevalence (n=203 responding).

Data Type	n	%
Quantitative	144	70.9
Databases	79	38.9
Audio	78	38.4
Non-digital text	68	33.5
Digital text	59	29.1
Digital images	56	27.6
Video	54	26.6
Metadata	25	12.3
Bio samples	23	11.3
Other	23	11.3
Geospatial	18	8.9
Non-digital images	18	8.9
Electronic lab notebooks	11	5.4
Gene sequences	4	2.0

Data Management Planning

The middle section of the survey contained questions related to long-term data management planning (see Table 4). We found that very few faculty were actively engaging in data management planning: only 24.6% had developed a data management plan in the last five years.

Additionally, faculty perceptions about long-term storage and access of data after a project or grant period were often idealistic – 42.9% said that the lifespan of their data was “as long as possible” – which may not always be practical or sustainable.

These findings indicate a current knowledge gap at our institution, and as such, are one of the main identified areas to focus initial attention by way of the internal working group and outreach/education initiatives.

Table 4. Data management planning practices.

	n	%
Developed data management plan in last 5 years (n=183)		
Yes	45	24.6
No	111	60.7
Don't know/Not sure	27	14.8
Ever published data alongside article (n=178)		
Yes	44	24.7
No	126	70.8
Don't know/Not sure	8	4.5
Ever used copyright with published data (n=176)		
Yes	14	8.0
No	103	58.5
Don't know/Not sure	59	33.5
Typical lifespan of data (n=205)		
A short, set amount of time (Less than five years)	51	24.9
As long as possible	88	42.9
Unknown/Hadn't considered	55	26.8
Other (Please explain)	11	5.4

Practices related to publication of research data were still very limited:

1. Only 24.7% had ever published data alongside an article.
2. Only 8.0% had ever used a copyright license with published data.

Methods of Backup and Storage

Consistent with prior studies, the most popular form of storage and backup for research data was a form of physical media such as disks, tape, hard drives, or USB drives (see Table 5). Within all colleges from which faculty answered this question, such physical media were either the top choice or tied for the top choice for data storage after project conclusion. From personal anecdotes, we surmise that this solution reflects what is either most available to the researcher without consulting outside units or is simply how data is stored during the project and remains so after the completion of the project or adjoining research paper.

Table 5. Strategies used to protect data from corruption or loss after the project has concluded.

Question	Yes		Total
	n	%	
Copies of datasets are saved on disk, USB, tape or hard drive	180	93.8	192
Copies of datasets are saved on web-based or cloud server	124	75.2	165
I make multiple copies in multiple locations (onsite or offsite copies)	96	64.4	149
Copies of datasets are saved on university server	89	56.3	158
Backup files are manually generated	87	62.6	139
Copies of datasets are saved on personal server	65	42.2	154
Backup files are automatically generated	60	42.9	140
Copies of datasets are stored in a data repository or archives	39	29.1	134
Other (Please specify)	8	17.4	46

Under 30% of the faculty had stored copies of their datasets in a data repository or archive. Interestingly, the proportion of faculty respondents using web-based or cloud servers was quite high at 75%, whereas the proportion of respondents saving data on a university server was about half. Personal hard disks and web-based storage permit a degree of control not present with university-owned servers; faculty may feel that they have easier access to data if it is stored on their own devices or cloud accounts.

In the open-ended responses, we found that surveyed faculty at our institution were often left to find their own storage methods for research data (both during and after research projects). This included a heavy use of external media, such as hard drives or thumb drives, or cloud storage (Dropbox and Google Drive being the most prominent). There is an immediate faculty need for a solution that allows easy data sharing, with options to restrict usage based on existing requirements or confidentiality needs. A system with embargo potential would also be of value for some researchers at Kent State University.

Web-based or cloud servers were the second highest choice for the College of the Arts, the College of Arts and Sciences, and the College of Education, Health, and Human Services. Faculty from the College of Communication and Information showed the highest rate of adoption for cloud storage to protect data after the conclusion of a project. Web-based or cloud servers were the second highest choice for the College of the Arts, the College of Arts and Sciences, and the College of Education, Health, and Human Services.

Barriers to Sharing

Respondents could select multiple barriers to sharing from a list (see Table 6). Overall, confidentiality requirements were overwhelmingly the biggest barrier to sharing data (64.7%). Other barriers selected by more than 25% of respondents, were lack of mechanism to share the data (28.2%), insufficient time to make data available (27.6%), lack of funding (27.1%), and the potential for data to be misinterpreted or misused by others (25.9%). When broken out by college, confidentiality was the top barrier for all colleges except for the College of the Arts; among the faculty in that college, the top barrier was lack of time.

Table 6. Barriers to sharing (n=170 responding).

Barriers to sharing	n	%
Confidentiality requirements	110	64.7
Lack of mechanism to share the data	48	28.2
Insufficient time to make data available	47	27.7
Funding	46	27.1
Potential for data to be misinterpreted or misused by others	44	25.9
Lack of mechanism to receive citation or credit once data are public	35	20.6
Opinion that the data shouldn't be made available	29	17.1
Lack of rights to make the data public	29	17.1
Opinion that people don't need the data	28	16.5
Inability to gain appropriate intellectual property rights protection	21	12.4
Lack of standards (data or metadata)	20	11.8
Other	18	10.6
Sponsor requirement regarding sharing	17	10.0

Discussion

RQ1: What questions should libraries ask of faculty when developing data management services? What sort of data management services would be the most appropriate for our faculty?

Our survey instrument contained questions about the type and volume of data generated; data management roles within the research team; data ownership and rights; long-term storage; and barriers to sharing. Our survey found that education was the most important service that the library could provide to faculty. Librarians can partner directly with faculty who supervise graduate assistants, and provide training on proper

data management practices. They can also provide feedback to faculty who are developing data management plans.

When conducting future research that aims to examine data sharing with outside researchers and scholars after a study is complete, researchers should make a distinction between data-sharing among co-researchers and data-sharing for re-use. Our survey items addressed data management from collection to long-term storage and/or destruction, yet we perceive from write-in answers and follow-up consultations that many respondents were focussed on data practices and sharing among research team members.

Though many respondents were aware of limitations to data-sharing for re-use, those who had chosen to share data most often opted to work with subject-specific data repositories that have mechanisms in place to ensure the long-term safety, storage, and accessibility of data files.

RQ2: What are the knowledge gaps among faculty about data management that the library could help fill?

The survey indicated that there is a huge gap present around data management planning, with 75% of surveyed faculty indicating they had either no plan in place, or had not considered the notion of data management. Most of the questions about data management practices included a “Don’t know/Not sure” option, and many individuals indicated this uncertainty. Data copyright had the greatest rate of uncertainty (33.5%, Table 4), followed by size of research data generated in year (21.1%, Table 2); developing a data management plan for any research projects in last five years (14.8%, Table 4); expected lifespan of data (5.4%, Table 4); and publishing data in conjunction with an article (4.5%, Table 4). These findings indicate a current knowledge gap at our institution, and as such are identified areas to focus attention by way of the internal working group.

Responses to the “Barriers to sharing” question, vocalized a need for a storage and access solution that is be robust, easy to use, and has a feature to easily share data, with options to restrict usage based on existing restrictions or confidentiality needs. Barring confidentiality requirements, time and funding were cited as the top reasons why researchers are not currently engaging in better research data management practices, and this is an area in which we perceive the library to be best situated to assist researchers.

Several faculty mentioned that they were self-trained on data management practices, and were now in the position of needing to train graduate students. One particularly notable extended response said that

“... I need better solutions and better training for students on good data mangement [sic] and documentation practices - but I was never trained in any of this myself.”

Through follow-up consultations with some survey respondents, we found that some disciplines, such as Biology and Library Science, were found to have some courses and workshops on research data available (or soon to be available), but that the majority of disciplines did not have inlets for budding (or more advanced) researchers to build and develop these skill sets. Therefore, this area is perhaps the place where well-defined research data support within the library could be of the most benefit for KSU researchers. Basic and advanced short library instruction courses could address aspects for researchers to develop skills around data management.

Additionally, researchers who trained and supervised graduate assistants noted that data-sharing practices involving constantly changing research assistants was a source of difficulty in consistency and continuity of practice around data management. We plan to

address this in future workshops and classes to assist researchers to develop better practices.

Overall, the general attitude of responses showed that faculty were more concerned with research data management during the research process than ‘long-term’ data management, the latter being an area where the library (particularly digital librarians with existing skill sets around digital media management) can be of assistance.

RQ3: How can the library identify collaborators on campus to participate in the development of data management services?

The process of developing and conducting the survey helped us to organically discover both institutional partners and faculty partners.

Within the survey, several faculty expressed interest in talking about topic of data management with us further. These faculty members agreed to meet with us informally to talk about their data management needs.

By way of both the survey and other meetings over the past year, the library has found two outside units currently working with researchers on aspects of the data management process. The Office of Research assists with grant applications and monitors research compliance, and Information Services provides systems support. In the short term, a collaborative working relationship has already increased referrals between the units, and in the long term, these connections can help to develop a comprehensive set of new services.

As we discovered in the analysis of the survey responses, the most candid write-in responses reflect many areas for our team to direct resources. In particular, we must strive to better communicate existing services for researchers, and consistently seek feedback through open library sessions. We must also strive to share feedback between units, especially with the Office of Research and Information Services.

RQ4: Are there any differences in data management practices and attitudes between disciplines that libraries should take into account when developing data management services?

A question at the beginning of our survey defined ‘research data’ as “the recorded factual material commonly accepted in the scientific community as necessary to validate research findings.” However, there were differing notions of the general concept of data between disciplines. Because the survey sample included faculty in all colleges, some respondents outside of natural sciences and social sciences “analogized” their closest equivalent to research data. For example, a faculty member from the Arts said: “I create and photograph artwork but do not have a system for managing it at this time.” We believe that this is a basic education point could be an element incorporated into all of the new library initiatives to better define data.

In terms of differences between disciplines and research attitudes, one respondent from the humanities was critical of some of the wording of certain areas in the survey, saying that the use of “research teams” in some of the questions was biased in favor of practices in the sciences. If researchers intend to address data management practices across disciplines using a broad definition of data, then care must be taken to avoid language that could be inferred as applicable only to certain disciplines.

Some of the notable similarities across disciplines were preferences for the use of physical media and cloud storage rather than university Web servers or repositories; and the perception of confidentiality as the greatest barrier to sharing. We also observed the need to improve communication about existing services within the institution.

Conclusion

The survey conducted at Kent State gave the University Libraries a starting place to define current needs at the institution and has led to the creation of a cross-departmental working group to address these needs (in conjunction with the Office of Research and Information Services). Survey results indicate a clear need of basic data management services and also gave indication of some specific education and systems needs of current researchers. Additionally, getting some of the more informal anecdotes in some survey questions provided an outlet for survey participants to share some of their previous experiences and frustration in the lack of having the needed support around research data. These anecdotes proved to be most useful in follow up conversations with outside units to relay these experiences to individuals who work in areas that can best address these needs. In subsequent conversations with the Office of Research, researchers in need of consultation with research data management plans for grant applications are now being referred to a member of the research data working group at the library.

General data management education, and research data management (and adjoining practices) are the top priorities identified in the survey. The information collected in the survey will assist in planning out and implementing new library services, and we also anticipate a follow-up survey once new initiatives are underway, as a point to ensure this work is well guided and also get feedback for continuing to refine services and address any gaps as needed.

Acknowledgements

The authors wish to thank Kent State University Libraries for support of this initiative, and provided funding of gift cards as an incentive for survey participants. This funding was made available through faculty research funds for University Library faculty members.

References

- Carlson, J. (2010). The Data Curation Profiles Toolkit: User guide. *Data Curation Profiles Toolkit*. doi:10.5703/1288284315650
- Chen, H.L. & Zhang, Y. (2017). Educating data management professionals: A content analysis of job descriptions. *The Journal of Academic Librarianship*, 43(1), pp. 18-24. doi:10.1016/j.acalib.2016.11.002
- Goben, A. & Nelson, M.S. (2018). Teaching librarians about data: The ACRL Research Data Management RoadShow. *College & Research Library News*, 79(7), pp. 354-357. doi:10.5860/crln.79.7.354
- Faniel, I.M. & Silipigni Connaway, L. (2018). Librarians' perspectives on the factors influencing research data management programs. *College & Research Libraries*, 79(1), pp. 100-119. doi:10.5860/crl.79.1.100

- Frank, E.P. & Pharo, N. (2016). Academic librarians in data information literacy instruction: A case study in Meteorology. *College & Research Libraries*, 77(4), pp. 536-552. doi:10.5860/crl.77.4.536
- Funari, M. (2014). Research data and Humanities: A European context. *Italian Journal of Library, Archives and Information Science*, 5(1), pp. 209-238.
- Joint Information Systems Committee (JISC). (2018). Research data management toolkit. Retrieved from <https://rdmtoolkit.jisc.ac.uk/>
- Kong, N., Fosmire, M. & Dewayne Branch, B. (2017). Developing library GIS services for Humanities and Social Science: An action research approach. *College & Research Libraries*, 78(4), pp. 413-427. doi:10.5860/crl.78.4.413
- Patel, D. (2016). Research data management: A conceptual framework. *Library Review*, 65(4/5), pp. 226-241. doi:10.1108/LR-01-2016-0001
- R Core Team. (2018). R: A language and environment for statistical computing. Retrieved from <https://www.R-project.org/>
- Sheehan, B. & Hogenboom, K. (2017). Assessing a patron-drive, library-funded data purchase program. *The Journal of Academic Librarianship*, 43, pp.49-56. doi:10.1016/j.acalib.2016.10.001
- Shen, Y. (2016). Strategic planning for a data-driven, shared-access research enterprise: Virginia Tech research data assessment and landscape study. *College & Research Libraries*, 77(4), pp. 500-519. doi:10.5860/crl.77.4.500
- Whitmire, A.L., Boock, M. & Sutton, S.C. (2015). Variability in academic research data management practices: Implications for data services development from a faculty survey. *Program: Data Technology and Applications*, 49(4), pp. 382-407. doi:10.1108/PROG-02-2015-0017
- Wittenberg, J., Sackmann, A. & R. Jaffe. (2018). Situating expertise in practice: Domain-based data management training for liaison librarians. *The Journal of Academic Librarianship*, 44, pp. 323-329.

Appendix

Survey Instrument

Q1.1 Our study will address the research data needs of faculty and introduce the potential of long-term access and preservation solution at Kent State. The survey should take between 10 to 15 minutes to complete, depending on the time spent answering the more qualitative, free text questions. The Institutional Review Board (IRB) at Kent State University has approved this study.

The answers provided in this survey will be collected anonymously, and we have ensured the software will not collect any personal or identifying information unless willingly provided by the individual for the purpose of post-survey follow-up or to qualify for the incentive. Any identifying information will be deleted upon completion of the designated survey period. As an incentive, we will give away a set number of (120) \$5 Starbucks gift cards to randomly selected people who complete the survey and opt into the drawing. Any information collected for the purpose of the gift card incentive will be kept in a separate Qualtrics survey, and will not be tied to the survey results in any manner.

The information we gather in the survey will be used to survey research data needs at Kent State University and may also be used in scholarly publication(s). The investigators listed below will be the only persons with direct access to the information provided by participants through the survey software, Qualtrics.

If you are 18 years of age or older, have understood the statements above, and freely consent to participate in the study, click on the "I Agree" button to begin the experiment. Please feel free to contact one of the investigators listed below if any additional information is needed, or contact the Kent State University Institutional Review Board at (330) 672-2704. We very much appreciate your time and we value your input.

Q2.1 How often do you generate original data as part of your research? Research data is defined as: The recorded factual material commonly accepted in the scientific community as necessary to validate research findings.

- Always
- Most of the time
- Sometimes
- Rarely
- Never

Q2.2 In a typical year, how many unique, completed datasets do you generate?

- 0
- 1
- 2-5
- 6 or more

Q2.3 Indicate the types of data your research generates. We have included examples of common file or document types for clarification. **Select all that may apply.**

- Quantitative, tabular or structured data (CSV, Excel, SPSS, JSON)
- Geospatial data (vector and raster data, shapefiles, geodatabases)
- Digital databases (surveys, census data, government statistics)
- Unstructured text in a digital format
- Digital images (.tif, .jpeg, .jp2)
- Data about biological, organic or inorganic samples or specimens
- Digital gene sequences or digital renditions of biological, organic, or inorganic samples or specimens
- Audio recordings (analog or digital)
- Video recordings (analog or digital)
- Electronic lab notebook(s)
- Non-digital text (handwritten notes, sketches, figures, paper lab notebooks)
- Non-digital images (photographs, etc.)
- Metadata (.xml, .rtf, .txt)
- Other (Please specify) _____

Q2.4 What would be your best estimate to the size of the research data generated per year (in total)? Note: 1 Terabyte (TB) = 1,000 GB

- Less than 1 GB
- 1 GB - 1 TB
- More than 1 TB
- Don't know/Not sure

Q2.5 Considering the research data you have generated in the last five years, who owns the data, and will the data be freely available for users to access? **Select all that may apply.**

- I own the data, and it is freely available for public use
- I own the data, but there are restrictions on use
- I am part of a team that created the data, and it is freely available for public use
- I am part of a team that created the data, but there are restrictions to use
- Ownership is by another party (Please describe if possible)
- _____
- Unknown

Q2.6 Some grants require that data must be made available for a set amount of time. When thinking about the research data you've generated in general, how long do you expect for it to be usable? That is, what is the "lifespan" of your data?

- A short, set amount of time (Less than five years)
- As long as possible
- Unknown/Hadn't considered
- Other (Please explain) _____

Q2.7 Please indicate which of the following strategies you use to protect your data from corruption or loss once the project has concluded.

	Yes	No
Copies of datasets are saved on disk, USB, tape or hard drive	<input type="radio"/>	<input type="radio"/>
Copies of datasets are saved on personal server	<input type="radio"/>	<input type="radio"/>
Copies of datasets are saved on university server	<input type="radio"/>	<input type="radio"/>
Copies of datasets are saved on web-based or cloud server	<input type="radio"/>	<input type="radio"/>
Copies of datasets are stored in a data repository or archives	<input type="radio"/>	<input type="radio"/>
I make multiple copies in multiple locations (onsite or offsite copies)	<input type="radio"/>	<input type="radio"/>
Backup files are automatically generated	<input type="radio"/>	<input type="radio"/>
Backup files are manually generated	<input type="radio"/>	<input type="radio"/>
Other (Please specify)	<input type="radio"/>	<input type="radio"/>

Q2.8 On your research team, who typically performs the tasks associated with data management? **Select all that may apply.**

	PI/Project Director	Graduate students	Internal (KSU) Researchers	External research staff	IT staff	Other	Not applicable
Data collection	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Data documentation (metadata creation)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Data cleaning (QA/QC)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Backing up data	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Data analysis	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Data storage and organization	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Data sharing outside of your research group	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Data archive or long-term storage	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Data disposal/destruction	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Q2.9 In the last five years, have you developed a data management plan for any of your research projects or related activities? A data management plan is a formal document that outlines what you will do with your data during and after you complete your research.

- Yes
- No
- Don't know/Not sure

Q2.10 Can you provide any details of this data management plan below. Such as: Name of tools used, plan to store and migrate data, etc.

Q2.11 What is your current storage and management solution for your research data? Please elaborate and describe below, particularly if this data management solution meets your data storage and management needs, and if it is easy to utilize.

Q2.12 How often do you share your data using the following methods?

	Always	Often	Sometimes	Rarely	Never	Unsure/Don't Know
Data are made publicly available on the principal investigator's personal website, or a project-specific website (non-University)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Data are made available through my organization's website (department, unit, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Data are made available through KSU's institutional repository, Digital Commons	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Data are made publicly available through a regional, national, or global network or repository (ICPSR, DataOne, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Data are made publicly available on a third-party platform (Dryad, figshare, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Data are not made publicly available, but I respond to individual requests	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Data are not made publicly available	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Other (please specify)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q2.13 Have you ever published research data in conjunction with an article? That is, have you ever made the “raw” data from a research study available alongside a publication?

- Yes
 No
 Don't know/Not sure

Q2.14 Please indicate which of the following limits the sharing of your data outside of your research team? Select all that may apply.

- Confidentiality requirements (privacy, human subject data, etc.)
 A lack of funding
 A lack of standards (data or metadata)
 An opinion that people don't need the data

- An insufficient time to make data available
- A lack of mechanism to share the data
- An opinion that the data shouldn't be made available
- A sponsor requirement regarding sharing
- A lack of rights to make the data public
- An inability to gain appropriate intellectual property rights protection
- The potential for data to be misinterpreted or misused by others
- A lack of mechanism to receive citation or credit once data are public
- Other (Please specify) _____

Q2.15 In addition to copyright, some creators use copyright licenses that permit others to reuse their works under certain conditions. Examples of these types of licenses include Creative Commons or the GNU General Public License.

A Creative Commons (CC) license is one of several public copyright licenses that enable the free distribution of an otherwise copyrighted work.

Is any of your data covered under such a license?

- Yes
- No
- Don't know/Not sure

Q3.1 What is your faculty status?

- Tenured
- Tenure Track
- Non-tenure Track
- Not applicable

Q3.2 What is your faculty rank? Please select the option that is most similar to your title.

- Full Professor
- Associate Professor
- Assistant Professor
- Senior Lecturer
- Associate Lecturer
- Lecturer
- Instructor
- Resident
- Adjunct
- Other (Please specify) _____

Q3.3 Please select your college from the list below.

- College of Applied Engineering, Sustainability and Technology
- College of Architecture and Environmental Design
- College of the Arts
- College of the Arts and Sciences
- College of Business Administration
- College of Communication and Information

- College of Education, Health, and Human Services
- College of Nursing
- College of Podiatric Medicine
- College of Public Health
- Honors College
- University Libraries
- Other/Interdisciplinary (Please specify) _____

Q3.4 Please select the campus where you are based.

- Kent
- Ashtabula
- College of Podiatric Medicine
- Columbiana
- Geauga
- Stark
- Trumbull
- Tuscarawas
- Other (Please specify) _____

Q3.5 What Kent State library services have you used in the course of your research?
Select all that may apply.

- Research consultation with a librarian who specializes in my subject area
- Reference Desk
- Digital Commons
- Special Collections & Archives
- Student Multimedia Studio
- Statistical Consulting
- Map It!
- Spark Innovation Studio
- Other (Please describe) _____

Q4.1 Would you like to be contacted about how the library can assist you with data management and curation?

- Yes (you will be redirected to another survey to enter your contact information)
- No

Q4.2 Would you like to be entered in the drawing for a Starbucks gift card incentive? (We will only contact you if you have been selected to receive a gift card.)

- Yes (you will be redirected to another survey to enter your contact information)
- No