Research as Design—Design as Research: Applying Design Thinking to Data Management Needs Assessment

Cinthya Ippoliti
Oklahoma State University, USA

Introduction
Data management has exploded into the world of higher education in the last five years and libraries have found themselves at the forefront of this movement. Some have quickly developed infrastructure and support for every activity ranging from data storage and curation to project management and collaboration, while others are now just beginning to think about addressing the needs of their researchers.

Regardless of which end of the spectrum they identify with, libraries are still seeking to understand the research landscape both within their organizations and on a national level and define their role in the process.

Institutional context
Oklahoma State University (OSU) has been classified as a research university with high research activity. The total student population at OSU’s primary campus in Stillwater, Oklahoma, is approximately 24,000, with an undergraduate population of approximately 20,000. Edmon Low Library is the primary library on campus and is used by undergraduates, graduates, and faculty.

The Research and Learning Services (RLS) division of the library includes an associate dean, ten librarians (six of which are liaisons, one is the director of the liaison program, two are directors of undergraduate/graduate student outreach, instruction and research respectively), and a first year experience librarian. The division also includes an instructional developer and a faculty coordinator for undergraduate outreach as well as two graduate student assistants and an intern.

The impetus for this project emerged out of a recent library reorganizational effort that sought to:

1. Develop strong internal and external partnerships in order to facilitate outreach efforts as well as build momentum and support for collaborative projects and initiatives.
2. Shift collection development duties from buying resources to reach a specific dollar amount to analyzing the collections and communicating with departments in order to better focus energies and expertise into building strong collections that can support the general needs of digital scholarship activities with targeted discipline-specific assets.
3. Provide systematic and ongoing internal training in areas such as data management and data information literacy, open access, tools such as Altmetric, and the new Framework for Information Literacy.
4. Update existing primary assignments to provide a way to determine our impact on university-wide strategic goals such as student learning and faculty productivity.
5. Literature review

Our study is not the first to address data needs of faculty, but it is the first to employ the design thinking analysis model for a qualitative-based process. Our results closely mirror those uncovered at Oregon State University where it was found that faculty generate broad data types, with relatively small data sets, and manage their own servers. In addition, we also had the same challenges in terms of raising awareness of our services and we found outreach to be a key component of our efforts. Jerry Sheehan, et al. discuss their two-pronged approach at Montana State University, which includes both surveys and interviews. Their findings of varying definitions and attitudes towards open data as well as the need for collaborative efforts at the campus level closely align with our own action items delineated below. The study conducted by Gu, et al. points to an interesting idea that has not appeared elsewhere in the literature, that is, the need to conduct a scan of existing services and resources prior to conducting interviews in order to help inform the conversation that will ensue during the interview and help raise further awareness of what is already available.
Similarly, the study conducted by Buys and Shaw at Northwestern University indicates that researchers struggle with “issues with long-term storage, data organization and management, knowledge of data management plans, and need for consultation and instruction.” Scaramozzino et al. discuss the interventions they implemented as a result of their data needs assessment to “broaden faculty awareness of data curation issues that span a wide array of disciplines.” Rolando, et al. continue this thread and speak to the need for libraries to develop “infrastructure, expertise, and services necessary to help researchers preserve and share their research data,” which is an important element to consider as part of a broad vision for the role of the library in data management policies and activities. Finally, Antell, et al. discuss the need for librarians themselves to be aware of both what skills they need in order to effectively collaborate with researchers as well as become aware of broader campus efforts in this arena. More specifically, they state that “science librarians have not come to a consensus about whether the data management role is a natural extension of their jobs, or a set of duties that would be better suited to librarians holding a different job title and assuming a different role within their respective organizations” amidst unfamiliarity regarding data management “assistance and initiatives on their own campuses or even within their own libraries,” which renders any type of needs assessment even more important in order to focus training and project collaboration based on campus priorities and appropriate expertise.

Study methodology
We worked with OSU’s Office of Institutional Research and Planning to identify new and tenure-track faculty who had not yet achieved tenure, which was a total of 259 members. We interviewed 31, or 12%, of the total number who could have responded to our invitation.

We obtained IRB approval to contact all new and tenure-track faculty in order to schedule individualized interviews. Our research goals were as follows:

1. What are faculty needs for research data support relating to data management plans, data analysis and storage, data information literacy, and research impact?
2. What is the library’s role in providing some or all of those services and resources?
3. Who else on campus is offering similar services and how can we create partnerships?
4. What training do librarians themselves need in order to provide quality services to support these needs?

Recruitment method (form and e-mails)
An e-mail was sent out to prospective participants:
You are invited to participate in a research project conducted by the OSU Library. We are requesting to schedule 30-minute interviews with you so that we can ask you about your research and data management needs in the hopes that this information will allow us to provide new services and improve existing services in these areas. Your responses will be kept completely confidential. No personal information will be collected except for your discipline so that we can sort the data by that variant. Participation is voluntary, and responses will be kept confidential. You have the option to not volunteer. Completion of the interview will be interpreted as your informed consent to participate and affirms that you are at least 18 years of age. If you have any questions about this research, please contact project investigator Cinthya Ippoliti. If you have any questions regarding your rights as a research subject or if you feel your rights have been violated during the course of this project, contact the IRB Office, 223 Scott Hall, 405-744-3377.

Faculty then had the option of filling out a short form to indicate their interest as well as their preferred date and time.

Interview format and survey questions
The interviewers consisted of a combination of the associate dean for research and learning services and the head of the library liaison program, as well as the respective liaisons for each faculty’s subject area. Interviews were scheduled depending on the faculty member’s availability, but the liaison was present whenever possible. The interviews were conducted in pairs, held in the faculty member’s office, and we collected hand-written notes that were later synthesized into an Excel spreadsheet where each of the questions represented the heading for a separate column paired with the answers for each faculty member. A sample portion of the spreadsheet is shown below.
We adapted the framing questions from the Data Curation Profile Toolkit\cite{10} and the University of Virginia Data Interview Protocol in the ARL Spec Kit 334.\cite{11} The entire group of library interviewers met prior to scheduling interviews to ensure we all understood the goals of the questions we were asking as well as the importance of the first step of the design thinking process that focuses on asking open-ended follow-up questions in keeping with the theme of developing an understanding of faculty challenges and needs.

1. What is your primary unit of affiliation?
2. Do any of your funding sources require you to draft a data management plan? If yes, how do you go about creating this plan?
3. What type of data do you generate?
4. Who performs the majority of the following activities (e.g., PI/director, graduate students, researchers, IT staff, other)? Data collection; data documentation (metadata); data cleaning; backing up data; data analysis; data storage and organization; data sharing outside of your research group; data archive or long-term storage; data disposal/destruction associated with your research?
5. Does personnel within your department/college typically provide technical support or assistance with your OSU research? If so, who and to what extent? If not, what do you do to get help?
6. Does your research include the analysis of data collection by others (also referred to as secondary data)? Do you acquire secondary data from public, non-public/restricted, other researchers’ data sources?
7. What has been the typical amount of digital data for a single project you have worked on in the past five years?
8. Do you generate metadata?
9. Could you please describe the system for version control that you have in place?
10. How often and how do you share your data with others?
11. What happens to your data after the research project has concluded?
12. How important do you think these services (data management planning; institutional repository; data storage; data carpentry; metadata support; research impact (Altmetric and ORCID) might be to you and your students and how likely would you be to use these services?
13. How important is it for your students to learn about data information literacy processes and tools?
Design thinking activities

Introduction to design thinking
Initiated at the Stanford D-School (http://dschool.stanford.edu/) and championed by IDEO CEO Tim Brown, design thinking is a way to approach a problem from the user’s perspective. It is highly iterative and non-linear, and failure is encouraged. The process consists of several steps that are designed to solicit user input prior to engaging in the design process.

Empathy: The goal of this step is to develop a deep understanding of user needs and therefore comprehend the situation from their point of view. Most often, this takes the form of in-person interviews and asking open-ended questions to tease out what the user’s challenges are in a particular context, whether it involves something physical, like lack of space, or more abstract, such as data management support.

Ideation: This step involves the distillation of the insights and needs into a compelling problem statement that can serve as a solution-generation springboard during the later steps. This stage is all about identifying patterns or problem statements as well as both explicit and implicit needs and focusing on generating as many “solutions” to solve them. These solutions are often framed as “How might we...” questions in order to open up the possibilities prior to thinking about constraints. For example, we might post a question such as “How might we assist faculty in storing their data sets in our institutional repository?” A final goal of ideation is to explore a wide range of solutions—both a large quantity of ideas and diversity among those ideas. From this grouping of ideas you can then build prototypes to test with users.

Prototyping: This step begins the process to narrow down solutions to those that are deemed most feasible and are ready to be piloted as a model—whether it is a service, tool, or program. Prototyping involves not only the creation of the model itself, but the activities surrounding it which consist of knowing what type of feedback might be sought, how it will be recorded, as well as a plan for letting go of what is not working and further developing what is.

Testing: Refine solutions and make them better based on the additional feedback received until a desired “final” result has been achieved. This final step comprises acting on the feedback received, communicating results or next steps with stakeholders and a wider audience as well as celebrating successes and seeing where improvements can still be made, and most importantly, trying again!

How we applied the design thinking process at OSU
Our empathy step had already been completed with our 31 faculty interviews across almost all disciplines: business, construction management, design, housing and merchandising, engineering, entomology, geology, history, sociology, and zoology. Conducting face to face interviews allowed us to connect with each faculty member on a personal level, which is not necessarily represented in the interview notes themselves, but was readily apparent during our conversations.

The next two steps involved ideation and prototyping to determine how we would define challenges and brainstorm solutions. We split up the faculty responses among the liaisons, the associate dean, and the head of the liaison program. Each person was responsible for examining responses on the spreadsheet and summarizing the results that were then recorded in a Google doc for ease of access. After that second-level triage was done, we met as a group and wrote each perceived challenge from the summary response onto a different Post-it note. Once this round of analysis was over, we quickly categorized our ideas (shown in the picture below) and realized that there were several commonalities.
During a separate meeting, we followed the steps to generate and narrow down solutions as part of the ideation phase. Individually, we wrote ideas prior to a broader group discussion. Subsequently, we categorized each set into broadly grouped themes based on the following instructions:

1. Re-frame challenges into “How might we...” questions.
2. Draw and be visual.
3. Keep all ideas on the table at first.
4. Switch to a new brainstorm question every fifteen to twenty minutes.
5. Set a goal for how many ideas you want to generate in total.
6. Look at each idea:
   a. What is at the core of your idea: what gets you excited about it? What is the most important value for your audience? What is the real need that this is addressing?
   b. Make a list of all the challenges and barriers you are facing with your idea. What are you missing? Who would oppose the idea? What will be most difficult to overcome?
   c. Think of additional possibilities that might satisfy the needs your idea responds to. Brainstorm how you might address some of these challenges. For example: how might we raise money to acquire furniture for a certain space?
   d. Discuss how you can change your concept based on your new ideas. How can you address the need differently? How can you work around the constraints you are facing?
   e. Let go of ideas that feel too difficult to implement, or that you are not excited about.

While we may not have uncovered completely original solutions, this sequence was important for our brainstorming process on a number of levels. First, it allowed us to engage in both individual and shared thinking that often does not occur, especially as part of a group project where discussion quickly ensues without time for reflection. Second, we were able to define the elements that were of most importance and interest to our researchers regardless of what we already knew about other trends. Third, we were able to move from random ideas to cohesive action items fairly quickly thanks to our pre-identified categories which allowed us to keep the conversation moving forward and sustain momentum. Finally, we identified internal training and professional development needs that will allow liaison librarians to collaborate with faculty in an informed and proactive manner.

This is an example of how we captured these ideas:
Challenges
We grouped challenges into the following broad categories in order to allow us to easily identify patterns and areas of similarities in terms of needs or perceived knowledge gaps:

• **Awareness**
  - Knowledge gap about data management plan elements and existing campus (library) services that can support them (if available)
    • Not aware of library support services (x4)
    • Lack of awareness of what a DMP is (x4)
    • Mostly sciences use DMP

• **Data sharing**
  - Knowledge gap about the ethical repurposing and sharing of datasets
  - Knowledge gap about mechanisms and processes of sharing data (i.e., via repositories)
    • Ethical use of data sources (x3)

• **Data storage**
  - Lack of long-term archiving and storage solutions and processes
  - Lack of systematic storage of data methods and processes during projects; mostly stored on external hard drives, USB drives, and lab computers
    • External hard drives (x6)
    • Store data in multiple places, not necessarily coordinated
    • Lack of centralized storage
    • What to do with old data?
    • Data storage needs improvement

• **Data organization/manipulation**
  - Knowledge gap in understanding what metadata is and how it can be applied to their data
  - Lack of usage of data organization standards (i.e., file naming conventions, Readme files, etc.)
  - Confidentiality (x3)
Many have their own specific system which may or may not be documented and shared.

Confusing metadata with keywords (i.e., like for journal article abstracts)

Training and support
- Lack of coordinated campus efforts to provide centralized training beyond small-scale efforts (i.e., one-credit course, one-off workshops)
- Need training on how to write a DMP
- Need training for themselves (and students) in data manipulation (software carpentry), data information literacy, data organization
  - Many do not have tech support (x2)
  - Need basic training on metadata (x2)
  - Data information literacy
  - Software carpentry/coding
  - Lack of training for graduate students

Data types/size
- Knowledge gap regarding the types of storage, organization, manipulation options for different data types
  - Tabular data (x2)
  - Digital images (x3)
  - Large amounts of quantitative data

Solutions and Testing
As a result of the design thinking sessions, the planning group identified several action items to test out as prototype partnerships, services, and programming opportunities:

Priority Level 1
- Link to existing datasets and repositories from library website
- See what existing training videos and learning objects we can borrow
- Join Center for Open Science and create an OSU landing page
- Develop a library/institutional policy for ShareOK (OSU's institutional repository) to enable faculty to upload and store datasets to comply with federal grant requirements as a way to fill a much-needed gap for small to medium storage and discoverability
- Create a campus-wide research data committee that would engage campus partners in the following activities:
  - Consulting and access to a referral network both within the libraries and across the university of dedicated experts
  - Centralize research data services and resources
- Offer workshops and training
- Ongoing researcher engagement and needs assessment
- Define cyberinfrastructure planning and support
- Offer programming and events such as a data forum

Priority Level 2
- Develop a data certificate in conjunction with a cohort-based program—possibly offer stipends for faculty to participate
- Develop a training exchange program with other institutions
- Provide training for both graduate and undergraduate students and faculty mentors via Office of Undergraduate Research and in collaboration with the Graduate College

Priority Level 3
- Hire a data specialist position
- Integrate data information literacy as part of concurrent enrollment course and explore offerings throughout the curriculum for both graduate and undergraduate students
- Coordinate an OSU data forum to bring together researchers on campus
- Coordinate a data conference for regional/national programming
- Create an OSU data center to store all types and sizes of data
- Obtain an institutional membership to Globus

The library has already approached the OSU High Performance Computing Center as well as the Center for Strategic Proposal Development and taken the lead in offering pilot-type services in the form of a research data services website (http://www.library.okstate.edu/research-support/research-data-services/) and some workshops (http://info.library.okstate.edu/workshops), so we decided to target at least one item from each priority level as a way to make progress at each level of complexity and planning. To date we are in the process of:

1. Refining our liaison outreach efforts to include on-demand workshops in writing data management plans, measuring research impact, and specific programming languages in addition to more formal Software Carpentry offerings;
2. Identifying a small cohort of faculty to assist us in piloting the depositing of datasets into our institutional repository and developing researcher profiles that are connected with Altmetric;
3. developing a workshop series for graduate students; and
4. beginning campus-wide conversations to form the research data committee in collaboration with the Office of the Vice President for Research.

In terms of training for librarians, we have developed several avenues of action that will assist us in developing our own skill set and delivering training for faculty in a consistent and strategic manner:

1. We developed an outreach toolkit for data consultations consisting of general information on how to initiate contact, build rapport, and questions to ask during a face-to-face meeting. For a data management consultation, we referred to the article “Suddenly... I'm Consulting on Data Management Plans! Data Management Plan Consultant Checklist,” which has a helpful step-by-step process for walking faculty through the various elements of developing a data management plan. Finally, we asked that all of our liaisons document their meetings to be added to our annual impact spreadsheet (and to assist with evaluations) and note any follow-up items that may need to be continued.

2. We also created a lesson plan and handout set that can be tailored to specific disciplinary or departmental needs when liaisons receive requests for workshops and presentations: [http://info.library.okstate.edu/c.php?g=401548&p=3595256](http://info.library.okstate.edu/c.php?g=401548&p=3595256). In addition, we have created a dynamic DMP template that blends generic language with specific items inserted based on the researcher’s responses, such as repository choices. This allows participants to fill out the form as the workshop unfolds and leave with a solid first draft of their plan: [http://forms.library.okstate.edu/machform/view.php?id=80500](http://forms.library.okstate.edu/machform/view.php?id=80500).

Limitations and lessons learned
Overall, the process went fairly smoothly, with no major scheduling or data analysis problems, but things can always be improved. We obviously reached a small number of faculty out of the total who could have participated, and we realize that their responses might not be representative of the entire faculty population at OSU. They do however closely align with those presented in other similar studies, so we feel reasonably confident we received accurate data upon which to base our planning.

We decided against recording the interviews for two main reasons. We wanted to encourage as much participation as possible, and by extension make this process as simple as we could, and we felt that if we added the element of recording, faculty would not be as willing to talk with us. In addition, we did not have the staff time that would be needed to transcribe and process each interview. In keeping with the design thinking theme, we wanted to let the faculty guide the conversation and we allowed for a more flexible and organic approach using the interview questions as reference points rather than strict guidelines. Given the already rich level of information we gathered, we are not sure that having transcripts of what was discussed would have yielded more meaningful data than we received, as the goal of design thinking is to look at broad patterns rather than focus on each detail.

Each interview took longer than the expected 30 minutes, largely due to faculty going into more detail about certain aspects of their research than we had anticipated, which yielded a deeper level of understanding about their activities, but which also made back-to-back interviews difficult to schedule. About two interviews per day was optimal, and we found that even doing one additional interview caused fatigue on the part of the interviewers.

Although the initial recruitment message came from the associate dean in the form of an e-mail (see below), it was not until the liaisons sent the same message that faculty began responding. This points to the importance of the way in which the information is distributed and to the true power of the relationships librarians have with their departments.
An added benefit of these discussions was developing relationships with faculty who had largely never met their library liaison, even for those who were not entirely new to the institution. We also raised awareness of the new services we are thinking of offering (based on national trends occurring at other libraries) and received validation that these would indeed be useful.

Closing the loop on a project of this nature is vital. Rather than sending out a thank you note with a summary of the results, we opted to host an in-person presentation where we invited all of the participants along with the liaison librarians and representatives from the Office of the Vice President for Research. This allowed us to showcase this more personalized approach, continue to further build upon the initial contact we made, and show a direct response to the needs that were identified in our conversations.

Conclusion

Data management practices and policies will evolve, and libraries must do the same in order to capitalize on the partnerships and collaborative activities that naturally arise from these various activities. Libraries have quickly addressed the challenges of data management by providing services and resources based on researcher needs that are similar across institutions of varying types and sizes. We must therefore be able to respond and provide guidance for faculty who are working in this arena by thinking well outside the proverbial box and envisioning policies for data plans, project management, and publication models that are still in their beginning stages but which will shape the landscape in the years to come.

—Copyright 2017 Cinthya Ippoliti
Endnotes


2. Ibid., 393.


