Data in context: Using case studies to generate a common understanding of data in academic libraries.

Suggested citation: Swanson J, Rinehart AK (2016) Data in context: Using case studies to generate a common understanding of data in academic libraries. Journal of Academic Librarianship 42(1): 97–101. https://doi.org/10.1016/j.acalib.2015.11.005

Preprint Version

Published Journal Article:
Swanson J, Rinehart AK (2016) Data in context: Using case studies to generate a common understanding of data in academic libraries. Journal of Academic Librarianship 42(1): 97–101.
https://doi.org/10.1016/j.acalib.2015.11.005

Title: Data in context: using case studies to generate a common understanding of data in academic libraries

Authors:
Juleah Swanson [a*]
Amanda K. Rinehart [b]

[a] Acquisitions Librarian for Electronic Resources, Ohio State University, 1165 Kinnear Road, Columbus, OH, 43210; Present Address: Head of Acquisitions Services, University of Colorado Boulder Libraries 184 UCB, 1720 Pleasant Street, University of Colorado, Boulder, CO 80309-0184

[b] Data Management Services Librarian, Ohio State University, 18th Ave Library, 175 W. 18th Ave, Columbus, OH 43210

***Corresponding author:** Juleah.Swanson@colorado.edu, 1+303-492-9632

Abstract:

As new expectations emerge in librarianship, librarians find themselves engaging with researchers throughout the entire research process. This includes during early stages, when research outputs are in their infancy. This shift means that any librarian might be faced with a ‘data question’ and be able to assist without necessarily being a ‘data’ expert. As libraries approach professional development in this field, additional difficulties occur as data cannot be easily understood without context. Instead of attempting to comprehensively cover this broad, nuanced, and sometimes vague topic, the authors took a different approach. In order to place ‘data’ in definable contexts, the authors created local, real-world case studies to introduce this topic to the library. This article describes the professional development event, complete with case studies, their development, discussion questions, and observations. As faculty and staff answered guided questions, they self-identified the value of existing librarian capabilities such as the reference interview, information location, and referral systems. This enabled library faculty and staff from across the library to engage positively and proactively, without any extensive background in this field.

Keywords: data, professional development, data management, change, case studies, emerging librarianship

Published Journal Article:
Swanson J, Reinhart AK (2016) Data in context: Using case studies to generate a common understanding of data in academic libraries. Journal of Academic Librarianship 42(1): 97–101.
https://doi.org/10.1016/j.acalib.2015.11.005

Introduction

It's hard to get away from the word 'data' these days. It's in the news, on our cell phone plans, and ubiquitously used when discussing academic research. Like many words that are co-opted for other purposes, data often has different meanings based on context. The word 'data' is sometimes used synonymously with the word 'information', while at other times it may only refer to numbers. Some have found that the word 'data' has been largely limited to mean numbers (Nicholson and Bennett, 2015), while others apply the word to a variety of non-numeric resources (Schreibman, Siemens, & Unsworth, 2004).The inexactness of the term 'data' and what it refers to presents a challenge in learning about the rapidly changing area of data. Even for the purposes of this article, data is defined broadly as "units of information observed, collected, or created in the course of research" (Erway, 2013).When using a definition a broad as Erway's, the term 'data' cannot easily be understood without additional context.

The ubiquity of the term 'data' also contributes to academic librarians' confusion when new positions are created with 'data' in their titles; constituents call for 'data' services, and strategic plans become more 'data' focused. In reference interviews, librarians must take the time to define the term in order to begin to tackle the real need, but when speaking with colleagues, librarians may not have a common framework that leads to mutual understanding. The flexibility in meaning of the word 'data' can get in the way of understanding the nuances in this emerging field of librarianship.

As librarians, it can be tempting to create an authoritative vocabulary around emerging fields in an effort to reduce confusion. However, outside the librarian community, authoritative vocabularies are rarely widely adopted, and the changing landscape may make terms irrelevant in a short period of time. Without being prescriptive, how do we begin to parse out what data means in academic libraries and, more importantly, to academic librarian positions?

'Data' needs context. When the term 'data' placed within the context of specific types of research, disciplines, and needs, it is becomes much easier to understand how the word is defined in that context, regardless of what it may mean in a different context. The approach taken in this article is to highlight a different avenue in creating a common understanding of data and data services in libraries, one that does not rely on developing nuanced language, but rather places 'data' in definable contexts and precise scenarios.By illustrating what data looks like to various types of patrons and in various scenarios, a framework of what data means in academic libraries can begin to emerge. More specifically, case studies, generated from real-life scenarios, can be used to

Published Journal Article:
Swanson J, Reinhart AK (2016) Data in context: Using case studies to generate a common understanding of data in academic libraries. Journal of Academic Librarianship 42(1): 97–101.
https://doi.org/10.1016/j.acalib.2015.11.005

illustrate the experiences that librarians encounter in serving and resolving patron's data needs. Ultimately, case studies can begin to establish a shared understanding of data and data services in academic libraries.

This article details how the authors designed and deployed an event that empowered library faculty and staff to:

- explore how 'data' was affecting their positions,
- recognize and identify colleagues across the libraries that were struggling with common themes, and
- collectively resolve case studies involving common, but nuanced data scenarios.

Instead of attempting to comprehensively cover the broad, nuanced, and vague topic of 'data in libraries', the authors used case studies developed from real-world situations encountered at the Ohio State University (OSU) Libraries to promote positive and proactive engagement in this emerging field. These case studies were fundamental in simplifying the overwhelming and rapidly changing data landscape into something that each library faculty and staff person could readily understand, discuss, and resolve. In the process, we discovered that data issues arise all across the spectrum of academic librarianship, sometimes in unanticipated areas.

What these case studies highlight is that the shift that is occurring in academic librarianship may not be simply a shift from organization of information to organization of data, but instead a shift in where libraries and librarians embed themselves in the research process. As Bracke (2011) points out, "librarians traditionally manage the end of the research process. Library collections are the products of research, and librarians create services based upon usage of collections" (p70). Bracke (2011) goes on to suggest that, "researchers need to locate and access data collections, not simply articles, monographs and gray literature. Meeting these needs requires involvement earlier in the research process, ideally as participants in research design" (p71). Thus, a shift in the roles of academic librarians and the emergence of data in library services stems from a need to engage with researchers throughout the research process, including early stages of the process, when research outputs are in their infancy.

Data services in libraries

The history of data services in libraries varies according to patron needs, the type of library, and librarian interests, skills, and position descriptions. Nearly every academic library has been providing some type of data service ever since census data began being distributed to federal depository libraries. Kellam & Peter (2011) detail the emergence of the demand for libraries to support numerical data services in the

twentieth century, the resistance to this demand as digital formats emerged in the mid-century, and the eventual acceptance in the 1990s. Although they recommend both a social science and a 'hard' science data librarian to support numerical data services in the academic library, it is difficult to provide that level of staffing in the current economic climate.

More recently data services in libraries has been expanded to include more than numerical datasets for patrons and occasional storage. Current interpretation of library data services may include education and technical capacity for several components of the research lifecycle: planning for data management, assistance with data collection from text sources, advice on data documentation and metadata, demonstrating impact of dataset publication, and providing tools for dataset discovery, access, and preservation. Some of these activities have been designated as 'data management services'.

It is not uncommon to find that the more traditional activity of locating numerical datasets for patrons and newer data management services are often conflated. The ARL Spec Kit 334 delineated data management services as services that focus on two aspects of the research lifecycle: data management planning at the grant proposal stage and data archiving at the project's end (Fearon & Association of Research Libraries, 2013). These services have arisen in direct relation to new federal requirements for receiving funding. As such, many research libraries had not deployed any research data management services prior to 2010, when new federal requirement were first debuted (Fearon & Association of Research Libraries, 2013; Tenopir, Birch, & Allard, 2012). Examples of library support for data management includes providing support for crafting data management plans, guidance on data management throughout the research lifecycle, and dataset archiving and dissemination. As of 2013, most (61-87%) of ARL libraries offered services surrounding mandated data management plans and 74% of libraries offered some sort of data archiving services (Fearon & Association of Research Libraries, 2013). An ideal representation of complementary library services, including data and more traditional services, has been detailed by the University of Central Florida Libraries Research Lifecycle Committee (2012).

In conjunction with these new services, some institutions have begun to offer professional development opportunities for their librarians. As early as 2012, Jisc supported several initiatives to upskill liaison librarians (Cox, Verbaan, & Sen). The University of Michigan Library conducted a three-part curriculum for their librarians that involves general data concepts, deep dives into specific areas, and advanced workshops (Martin, 2014; Martin & Oehrli, 2015).

Published Journal Article:
Swanson J, Reinhart AK (2016) Data in context: Using case studies to generate a common understanding of data in academic libraries. Journal of Academic Librarianship 42(1): 97–101.
https://doi.org/10.1016/j.acalib.2015.11.005

Involvement of all librarians

The impact of data and data services in academic libraries is not limited to specialized librarians, such as data management or data services librarians. Librarians in both public and technical services are providing services that support researchers' emerging data needs. In fact, a survey in 2012 by Tenopir, Birch, and Allard, found that it was twice as likely for institutions to reassign existing staff to provide research data services as to hire a person. As well, only 12 out of 160 academic libraries that offered research data services had a specialist librarian dedicated to the task. It was far more common for individual discipline librarians and staff to provide research data services (Tenopir, Birch, & Allard, 2012).

Even librarians with less of a public-facing role, than their subject, reference, or liaison librarian counterparts, are becoming involved in data issues and services in their line of work. For example, acquisitions librarians and electronic resources librarians are finding expanded roles in negotiating license agreements to include terms that enable researchers to data and text mine library licensed content (Williams, et. al., 2014, & Ruttenberg, 2013). Collection development librarians are finding themselves asked to purchase more datasets and subsequently to define how those decisions are guided (Michigan State University, 2014). What the emergence of data-focused services suggests for all librarians is a deeper level of engagement with students, faculty and researchers in order to best support the nuances of individual data needs.

Data services in libraries do not necessarily require new skill sets or more technical acumen, but rather commands broader application of the existing competencies of librarians. "Along with the specialty area of data librarianship, the broader information profession field must be aware of growing data-based research and issues involved in its storage, processing, and use" (Reinhalter & Wittmann, 2014, p364). A little guidance on both resources and how established librarian skills translate makes the realm of research data management more accessible to all librarians.

Materials and Methods

The Case Study Activity

At Ohio State University Libraries, librarian engagement with academic disciplines and researchers is a priority. In the fall of 2011, a cross-unit group of librarians was charged with identifying strategic opportunities for subject librarians to be involved with the research, teach and learning, services, and outreach goals of the university. Expectations for competencies and best practices for subject librarians were outlined in the Framework for the Engaged Librarian (Engaged Librarian Framework Working

Published Journal Article:
Swanson J, Reinhart AK (2016) Data in context: Using case studies to generate a common understanding of data in academic libraries. Journal of Academic Librarianship 42(1): 97–101.
https://doi.org/10.1016/j.acalib.2015.11.005

Group, n.d.) and a forum was created to discuss, disseminate, and provide professional development to librarians in the competency areas and best practices (Ohio State University Libraries, 2014). The case studies presented in this article originated as an activity at an Engaged Librarian Forum (ELF) in September 2014 that focused on the emerging role of data in the OSU Libraries. Both faculty and staff are welcome to attend ELF sessions as they are open to the entire library.

For the ELF on data in libraries, the authors crafted an hour and a half long session that incorporated a short overview lecture on data, a breakout session into small groups to read, analyze, and discuss the case studies, and a review of the case studies among all participants. This allowed for attendees to participate in multiple forms of learning, discussion, and analysis, and to be highly engaged in the forum.

Preparation:
Prior to the ELF, the Data Management Services Librarian conducted an informal survey of library faculty to determine what their knowledge and interest in ‘data’. Specifically, faculty were asked how they defined ‘data’ and to rank their interest in various ‘data’-related topics. This informal survey provided context to the presenters and gauged the level of knowledge the audience may have on the topic.

To develop the case study activity, the Business Librarian, Head of Research Services, Acquisitions Librarian for Electronic Resources, and Data Management Services Librarian were all solicited to provide examples of interactions with patrons who had a data need. These librarians were selected simply based on the assumed likelihood they had come across a scenario with a patron who had a data need, though it is likely that many other librarians could have also provided examples as well. The examples collected from the librarians were then derived and developed into case studies, specifically written to tease out many of the nuances of each scenario. All cases came to the attention of the librarians through existing, traditional roles of librarianship, such as research consultations, meetings with faculty constituents, negotiating with library vendors, or responding to reference questions.

The Case Studies:
Five case studies were crafted for the forum, based on the varied real-life situations OSU Librarians had experienced.

After each case study, the following questions were listed for groups to discuss and report out on:

Published Journal Article:
Swanson J, Reinhart AK (2016) Data in context: Using case studies to generate a common understanding of data in academic libraries. Journal of Academic Librarianship 42(1): 97–101.
https://doi.org/10.1016/j.acalib.2015.11.005

1) Summarize some of the nuances within this case study
2) What resources or services exist within the libraries?
3) Does this situation need to be referred elsewhere? If so, who?
4) How might you assist the patron?
5) What can we do better as a library to support this case study?

To protect our patrons and make it a little more fun, de-identification included substituting patron names with child stars from the 1990s and early 2000s. Attendees were further challenged to identify what show made the child star famous and where they went to college. This playfulness underscored the exploratory nature of the exercise and helped to alleviate the stress of potential new obligations. There is no intended connection between the scenario in each case study, and the name of the child star used.

Below is a brief synopsis of each case study, however the full case studies (found in the Appendix) illustrate the complexity and nuances of each situation.

Case study #1 focuses on the story of a graduate research assistant who has been asked to find citation information for all articles published by faculty from within the Department of Communications. He compiles the data manually, saving each result individually as an HTML file. The corpus, however, needs to be compiled into a format that can be analyzed, and the graduate student does not know how to parse the data from the HTML files into an analyzable format.

Case Study #2 tells the story of an assistant professor who needs a dataset of metadata from tens of thousands of articles in order to test the effectiveness of software he has developed. The publisher he has contacted has agreed to provide the desired dataset, free of charge, but under the condition that the faculty member agrees to a royalty-free, non-exclusive license for the software that is developed based on the provided dataset.

Case Study #3 illustrates the common situation of a faculty member who is frustrated by the lack of basic knowledge and skills her graduate assistants have regarding managing data and wants to incorporate basic data management training into her courses.

Case Study #4 details the results of a graduate student who created her own script to crawl LexisNexis Academic in order to text mine thousands of newspaper articles for

Published Journal Article:
Swanson J, Reinhart AK (2016) Data in context: Using case studies to generate a common understanding of data in academic libraries. Journal of Academic Librarianship 42(1): 97–101.
https://doi.org/10.1016/j.acalib.2015.11.005

her dissertation research. LexisNexis shuts down institutional access to LexisNexis Academic upon detection of the unauthorize script crawling the database.

Case Study #5 is the report of a graduate student who sends three reference question e-mails asking to locate economic and financial data, ranging from data that exists in a database that the library does not subscribe to, to data that is not readily compiled into an existing dataset.

The event:
The Engaged Librarian Forum began with a short introduction of the definition of the word ‘data’, the research lifecycle, and results from the library faculty survey. From the survey, the library faculty ranked the following data-related topics from most interested to least: data information literacy, data visualization, locating data, the impact of data sharing, citing datasets, practical data management, data ownership and rights, data ethics, data management plan requirements, and metadata standards and schema. Volunteers to help conduct future workshops on these topics were solicited at the end of the event.

After the short (approximately 15 minutes) introduction, attendees were separated into five groups. There were approximately twenty-five attendees, allowing each group to have four to five members. While the majority of the attendees were subject librarians, most groups were a mix of positions, which promoted lively cross-unit discussions. Each group was given a scenario, a large writing surface and instructions to read through the scenario, discuss the attached case study questions, and to jot down notes or summaries of answers. At the end of approximately twenty minutes, the groups were asked to report out.

Published Journal Article:
Swanson J, Reinhart AK (2016) Data in context: Using case studies to generate a common understanding of data in academic libraries. Journal of Academic Librarianship 42(1): 97–101.
https://doi.org/10.1016/j.acalib.2015.11.005

Engaged Librarian Forum. Small groups discussing their respective Case studies.

Results and Discussion:

When the groups reported out on the analysis of their assigned case study, some common responses and themes began to emerge. While some of the proposed resolutions to the case studies involved consulting with or connecting researchers to campus partners or colleagues with non-traditional librarian roles, the pathway to resolving each case study was founded in core competencies of librarianship. For example, several groups noted the need to clarify the patron's ultimate goals, and two of the groups specifically noted the need to address the intent and purpose of the researcher's request. Understanding the needs of patrons, and getting to the core of that need is a fundamental librarian competency.

Connecting researchers to additional resources, whether they are books or online materials, or even other people, is also a well-trodden skill that frequently came up in the discussion. In particular, many academic libraries, like OSU Libraries, have a strong internal referral system, where questions and patrons are formally or informally referred to other experts within the library. While the intent of the case study activity was to empower all librarians to increase their level of comfort with data issues that they may come across, it also reinforced the practice of seeking out additional expertise when needed. So it was no surprise that some groups suggested consulting online resources within the libraries such as our customized version of the DMPTool (California Digital Library, Regents of the University of California), while other groups suggested referrals to librarian colleagues such as the data management services librarian (a popular response), the copyright librarian, or even the business librarian. In addition, the majority of groups recommended reaching out to external partners for guidance and

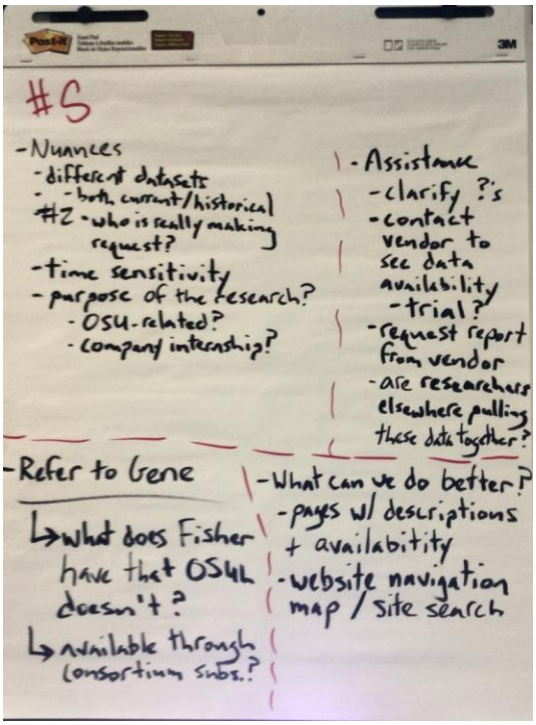
Published Journal Article:
Swanson J, Reinhart AK (2016) Data in context: Using case studies to generate a common understanding of data in academic libraries. Journal of Academic Librarianship 42(1): 97–101.
https://doi.org/10.1016/j.acalib.2015.11.005

expertise. These suggestions included campus resources, such as college-level grants administrators, and central and departmental IT staff.

Of particular concern among the groups were the themes of legal implications and restrictive access to data.These concerns ranged from potential violation of consortial license agreements to concerns regarding data creators' rights to the derivative datasets. Intellectual property considerations and possible impacts on the university were discussed. Three out of five of the groups noted other publishers, trial subscriptions, license add-ons, and consortial agreements as possible alternative means of procuring the requested datasets. As a whole, the larger group discussion contributed to a growing recognition of the legal complexities of data rights, ownership and access that presents unique challenges in supporting researcher needs.

All of the groups had suggestions for future library improvements. These included the need to develop education materials on best practices and standards, the development of service models to support new requests, and the need to integrate these new models with existing IT workflows. However, because each scenario was based on a real situation, the ELF facilitators and the librarians who were involved in each real-life situation were able to recap to the audience what actually transpired. Spending time discussing the real-life conclusion of each scenario allowed librarians to admit that less-than-ideal resolutions were acceptable and sometimes unavoidable. In general, the attitudes from librarians reflected recognition that the OSU Libraries has a collective responsibility to address data issues, that there are multiple routes to resolution, and that the group already possesses several applicable areas of expertise. While additional standards, best practices, and training would be beneficial, the case studies revealed that librarians can begin to tackle researcher's data needs without formal initiatives or programs that focus on data specifically.

Published Journal Article:
Swanson J, Reinhart AK (2016) Data in context: Using case studies to generate a common understanding of data in academic libraries. Journal of Academic Librarianship 42(1): 97–101.
https://doi.org/10.1016/j.acalib.2015.11.005



Engaged Librarian Forum. Notes from Case Study #5 small group discussion.
To be published as 1 column wide.

Conclusions:

All the case studies were based on real, local situations from across the library organization, presenting the reality that any librarian might be faced with a 'data question' in their work and any librarian can contribute to a resolution without necessarily knowing new concepts, techniques, or being a 'data' expert. The intent of presenting the issue of data services through cases studies was for attendees to leave the ELF feeling empowered to contribute to data librarianship, comfortable talking about these new situations and challenges, and better informed about existing resources. By placing context around the term 'data' in ways that were intimately familiar to librarians, and based on situations that they could recognize in their own work or the work of their colleagues, librarians could begin to understand what data and data services look like in academic libraries.

Whether adapting the case studies presented here, or replicating the process and gathering case studies from one's own institution, the activity presented in this article can be duplicated at academic libraries of all sizes. Because these case studies all arose out of traditional interactions with patrons, and have resolutions that do not necessarily require the assistance of a data services or data management librarian, this

activity can serve as a means to cultivate a local understanding of data needs as well empower all librarians to address these needs.

Since the beginning of any research process is an observation that leads to collecting evidence in the form of data, many academic librarians, regardless of their position, find themselves involved with research data. As this shift continues, most librarians may find themselves to be data librarians. The use of case studies, derived from real-life situations can demonstrate to librarians that our current practices are not that far removed from skills required to support data needs, and that we all *can* become data librarians.

Reference List:

Cox, A., Verbaan, E., & Sen, B. (2012). Upskilling Liaison Librarians for Research Data Management. *Ariadne,* 70. Retreived from http://www.ariadne.ac.uk/issue70/cox-et-al

Bracke, M. S. (2011). Emerging Data Curation Roles for Librarians: A Case Study of Agricultural Data. *Journal of Agricultural & Food Information, 12,* 1, 65-74.

Engaged Librarian Framework Working Group. (n.d.). *A Framework for the Engaged Librarian: Building on our Strengths*. Retreived from https://carmenwiki.osu.edu/download/attachments/37010332/Engaged+Librarian+Document.pdf

Erway, R., & OCLC Research. (2013). *Starting the conversation: University-wide research data management policy*. Dublin, Ohio: OCLC Research. Retrieved from http://www.oclc.org/content/dam/research/publications/library /2013/2013-08.pdf

Fearon, D., & Association of Research Libraries. (2013). *Research data management services*. ARL SPEC Kit 334. Washington, DC : Association of Research Libraries.

Nicholson, S., & Bennett, T. (2015). *Dissemination and discovery of diverse data: Do libraries promote their unique research data collections?* Poster session presented at Annual Association of College and Research Libraries Conference, Portland, OR.

Michigan State University. (August 2014). *Collection Development Policy Statement: Data Services (Numeric Data)*. Retreived from: http://libguides.lib.msu.edu/dataservicescollectiondevpolicy

Published Journal Article:
Swanson J, Reinhart AK (2016) Data in context: Using case studies to generate a common understanding of data in academic libraries. Journal of Academic Librarianship 42(1): 97–101.
https://doi.org/10.1016/j.acalib.2015.11.005

Kellam, L. M., & Peter, K. (2011). *Numeric data services and sources for the general reference librarian*. Oxford: Chandos Publishing.

Martin, Scott. (2014, March 6). *Deep dive into data* [web log post]. Retrieved from https://mlibrarydata.wordpress.com/2014/03/06/deep-dive-into-ecology-data/

Martin, Scott & Oehrli, Angela. (2015). *Diving into data: developing skills in research data services.* Poster session presented at Annual Association of College and Research Libraries, Portland, Oregon.

Ohio State University Libraries. (2014, October 14). *Engaged Librarians Forum*. Retreived from https://library.osu.edu/document-registry/docs/340/stream

Reinhalter, L., & Wittmann, R. J. (2014). The Library: Big Data's Boomtown: Edited by Rick J. Block. *Serials Librarian, 67,* 4, 363-372.

Ruttenberg, C. (2013). Technical Services Report: Report of the ALCTS Electronic Resource Management Interest Group. American Library Association Midwinter Meeting, Seattle, January 2013. *Technical Services Quarterly, 30,* 4.

Schreibman, S., Siemens, R., & Unsworth, J. ed. (2004). *A Companion to Digital Humanities*. Oxford: Blackwell. Retrieved from http://www.digitalhumanities.org/companion/

Tenopir, C., Birch, B., & Allard, S. (2012). *Academic libraries and research data services: Current practices and plans for the future; an ACRL white paper.* Chicago: Association of College and Research Libraries, a division of the American Library Association.

University of Central Florida Libraries Research Lifecycle Committee. (2012). *The research lifecycle at UCF* [Online Graphic]. Retrieved from library.ucf.edu/ScholarlyCommunication/ResearchLifecycleUCF.php

Appendix A

Engaged Librarian Forum Data Case Studies

September 18, 2014

Case studies authored by:
Juleah Swanson[1]
The Ohio State University Libraries

Original published in:
Ohio State University Knowledge Bank
http://hdl.handle.net/1811/70885

Case Studies: Patrons and Data

We have compiled five case studies involving student and faculty requests, use, or engagement with data in their research or study.

Cases are based on real situations that have been addressed by a number of librarians here at the OSU Libraries. You may recognize some details of these cases as your own, or you may be familiar with similar situations.

Events have been dramatized and all names have been changed from patron names to the names of former child stars of the 1990s and early 2000s who have since grown up and went on to college. *Bonus points for anyone who can identify what made the child star famous and where the child star went to college* ☺

Instructions:

- **Break up into five groups.**

---

[1] Current author affiliation and contact information:
Juleah Swanson
Assistant Professor, Head of Acquisition Services
University of Colorado Boulder Libraries

184 UCB, 1720 Pleasant Street
Boulder, CO 80309
juleah.swanson@colorado.edu

- **Read through your assigned case study.**
- **Discuss the case study questions as a group.**
- **Jolt down notes or a summary of your answers on the flipchart**
- **Report out findings to the room**

Case Study Questions

- Summarize some of the nuances within this case study
- What resources or services exist within the libraries? Does this situation need to be referred elsewhere? If so, who?
- How might you assist the patron?
- What can we do better as a library to support this case study?

Case Study #1

Under-caffeinated graduate research assistant Jonathan Taylor Thomas has been asked by his advisor in the Department of Communications to work on a special project to identify *all* the articles faculty members within Communications cite in their publications. He was advised to use Web of Science to look up the Department of Communications faculty publications. His advisor also taught him how to capture each Web of Science result as an html page.

Over the summer, Jonathan spent hours manually looking up thousands of citations in Web of Science and saving thousands of html pages containing the data he needed.

With two weeks left in the summer Jonathan's advisor took off on a well-deserved vacation to hike around the remote Isle Royale National Park with no access to the internet, cell service, or other comforts of modern civilization.

Jonathan is nearly complete with his research on Web of Science, but now needs to parse and compile all of the data from the saved html pages into a format that can be analyzed. He has no idea how to do this and doesn't know who to ask.

In a moment of desperation, Jonathan calls up the OSU Libraries Reference Desk because he figures, librarians are smart people, they can help.

You pick up the phone when Jonathan calls.

Case Study Questions

- Summarize some of the nuances within this case study
- What resources or services exist within the libraries? Does this situation need to be referred elsewhere? If so, who?
- How might you assist the patron?
- What can we do better as a library to support this case study?

Case Study #2

Assistant Professor Haley Joel Osmond is working to build a sophisticated article recommendation and analytics system built on top of rich article metadata. It would be something like a read-alike that would present something similar to a user:

> "*If you liked this article on Carbon Nanotube-Based Cells for Yield Analysis maybe you'll like this article on System Level Benchmarking with Yield-Enhanced Standard Cell Library for Carbon Nanotube VLSI Circuits."*

Awesome.

Assistant Professor Osmond needs metadata from tens of thousands of articles to further refine and develop his software. He has heard from colleagues at other universities working on similar research that they have been using metadata from the Association for Computing Machinery (ACM).

Assistant Professor Osmond is smart, so he knows he could easily create a script that would likely be undetectable to the security software that ACM has in place and crawl their website for the article metadata.

But, he also knows that ACM likely has all this metadata in a retrievable data file on their back end, which they could easily send along to him.

Osmund is told by a colleague that he needs to talk to the Libraries in order to work with ACM to acquire the metadata he needs. So he approaches you.

You speak briefly to ACM about the request for a file of ACM article metadata. As a not-for-profit, association publisher, they are more than willing to provide the data to the faculty member for free!

ACM just requests that the faculty member sign a research agreement that mostly stipulates how the data is to be accessed, used, and distributed, but also includes the following clause requesting the faculty member grant ACM a royalty-free, non-exclusive license for the software that is developed based on the ACM article metadata:

> *"Provide the ACM a copy of any software developed in the course of the research that relates to the management and analysis of the ACM-supplied Material, with a royalty-free, non-exclusive, non-transferrable perpetual license to use it as is, or with modifications, solely in the context of ACM's Digital Library service."*

Assistant Professor Osmund is open to the idea of a letting ACM use the software he is developing with their metadata, but is unsure about whether he can sign the agreement. He approaches you again for consultation.

Case Study Questions

- Summarize some of the nuances within this case study
- What resources or services exist within the libraries? Does this situation need to be referred elsewhere? If so, who?
- How might you assist the patron?
- What can we do better as a library to support this case study?

Case Study #3

Faculty member Danica McKellar manages a small lab in the College of Veterinary Medicine. Teaching and conducting research at OSU is a second career, as she spent many years in industry managing a much larger lab with many long-term lab assistants.

She has recently become frustrated with the high turnover of graduate assistants in her lab and the lack of basic knowledge and skills that her new graduate assistants have on managing data. Files are poorly organized. Data is being saved all over the place, in secure and unsecure locations. And her students do not seem to understand the importance of formatting, saving, and storing the data for future use and long-term preservation.

She wants to introduce basic data management training into her courses to better prepare her students for current or future work in the lab or field.

You have a meeting scheduled with Danica to discuss her current lab work and research. In this meeting she brings up her frustration and concern over the lack of basic data management skills that she sees among her students.

Case Study Questions

- Summarize some of the nuances within this case study
- What resources or services exist within the libraries? Does this situation need to be referred elsewhere? If so, who?
- How might you assist the patron?
- What can we do better as a library to support this case study?

Case Study #4

Enterprising graduate student Andrea Barber needs to text mine millions of news articles from the past twenty years for her dissertation in Political Science. She is first looking for specific word pairings that exist in news articles and then is looking at the greater context around those word pairings to identify themes and tone, in order to determine changes over time.

Her ideal source for accessing the millions of news articles is the LexisNexis Academic database because it has over 3,000 global newspapers and she was easily able to create a Python script for LexisNexis that crawled for her word pairings and retrieved the data she was looking for.

Andrea launches her Python script, sits back and watches her dissertation data flow in. She is on her way to PhD success and dissertation fame. Maybe she'll even get a book deal from a university press like Harvard or NYU.

Thirty minutes in to running her Python script, the script fails. Or at least it stops running and she's not quite sure why. Andrea sets it aside to try again tomorrow.

The next day, Andrea receives a stern e-mail from the OCIO's Senior Security Engineer informing her that per the OSU Libraries license agreement with LexisNexis, the scripted search she was conducting is not allowable, and has resulted in a temporary suspension of LexisNexis Academic for the entire University.

Andrea writes back:

"I apologize - I didn't know that scripted searches were not allowed. I will stop it immediately and remove the scripts from the machines.

That said, I wonder if I could have the email or phone number of your contact at LexisNexis? The thing is, I need to text-mine millions of news articles in order to do my dissertation. It would be impossible to download everything manually (I need to perform over 1.4 million searches on the LexisNexis database). So I'd like to talk to them, explain my situation, and see if they can make an exception in my case."

The OCIO Senior Security Engineer knows this is something the Libraries should handle, but doesn't know who. He forwards the e-mail along to you for assistance.

Case Study Questions

- Summarize some of the nuances within this case study

- What resources or services exist within the libraries? Does this situation need to be referred elsewhere? If so, who?
- How might you assist the patron?
- What can we do better as a library to support this case study?

Case Study #5

It's Monday morning, it's raining, and you've only had one cup of coffee instead of your requisite two. Things are moving slowly as you enter your office and fire up the computer to check your e-mail.

> ***Unread Message #1***
>
> *From: watson.754@osu.edu*
> *Sent: Sunday 11:59 PM*
> *Subject: NAV data*
>
> *Hi I'm a PhD candidate in Political Science and I'm looking for price and net asset value (NAV) data for close-ended, delisted exchange-traded funds (EFTs). Does the library have this data?*
> *Thanks!*
> *Emma Watson*

After some swift flexing of your librarian-ing skills you find that this data is available in a module of Global Financial Data that OSU Libraries does not currently subscribe to. Last time you check it was about $10,000 to add on the module.

You also find that the price data is is available through the Datastream terminal at Fisher, but not the net asset value data. The Net Asset Value data is available through Morningstar Direct, another terminal based product that OSU Libraries does not currently subscribe to.

You add this request to your to-do list for the morning and move on to the next e-mail.

> ***Unread Message #2***
>
> *From: watson.754@osu.edu*
> *Sent: Monday 12:02 AM*
> *Subject: Price index for recyclables*
>
> *Hi Again!*
> *I'm also working with a faculty member at Fischer so I also need the price index listing market prices for common recycled commodities, such as recycled aluminum. I'd like current and historical prices.*
> *Thanks!*
> *Emma*

Again, you flex your librarian-ing-skills but this time come up empty-handed. Nothing you find perfectly matches this request. There's the Global Financial Data database that OSU subscribes to which has dozens of commodity indexes, but nothing specifically on recycled materials. There's the USGS Mineral Industry Survey available online which contains buying prices for scrap metal, and there is the American Metals Market, which OSU does not subscribe, and you're not sure whether AMM even offers an institutional subscription.

You also put this e-mail on your to-do list to investigate further.

> ***Unread Message #3***
>
> *From: watson.754@osu.edu*
> *Sent: Monday 12:09 AM*
>
> *Subject: Foreign Direct Investment by State*
> *Oh...and I also want Foreign Direct Investment (FDI) data in the US at the state level, state by state. This one is time sensitive. I need the data by the end of the week. Where can I find this?*
> *Emma*

Everyone seems to want Foreign Direct Investment data, so you have a good handle on where this data might exist. Each state's commerce page has data available on FDI, but that would require the student to look at the commerce sites for all 50 states. There's also the resource ForeignDirectInvestment.com that has some freely available datasets, though more enhanced reports and datasets requires purchase.

You also put this e-mail on your to-do list to investigate and to reply.

Case Study Questions

- Summarize some of the nuances within this case study
- What resources or services exist within the libraries? Does this situation need to be referred elsewhere? If so, who?
- How might you assist the patron?
- What can we do better as a library to support this case study?