<div align="center">**One-Page Abstract**</div>

Research as Design-Design as Research: Developing a Researcher-Driven Collaborative Model for Data Services

The Oklahoma State University (OSU) Library, under project co-leads Cinthya Ippoliti and Nicole Sump-Crethar, proposes a one-year $24,835 Sparks Grant to develop an innovative proof-of-concept qualitative research model and project management toolkit to assist libraries in developing workflows and services, in collaboration with researchers, for research data management throughout the entire research lifecycle. Specifically, our project consists of working with a total of 27 OSU faculty across STEM fields as well as the Social Sciences https://info.library.okstate.edu/imls to document their goals, processes, and habits during the entire duration of the research lifecycle https://cos.io/our-products/open-science-framework/ in order to answer our main research question, which is "What is the role and impact of the library in helping researchers manage their data along an entire project lifecycle?" As part of that work, we will:

1. Establish the following Advisory Board comprised of smaller, Hispanic-serving, Historically Black Colleges and Universities (HBCU), and tribal libraries: Natalie Dorfeld, Assistant Professor of English, Florida Institute of Technology; Carrie Forbes, Associate Dean for Student and Scholar Services, University of Denver; Karen Haught, Librarian, College of the Muscogee Nation; Amy Koshofer, Science Informationist, University of Cincinnati; Pamela Louderback, Library Director, Northeastern State University; Devin Savage, Assistant Dean for Assessment and Scholarly Communication, Illinois Institute of Technology; Sandy Tharp-Thee, Digital Inclusion Corps Member, National Digital Inclusion Alliance;
2. Implement two types of qualitative research methods (customer journey mapping and design thinking);
3. Develop a free and completely accessible toolkit housed within the Open Science Framework (OSF) https://osf.io/; and
4. Integrate a strong assessment component that will include the results of the customer journey mapping and design thinking activities, as well as a pre- and-post assessment survey to measure the change in researcher attitudes (i.e., awareness of library services) and behaviors (i.e., using library services) as a result of this project.

Many library surveys and interviews have been conducted to understand researcher behavior, motivation, and habits related to data management, but they have two significant flaws: 1. Surveys try to condense complex issues to a few broad categories that may not be applicable to researchers even within the same discipline. 2. Proposed solutions to the uncovered challenges are seldom designed in concert with researchers' input. This lack of deep understanding results in researchers viewing library services as an add-on rather than an integral part of the research lifecycle. Our proposed project intends to address these issues by working together with our Advisory Board, library colleagues, and researchers to identify the researchers' goals, processes, and habits throughout the research lifecycle and collaboratively develop shared solutions in the form of services, tools, and programs. In addition, we are hopeful that this project will have a positive impact on the libraries and institutions such as the ones we are hoping to reach initially via our Advisory Board. We are optimistic that this project would enable these libraries to begin conversations on their campuses via the methods described, and serve as a tool to empower libraries of any size to engage in conversations regarding research data services with direct input from their stakeholder groups. We also anticipate that researchers will take an active role in determining their data management needs and engage with each other and the library community to develop shared problem-solving techniques.

We propose the following timeline: July 2018: Train Graduate Assistant on customer journey mapping and design-thinking methods and prepare all materials. Convene and brief Advisory Board and establish a regular meeting schedule; August-December 2018: Administer pre-assessment. Researchers complete customer journey mapping exercise; January 2019-February 2019: Complete first round of data analysis to prepare for design thinking sessions; February 2019-March 2019: Host design thinking brainstorming sessions, begin developing toolkit, and administer post-assessment; April 2019-June 2019: Complete the second round of data analysis and develop the list of action items to address solutions identified during the design thinking phase. Launch toolkit in the OSF.

Research as Design-Design as Research: Developing a Researcher-Driven Collaborative Model for Data Services

STATEMENT OF NATIONAL NEED

The Oklahoma State University (OSU) Library proposes a one-year $24,835 Sparks Grant to develop an innovative proof-of-concept qualitative research model and project management toolkit to assist libraries in developing workflows and services for research data management in collaboration with researchers throughout the entire research lifecycle. Specifically, our project consists of working with a total of 27 OSU faculty across STEM fields as well as the Social Sciences https://info.library.okstate.edu/imls to document their goals, processes, and habits during the entire duration of the research lifecycle https://cos.io/our-products/open-science-framework/ in order to answer our main research question, which is "What is the role and impact of the library in helping researchers manage their data along an entire project lifecycle?" As part of that work, we will:

1. Establish an Advisory Board comprised of smaller, Hispanic-serving, Historically Black Colleges and Universities (HBCU), and tribal libraries. Our Advisory Board members are:
    a. Natalie Dorfeld, Assistant Professor of English, Florida Institute of Technology
    b. Carrie Forbes, Associate Dean for Student and Scholar Services, University of Denver
    c. Karen Haught, Librarian, College of the Muscogee Nation
    d. Amy Koshofer, Science Informationist, University of Cincinnati
    e. Pamela Louderback, Library Director, Northeastern State University
    f. Devin Savage, Assistant Dean for Assessment and Scholarly Communication, Illinois Institute of Technology
    g. Sandy Tharp-Thee, Digital Inclusion Corps Member, National Digital Inclusion Alliance;

2. Implement two types of qualitative research methods (customer journey mapping and design thinking) that will allow us to work with researchers to identify challenges and potential solutions for their data management needs;

3. Develop a free and completely accessible toolkit housed within the Open Science Framework (OSF) https://osf.io/ that will consist of all of the materials and directions necessary for libraries to implement this project within their institutions; and

4. Integrate a strong assessment component that will not only include the results of the customer journey mapping and design thinking activities, but those of a pre- and-post assessment survey that will capture to what degree researcher attitudes (i.e., awareness of library services) and behaviors (i.e., use of library services) changed as a result of this project.

An important note is that our emphasis for the purposes of this grant is on the process that other institutions can replicate, as opposed to the results themselves, which might or might not be useful outside of the OSU environment. We are not attempting to provide a one-size-fits-all approach, but rather to provide librarians and researchers with a model for structuring their data management practices in a way that is scalable and applicable to any academic environment because of its flexible and highly collaborative nature. This project aligns with the National Digital Platform Project Category in several ways.

- First, we are building on existing studies and taking a more in-depth approach than previously done;
- Second, we are relying on both our library and researcher colleagues from OSU as well as other institutions to provide valuable input into the methods and activities for this project;
- Third, we are intentionally using qualitative research methods to allow for the greatest flexibility, transparency, and collaboration possible during all project phases; and
- Finally, we are designing our processes and materials to have wide applicability and scalability, as the intent is for any library to use the training materials and toolkit to adapt the information to their specific needs.

Based on our research, we believe this is the first project to partner *with* researchers to undertake such an effort, and we believe this embodies the IMLS priority of placing the learner at the center, as researchers work alongside librarians to identify challenges and solutions to shared goals.

Research data management (RDM) gained global attention in 2013 when funding agencies like the National Science Foundation established requirements making research data openly available and the research behind it replicable. The changing landscape of the scholarly environment and new forms of authorship pose new challenges for libraries as students and faculty curate, repurpose, and share digital information, and librarians help them develop the skills to engage in inquiry, demonstrate respect for intellectual property, and communicate in ways that encourage original thinking. Larger questions remain about what librarians should communicate and how, especially when paired with the emergence of various platforms, measurements, and systems. In the years following the Office of Science and Technology Policy Memo, libraries have steadily expanded research data services, including data management, data publishing/archiving, and data curation. A gap exists in practice, however, which does not consider how researchers develop internal processes in managing their data. This lack of deep understanding results in researchers viewing library services as an add-on rather than an integral part of the research lifecycle. As noted by Ixchel Faniel and Lynn Connaway in their study of 36 academic library professionals regarding their perceptions of and challenges with providing research data services, less than a quarter "mentioned communicating, coordinating, and collaborating with researchers to facilitate RDM support, which led them to conclude that "more and different connections between librarians and researchers need to be explored to not only build awareness of researchers' needs and enhance RDM services, but also to build researchers' awareness of librarians' support and to enhance their perception of librarians' abilities to perform RDM services"(2018, p.114). Alison Mackenzie and Lindsey Martin, the editors of "Developing Digital Scholarship," mention that apart from the work of Antonella Esposito, "the literature does not reveal librarians gaining a research-informed, direct understanding of the behaviors and workflows of researchers" (2017, p.9). Even Esposito's study, whose aim was to determine the current uses of "digital tools and environments for research purposes" as well as what "researchers think of a likely impact of digital tools and environments towards an 'open scholarship' approach" (2013), primarily studied researcher behavior, leaving us to extrapolate and envision a role for the library.

Our proposed project also complements existing scholarship and practice in this area. At Oregon State University, it was found that faculty generate broad data types, with relatively small data sets, and manage their own servers (Whitmire 2015, p.391-392). The findings at Montana State University include varying definitions and attitudes towards open data and the need for collaborative efforts at the campus level (Sheehan 2015, p.11). The study conducted by Gu et.al. points to an interesting idea that has not appeared elsewhere in the literature--the need to conduct a scan of existing services and resources prior to conducting interviews (Gu et.al., 2012, p.20) in order to help inform the conversation that will ensue during the interview and help raise awareness of services 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 (DMPs), and the need for consultation and instruction" (2015, p.2). 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" (Scaramozzino et. al., 2011, p.362). 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" (2013, p.5), 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 to be aware of skills they need to effectively collaborate with researchers and to become aware of broader campus efforts in this arena. 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" (Antell et. al., 2013, p.570) amidst unfamiliarity regarding data management "assistance and initiatives on their own campuses or even within their own libraries" (Antell et. al., 2013, p.569). This renders any type of needs assessment even more important in order to focus training and project collaboration based on campus priorities and appropriate expertise.

Moreover, our results build on the work of Christine Borgman and Carole Palmer, who have studied researcher practices and behaviors regarding data sharing and dissemination and who focused on the collaborative nature of data curation workflows and re-imagining today's knowledge infrastructures, including underlying mechanisms, attitudes, and politics surrounding their collection, preservation, and dissemination. Mackenzie and Martin cite Borgman's book, entitled *Scholarship in the Digital Age*, which emphasizes that much of the library literature "has focused on the importance of

infrastructure," leaving us with "limited research to provide an evidence-based understanding of practice" (Mackenzie and Martin, 2017, p.13). It is on this missing element that we turn our attention to fill this void with a better understanding of practices that can inform infrastructure, policy, and workflows.

In addition to turning to the literature, and as a way to determine to what extent these issues were present at OSU, in Spring and Summer of 2016, we engaged in an internal design-thinking based data needs assessment where we interviewed 31 researchers across disciplines ranging from Engineering to History to identify:

- Faculty needs for research data support relating to DMPs, data analysis and storage, data information literacy, and research impact;
- The library's role in providing those services and resources;
- Who else on campus offers similar services, and how we can create partnerships; and
- The training librarians need to provide quality services to support these needs.

Our results mirrored the national trends discussed earlier:
1. Awareness: Knowledge gap about DMP elements and existing campus (library) services to support them;
2. Data sharing: Knowledge gap about ethical re-purposing and sharing of data sets and mechanisms and processes of sharing data, i.e., via repositories;
3. Data storage: Lack of long-term archiving and storage processes and lack of systematic storage of data during projects; mostly stored on external hard drives, USB drives, and lab computers;
4. Data organization/manipulation: Knowledge gap in understanding what metadata is and how to apply it to data and failure to use data organization standards, i.e., file naming conventions, Readme files, etc.;
5. Training and support: Lack of coordinated campus efforts to provide centralized training beyond small scale efforts, i.e., one credit course, one-off workshops. The need for training on how to write a DMP and data manipulation (software carpentry), data information literacy, and data organization; and
6. Data types/size: Knowledge gap regarding the types of storage, organization, and manipulation options for different data types.

This previous project allowed us to develop a robust initial set of Research Data Services encompassing a campus-wide research data services and cyberinfrastructure committee-Coalition for Advancing Digital Research and Education (CADRE)-to centrally coordinate activities; on-demand workshops, and consultations on software carpentry, research impact, and building a DMP; plus access and training for tools such as ORCID, Altmetric, and OSF. We also gained experience with the tools and processes involved in RDM practices, and presented our results at the 2016 Association of Research Libraries Assessment Conference.

PROJECT DESIGN
The main research question we intend to answer is what is the role and impact of the library in helping researchers manage their data along an entire project lifecycle? Additional questions which are a subset of our broader theme are:
- Do researchers have a clear set of goals regarding their data management practices?
- What are general perceptions regarding the role of the library as a facilitator of these goals?
- Are researchers aware of existing library services? If so, are they using them-why or why not?
- What are the specific barriers preventing them from using library services?

GOALS AND OUTCOMES:
1. Researchers will take an active role in determining their data management needs and challenges: Researchers will document their own data management practices and processes to develop a meta-cognitive scaffold with which to understand their own behaviors and habits. Researchers will make explicit their own goals in terms of managing their data.

2. Researchers will engage with each other and the library community to develop shared problem-solving techniques to assist them in managing data: Researchers will take part in generating and implementing solutions to address

the very challenges they identify as a way to crowd-source ideas that promote interdisciplinary engagement and discussion.

3. Librarians will demonstrate their capacity to work alongside researchers as partners, not simply supporters, of their data management processes: Researchers will exhibit a greater awareness of library resources and services as well as increased understanding of how they can assist them in achieving their research goals. Librarians/library resources will be considered a key component of researcher data practices.

The crux of this project is collaboration, feedback, and discussion during all steps. We had an overwhelmingly positive response to our call for interested faculty who said they saw value in what the library was trying to do and would find this collaboration helpful. Instead of offering a monetary incentive, which is not sustainable for the long term, we are offering in-depth support for the cohort's research projects, which clearly resonated, given that all of the faculty we asked enthusiastically agreed to participate. In addition, the Vice President for Research and the Dean of Libraries have agreed to award these and other faculty cohorts who participate in our library research data services program a co-sponsored "Working Open" certificate that faculty can display in their scholarly profiles, in addition to the comprehensive support mechanism we are offering as part of the project. This is in keeping with Weller's criteria for the transformation of scholarly practice-digitization of content, peer and content networks, as well as openness as an ideology and actual technical reality (2011, p.50). This is significant because it allows us to look at the digital products of research (i.e., researcher data sets), and develop interdisciplinary networks, by bringing together subject matter colleagues with different perspectives as well as work products, which are in and of themselves collaboratively developed and freely available. Our assumptions are that researchers have not taken full advantage of library services, and they have resorted to their own mechanisms for dealing with their data needs, which are less than ideal. We assume there will be areas of overlapping needs among disciplines that will be universal across institutions, but there will be some individualized needs that require additional attention on a case-by-case basis and are dependent upon localized policies and processes.

METHODOLOGY PART 1: Customer journey mapping is a form of participatory design and a staple of user experience studies. It differs dramatically from a reference interview because the researcher *personally documents steps taken at each phase of the project* via charting and journaling instead of simply talking about them. A journey map allows us to document the researcher's experience step by step, identify at what point (if at all) the researcher interacted with the library, and determine their attitudes and goals along the way. We are documenting researcher behavior at each step of the lifecycle of their project and, more importantly, what their goals were at each phase that might have benefitted from working with the library (if they did not do so). If they did, we would be able to determine what that interaction entailed and if any improvements could be made.

If researchers did not contact the library, this would allow them to tell us why not, and if it would make sense for them to have contacted us, and how can we change that for the future as we seek to match the appropriate library interventions to help them meet each of these goals. A sample journey map is included to clarify how a researcher documents the actions and thoughts present at each step of the research cycle. We would develop a separate chart for each major step so that we can cluster and cross-tabulate responses across researchers and disciplines for each portion of the cycle. We will first meet with all researchers involved in the project to provide an overview of the project goals, activities and deliverables, and timeline. They will then receive their customer journey maps which they will fill out for one specific project for the duration of the grant. In our conversations with our researcher partners, they are open to utilizing either a previous or a current project for this activity, and we will ensure that we have as close as possible to a 50-50 split between these two project types so that we can compare how responses differ between them, which will give us an opportunity to identify commonalities as well as more distinct elements.

We will collate the data collected as part of each journey map and use spreadsheets, and as appropriate, visualization software to develop a researcher dashboard that will allow us to identify common elements found in the journey maps for each step across researchers and disciplines, document individualized needs and goals, rank and prioritize opportunities, and isolate areas where a deeper dive into the issues is needed. For the areas where opportunities and deeper issues are identified, we will meet with researchers to walk through the design-thinking process of data analysis where we take

information from the customer journey mapping and use the design-thinking process defined below to generate solutions to the most pressing collective challenges we have identified.

METHODOLOGY PART 2: Design thinking is a creative process generating solutions based on user-identified challenges. It is a highly repetitive process relying on user input and feedback throughout the design process. Design thinking was initiated at the Stanford D-School: https://dschool.stanford.edu/resources/getting-started-with-design-thinking. It is a way to approach a problem from the user's perspective and integrate elements of equity and empathy in design. It has become a well-documented process and is increasingly being applied to the library world. It was a recent topic of discussion at the 2017 ALA Conference for which the project's principal investigator, Cinthya Ippoliti, was a panel reactor: *From MLS to MLD: It's time to Integrate Design Thinking and Philosophy into LIS Education.* Alison Carr-Chellman and Michael Savoy discuss their view of user-design research and issues, and they make a clear distinction between user-centered design and user design. User-centered design is one which results in "something being done to rather than with the learner." By contrast, user design "empowers the users to engage authentically in the decision-making process that is design" (2004, p.702). Gretchen Rossman and Sharon Rallis describe the most inclusive form of design, emancipatory design, in which participants "are not generating knowledge simply to inform or enlighten an academic or social science community, but are collaboratively producing knowledge to improve their work and their lives" (2003, p.15). Because of the potential for this type of methodology to transform the way in which both libraries and researchers approach data management practices, we are attempting to make this as authentic an emancipatory design experience as possible. Here are the steps in the design thinking process:

Empathy (This step is already addressed with the customer journey mapping): The goal of this step is to develop a deep understanding of user needs and therefore comprehend the situation from their point of view.

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 a diversity among those ideas.

Prototyping: This step begins the process to narrow down solutions to those that are deemed most feasible and are ready to be piloted either as 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!

Each step of the process consists of guided brainstorming sessions where we identify our overarching problems or challenges that need to be addressed based on the customer journey mapping, and we proceed to generate as many solutions as possible by writing each new idea on a post-it note. At first, no solution is discounted as the focus is on generating as many different approaches as possible. During the next session, we will go through a similar process to narrow the list down to the solutions that are most feasible and useful to implement. Prototyping and testing will occur during the second phase of the project as they require actual implementation and refinement of these solutions. We will also draft a final report to be included on the project website that will address all of the elements of the project as well as lessons learned--all of which we plan to disseminate widely at conferences as well as through publications.

TOOLKIT: Our online toolkit within the OSF platform will consist of all of the materials and directions necessary for libraries to implement this project within their institutions. We plan to use the OSF as the platform for our materials to ensure they can be stored, updated, viewed, and disseminated to the widest audience possible for free. We decided on OSF as opposed to creating a regular website or a Libguide because OSF is a growing project management platform that researchers themselves are using and because of its ease of use and functionality in tracking versions, adding content, and creating new project copies as other institutions re-purpose this content for their needs. We will turn on commenting so individuals can comment on the project and add valuable anecdotal information. This will provide a more transparent method of communicating amongst ourselves and our colleagues. We will document our own process and provide explicit details on each step of the project so that they have everything they need to get started, including:

1. Recruitment messages and tips;
2. Background information on customer journey mapping and design thinking;
3. Templates and notes for how to conduct each phase of the process, tools needed, and tips for best practices;
4. Lessons learned and ideas for improvement;
5. Aggregate raw result sets (in compliance with Institutional Research Board guidelines)
   - Customer journey maps which will be filled in via Word document and stored on the project's Google Drive
   - The pre- and post-assessment questions which will be collected via Qualtrics
   - The post-it and other notes from the in-person meetings that will be photographed and summarized, then added to the project's Google Drive site;
6. Data from other libraries who undertake this project as our partners or on their own; and
7. Our final report.

As with all projects, we expect there will be some challenges involved. One risk we undertake is that we will not uncover any new information. We might also learn that the customer journey mapping and subsequent design-thinking methodologies were not the correct approaches to take in this instance. Another risk is that we will not be able to provide the researchers with the tools and services they need to successfully meet their goals once we uncover what they are. To account for these potential issues, we will conduct pre- and post- assessment activities (detailed in the assessment section) to determine researcher goals and awareness of and engagement with the library prior to the start of the project and afterwards. We will also rely on our internal partners such as the High Performance Computing Center and the Office of the Vice President for Research as part of our campus-wide CADRE group to assist us in addressing any unmet needs.

TIMELINE:
1. July 2018: Train Graduate Assistant on customer journey mapping and design-thinking methods and prepare all handouts, spreadsheets, and materials. Convene and brief Advisory Board and establish a regular meeting schedule.

2. August-December 2018: Administer pre-assessment. Researchers complete customer journey mapping exercise. Each of the sections of the research lifecycle will have interim mini-deadlines by which they have to be completed, and we will offer a combination of group and individual consultations so that maps are filled in as consistently as possible among all researchers.

3. January 2019-February 2019: Complete first round of data analysis to prepare for meetings to work through the design thinking process.

4. February 2019-March 2019: Host design thinking brainstorming sessions, begin developing toolkit, and administer post-assessment.

5. April 2019-June 2019: Complete the second round of data analysis and develop the list of action items to address and implement the solutions (as they relate to services, programs, tools) identified during the design thinking phase. Launch toolkit in the OSF.

Project co-leads Cinthya Ippoliti and Nicole Sump-Crethar will have project management responsibilities consisting of developing project goals and documentation, setting up meetings with our cohort, and overseeing data analysis and the development of the toolkit. Our graduate assistant will assist in data analysis and toolkit development. With full support of our Dean, we have allocated 5 hours per week each to fully devote to this project. We will need funds to hire a part-time graduate assistant to help us collect and analyze qualitative data and to provide a stipend for our Advisory Board who will participate virtually in all aspects of the project. Advisory Board members will be expected to provide guidance during every phase of this project not only from a subject matter expert perspective as relating to data practices, but also from the perspective of the researchers and communities they work with as a way to help expand the applicability of this work to any type of institution. They will virtually participate in meetings to give us guidance on the project activities and materials as we implement each phase, and they will also virtually attend the design thinking sessions and participate in discussions related to our data analysis. The graduate assistant will synthesize results from the customer journey mapping and organize them in spreadsheets so that we can determine patterns and outliers and, if appropriate, develop a visual display of the data using a tool such as Tableau. This individual will also assist us in organizing our design thinking data generated from our brainstorming sessions. The main tasks in this analysis are identifying patterns for each category based on researcher responses for the customer journey mapping and the design thinking session. For example, our customer journey maps may uncover a pattern that researchers routinely do not create ReadMe files as part of their project workflow. This would become one of our major problem statement categories, and we would determine how many researchers brought up this issue. We would note variations surrounding this problem so that we develop a complete a picture of this issue. Finally, this individual will add content to our OSF project site, where we will house all materials and results.

ASSESSMENT: Further assessment information, including performance goals and benchmarks, are included in the National Impact section below.

DISSEMINATION: We plan to present the project and its findings in person at a wide variety of conferences ranging from the Coalition for Networked Information, to the Research Data Access and Preservation Summit, as well as the annual Association of College Research Libraries and American Library Association conferences. Finally, we plan to publish the results in journals such as *College and Research Libraries*, *Journal of New Librarianship*, and *Journal of eScience Librarianship.* Results will be available via the project toolkit and freely accessible and available on the OSF platform.

DIVERSITY PLAN
We have attempted to include diversity in each aspect of our project in terms of the individual ethnic and gender diversity of our faculty participants and Advisory Board members, and via the types of research they conduct, which range from applied to theoretical and includes both small and large quantitative and qualitative data sets. We deliberately sought out Hispanic-serving, small, tribal and HBCU organizations for our Advisory Board who may benefit from this collective effort and who may not have the resources to conduct this type of research on their own. We reached out to colleagues with whom we have had prior working relationships, and who have demonstrated a willingness to take this project as seriously as we do.

We are also hopeful that this project will have a positive impact on the libraries and institutions such as the ones we are hoping to reach initially via our Advisory Board, then subsequently via our scaled project. According to the study by Faniel and Connaway, staffing and training were identified by over 78% of the librarians as the biggest barriers to being able to scale and integrate these services into their current jobs (2018, p.106). By developing services that are relevant and take into account institutional factors, this will enable libraries to direct their existing resources in a much more meaningful and purposeful manner than if they attempted to provide a generalized suite of offerings that will ultimately go unused. Smaller libraries have not been as involved (or not at all) in these conversations, per our literature review where articles and books are mainly published by large research institutions. We hope that this project would enable these libraries to begin conversations on their campuses via the methods described, and serve as a tool to empower libraries of any size to engage in conversations regarding research data services and help them develop solutions with direct input from their stakeholder groups.

Our university's commitment to excellence in diversity and inclusion was recognized with the 2016 Institutional Excellence Award from the National Association of Diversity Officers in Higher Education. We also received the nationally prestigious

Higher Education Excellence in Diversity Award for the fourth consecutive year. Because diversity is a core value, we intend to encourage inclusiveness in the project not only by engaging a diverse range of faculty partners and institutional partners, but also by recruiting from a diverse pool of graduate assistants. OSU is an Affirmative Action/Equal Opportunity/E-verify employer committed to diversity, and all qualified applicants will receive consideration for employment and will not be discriminated against based on age, race, color, religion, sex, sexual orientation, genetic information, gender identity, national origin, disability, protected veteran status, or other protected category.

NATIONAL IMPACT

Other library surveys and interviews have been conducted to understand researcher behavior, motivation, and habits related to data management, but they have two significant flaws: 1. Surveys try to condense complex and often discipline-related issues to a few broad categories that may not be applicable to researchers even within the same discipline. In addition, the data collected requires a secondary layer of interpretation by librarians who bring their own issues and biases into that analysis. Interviews go a step further in providing more individualized details, but this format also requires the librarian to interpret notes which may not accurately reflect what the researcher intended to say about an issue. 2. Proposed solutions to the uncovered challenges are seldom designed in concert with researchers' input, and librarians often rely on what other libraries are doing and a broad set of best practices as opposed to working directly with their user bases. Often, services are developed, rolled out, and assessed typically after little engagement with said services. As a result, there is a low level of buy-in from researchers, and services can only be developed very generally because librarians are trying to provide support at the baseline level to address the most generic of needs.

We will measure progress via both quantitative and qualitative methods where we will establish mutually agreed upon hard deadlines by which each component of the project will be completed, with the understanding that if faculty cannot meet the timeline, they may be asked to step aside from the project. We will send reminders and provide opportunities for individualized assistance so faculty have a high degree of success in completing the customer journey maps. For the design thinking aspect, we will require in-person attendance during our brainstorming sessions. We will also collect the completed materials such as the customer journey maps and the design thinking notes (which will consist of post-its on large pads and detailed photos, plus written transcriptions of each one) which provide direct evidence of progress, as we cannot move forward with the next phase of the project until the previous one is completed. Finally, we will work on the toolkit in parallel with the project. Once our final set of data is collected, we will then fast-track the toolkit development to complete the project on time.

In addition to collecting artifacts from the customer journey mapping and the design-thinking sessions, we will administer a pre- and post-assessment in which we will ask researchers about their level of involvement with the library as part of their daily workflows and reasons why they did or did not involve the library during a particular step. We will ask questions about their habits relative to data management which will also map onto the research lifecycle model so that we can more easily correlate answers for each stage of the project. For example, under the category of Store Data, we might ask, "Have you considered depositing your finalized data sets into our Institutional Repository?" with a follow-up question to clarify why or why not. We will then compare these answers to the goals researchers identified as part of the corresponding customer journey map to give us the benchmark for how these two elements fit together and where there are significant gaps. We will ask researchers to complete the same assessment at the end of this project and measure the difference in responses against our established benchmark. The sample questions are included here: https://info.library.okstate.edu/imls

We will also report the total number of participants, how many responded to our pre- and post-assessment questions, and how many did not respond. We will also ask three broad questions about their overall interest and understanding of this project based on their experience:
1. My understanding has increased as a result of this program/training;
2. My interest in this subject has increased as a result of this program/training; and
3. I am confident I can apply what I learned in this program/training.

Benchmarking will consist of measuring the initial correlation between researcher attitudinal data and behavioral data as mapped onto their research project goals. While we do not expect to conduct statistical analysis such as C-SAT which relies

more on customer satisfaction-type measures, we intend to collect evidence that both attitudes (i.e., awareness of library services) and behaviors (i.e., using library services) have improved as a result of this project. Our first form of benchmarking will consist of the post-assessment taken by researchers at the end of the grant period to ascertain if there is an increase in awareness of library services and an increased understanding of the role the library can play within the scope of their research projects. To minimize variables, we will administer the same survey as in the pre-test and note responses that indicate an upward trend in these categories in terms of how many faculty responded positively to these questions after being involved in this project. Our second form of benchmarking will occur when we solicit additional feedback as we implement the solutions identified during the design thinking process. By virtue of this type of participatory design, where we are integrating (and therefore influencing) researcher attitudes and behaviors into our service design, the set of services we are currently offering will serve as our starting point from where we can measure to what degree the new tools and activities are making an impact in these the areas (attitudes and behaviors) by assessing the success of these updated solutions.

If the proof of concept is successful, the scaled project would consist of a multi-layered approach that trains librarians and utilizes the toolkit materials to guide their work in conducting a similar project within their institutions. We are also relying on researchers and libraries on a national level as part of an expert community of practice so that conversations and artifacts emerging from individual campus projects are shared across institutions. This would be accomplished via:

1. An on-demand, online professional development curriculum delivered synchronously (repeated a few times throughout the academic year) and asynchronously. We will offer an on-demand webinar series so that librarians can watch the recordings and walk through the materials on their own time and ask any questions via the Google group mentioned below;
2. The toolkit and implementation guidelines within the OSF platform where we will gather and store data from all subsequent library groups who conduct this research within their respective institutions; and
3. Google Groups where we will open participation to any librarian interested in the discussion. We will hold virtual calls, post issues and questions, and brainstorm ideas for additional programming and activities.

We will also formally invite library institutional partners interested in participating. Advisory Board members will have the opportunity to transition from a consulting capacity to implementing the project within their institutions. We will solicit participation from diverse organizations like Hispanic-serving, HBCUs and non-ARL institutions via listservs, the ACRL Digital Scholarship Section, Greater Western Library Alliance Research Data Management committee (of which we are a member), and where we have institutional representation and personal networks. Partners will be expected to add their data to the project so that we can continue to build the knowledge base for this project and develop further refinements or identify new directions. Others who undertake this type of assessment can use the same mechanism for sharing their data back into the project, but it is not an expectation outside of the formal collaboration agreement. The scaled project would also include block grants (similar to the Mellon NCSU grant: https://www.lib.ncsu.edu/news/ncsu-libraries-brings-in-major-mellon-grant-for-visualization) for these libraries so that we can minimize barriers for institutions who may not have the initial resources to undertake such a project.

Looking towards the longer term, we are hopeful and optimistic that once we have collected this larger data set, we will then be able to address these issues at a national level and bring them to the attention of organizations such as the Association of College and Research Libraries that are in a position to develop and influence policy and harness shared resources across institutions to develop a much richer, scalable, and sustainable set of best practices and workflows than are currently available.

Schedule of Completion

Activities	Jul '18	Aug '18	Sept '18	Oct '18	Nov '18	Dec '18	Jan '19	Feb '19	Mar '19	Apr '19	May '19	Jun '19
Train Graduate Assistant on customer journey mapping and design-thinking methods and prepare all handouts, spreadsheets, and materials. Convene and brief Advisory Board and establish a regular meeting schedule.	■											
Administer pre-assessment. Researchers complete customer journey mapping exercise. Each of the sections of the research lifecycle will have interim mini-deadlines by which they have to be completed, and we will offer a combination of group and individual consultations so that maps are filled in as consistently as possible among all researchers.		■	■	■	■	■						
Complete first round of data analysis to prepare for meetings to work through the design thinking process, and then analyze that data.							■	■				
Host design thinking brainstorming sessions, begin developing toolkit, and administer post-assessment.									■			
Complete the second round of data analysis and develop the list of action items to address and implement the solutions (as they relate to services, programs, tools) identified during the design thinking phase. Launch toolkit in the OSF.										■	■	■

DIGITAL PRODUCT FORM

Introduction

The Institute of Museum and Library Services (IMLS) is committed to expanding public access to federally funded digital products (i.e., digital content, resources, assets, software, and datasets). The products you create with IMLS funding require careful stewardship to protect and enhance their value, and they should be freely and readily available for use and re-use by libraries, archives, museums, and the public. However, applying these principles to the development and management of digital products can be challenging. Because technology is dynamic and because we do not want to inhibit innovation, we do not want to prescribe set standards and practices that could become quickly outdated. Instead, we ask that you answer questions that address specific aspects of creating and managing digital products. Like all components of your IMLS application, your answers will be used by IMLS staff and by expert peer reviewers to evaluate your application, and they will be important in determining whether your project will be funded.

Instructions

- Please check here if you have reviewed Parts I, II, III, and IV below and you have determined that your proposal does NOT involve the creation of digital products (i.e., digital content, resources, assets, software, or datasets). You must still submit this Digital Product Form with your proposal even if you check this box, because this Digital Product Form is a Required Document.

If you ARE creating digital products, you must provide answers to the questions in Part I. In addition, you must also complete at least one of the subsequent sections. If you intend to create or collect digital content, resources, or assets, complete Part II. If you intend to develop software, complete Part III. If you intend to create a dataset, complete Part IV.

Part I: Intellectual Property Rights and Permissions

A.1 What will be the intellectual property status of the digital products (content, resources, assets, software, or datasets) you intend to create? Who will hold the copyright(s)? How will you explain property rights and permissions to potential users (for example, by assigning a non-restrictive license such as BSD, GNU, MIT, or Creative Commons to the product)? Explain and justify your licensing selections.

Copyright will be held by the creators of the content. All digital content created for this project will have a Creative Commons - Attribution (https://creativecommons.org/licenses/by/4.0/) license. Users of digital content created for this project will be encouraged to reuse, remix and improve on the toolkit.

A.2 What ownership rights will your organization assert over the new digital products and what conditions will you impose on access and use? Explain and justify any terms of access and conditions of use and detail how you will notify potential users about relevant terms or conditions.

At Oklahoma State University (OSU), the creator of the content holds the copyright if the work is not a direct assignment for their job. As copyright holders, the principal investigators will place no restrictions on use of created content beyond that outlined in the Creative Commons Attribution 4.0 license.

A.3 If you will create any products that may involve privacy concerns, require obtaining permissions or rights, or raise any cultural sensitivities, describe the issues and how you plan to address them.

No content created for this project will expose private information of project participants. Data collection will be approved by the OSU's Institutional Review Board and will comply with human subjects research protection requirements.

Part II: Projects Creating or Collecting Digital Content, Resources, or Assets

A. Creating or Collecting New Digital Content, Resources, or Assets

A.1 Describe the digital content, resources, or assets you will create or collect, the quantities of each type, and format you will use.

Digital content we create will include spreadsheets, data visualizations, assessments/surveys, photographs/scans of analog materials created by participants, and text documents. The requirements and input of the 27 pilot participants will determine the quantity of material collected and created. The data collected for this project is not expected to exceed 20 GB. We will prioritize the use of open formats wherever possible and plan to convert any data generated in proprietary formats into an open format for public use.

A.2 List the equipment, software, and supplies that you will use to create the content, resources, or assets, or the name of the service provider that will perform the work.

Software and platforms used in the project will include the Open Science Framework (OSF), Google Drive, Tableau, Excel, Adobe Acrobat Pro, Adobe Photoshop, Qualtrics, Microsoft Word, and Microsoft Excel.

A.3 List all the digital file formats (e.g., XML, TIFF, MPEG) you plan to use, along with the relevant information about the appropriate quality standards (e.g., resolution, sampling rate, or pixel dimensions).

DATA TYPE	FILE FORMAT	QUALITY STANDARD
TABULAR DATA	XLS, CSV	
TEXTUAL DATA	DOC, RTF	
IMAGE DATA	TIFF, JPEG, PDF/A	Uncompressed (TIFF 6.0), JPEG originals only, scans 300 PPI lossless in either black and white or grayscale unless color is necessary for interpretation of the image
DOCUMENTS	PDF/A	

B. Workflow and Asset Maintenance/Preservation

B.1 Describe your quality control plan (i.e., how you will monitor and evaluate your workflow and products).

All parties involved in data collection will undergo training guided by the principal investigator. The principal investigators will oversee the work completed by the graduate assistant, with careful attention to the project timeline and work quality. Data analyses in the form of visualization will assist with error detection. In addition, the advisory board will have access to the work in OSF and Google Drive as the project progresses and will be encouraged to provide feedback on all resulting products and data associated with the project.

B.2 Describe your plan for preserving and maintaining digital assets during and after the award period of performance. Your plan may address storage systems, shared repositories, technical documentation, migration planning, and commitment of organizational funding for these purposes. Please note: You may charge the federal award before closeout for the costs of publication or sharing of research results if the costs are not incurred during the period of performance of the federal award (see 2 C.F.R. § 200.461).

All collected data and products, even if they are stored in the cloud for collaborators' access, will be backed up to local storage in the form of network drives housed on servers in the library as well as on back up servers housed by campus IT in another location. In addition to OSF and Google Drive, the toolkit and associated data will be accessible in the SHAREOK.org repository, where it will receive both a Handle and a Digital Object Identifier (DOI), and will be preserved in the OSU Library's digital management and preservation system, Rosetta. We will use Rosetta to preserve any technical and descriptive metadata, check the integrity of the files, and help in the migration of data to new formats as necessary.

C. Metadata

C.1 Describe how you will produce any and all technical, descriptive, administrative, or preservation metadata. Specify which standards you will use for the metadata structure (e.g., MARC, Dublin Core, Encoded Archival Description, PBCore, PREMIS) and metadata content (e.g., thesauri).

The project will adhere to the OSU Library Metadata: A Guideline to Best Practices, https://info.library.okstate.edu/metadata/home. The OSU Library uses Dublin Core for descriptive metadata and a combination of locally defined lists of terms and the Faceted Application of Subject Terminology (FAST). Technical

metadata will be captured at the time of creation and saved with the files for preservation in the Rosetta system. A metadata dictionary will be prepared for the project to ensure consistency.

C.2 Explain your strategy for preserving and maintaining metadata created or collected during and after the award period of performance.

All metadata will be stored with the files on local library servers and saved with preservation masters in Rosetta. We will also make descriptive metadata openly available with the toolkit on the OSF and in SHAREOK at the completion of the project.

C.3 Explain what metadata sharing and/or other strategies you will use to facilitate widespread discovery and use of the digital content, resources, or assets created during your project (e.g., an API [Application Programming Interface], contributions to a digital platform, or other ways you might enable batch queries and retrieval of metadata).

The project will be openly available on the OSF. We will also make the toolkit and associated data available through SHAREOK, which has high rankings in Google search results, is indexed by SHARE (http://www.share-research.org/) and is listed in the Directory of Open Access Repositories.

D. Access and Use

D.1 Describe how you will make the digital content, resources, or assets available to the public. Include details such as the delivery strategy (e.g., openly available online, available to specified audiences) and underlying hardware/software platforms and infrastructure (e.g., specific digital repository software or leased services, accessibility via standard web browsers, requirements for special software tools in order to use the content).

The toolkit and associated data will be openly available through the OSF that strives to make research and data openly accessible. By employing the built-in tools and extending the framework with add-ons, especially Google Drive, we can make the toolkit widely accessible *and* easily reused and remixed. The OSF is accessed via standard web browsers and, because our project will be public, no login will be required to access the content. Every effort will be made to use open file formats. If an open file format is not available, the project will use commonly available file formats. All PDF content will be processed using optical character recognition to make full text searching possible. In addition to the OSF, all data and materials for the toolkit will be made available in SHAREOK, which is built on DSpace and is easily accessed via standard web browsers.

D.2 Provide the name(s) and URL(s) (Uniform Resource Locator) for any examples of previous digital content, resources, or assets your organization has created.

The OSU Library maintains a large number of digital collections. Most of these are on the library's website, https://library.okstate.edu/search-and-find/collections/digital-collections/. Other examples of digital collections are in SHAREOK, https://shareok.org/handle/11244/6231. In addition to our digital collections, we use the OSF to host our data management planning materials and workshop handouts, https://osf.io/83s6g/.

Part III. Projects Developing Software

A. General Information

A.1 Describe the software you intend to create, including a summary of the major functions it will perform and the intended primary audience(s) it will serve.

A.2 List other existing software that wholly or partially performs the same functions, and explain how the software you intend to create is different, and justify why those differences are significant and necessary.

B. Technical Information

B.1 List the programming languages, platforms, software, or other applications you will use to create your software and explain why you chose them.

B.2 Describe how the software you intend to create will extend or interoperate with relevant existing software.

B.3 Describe any underlying additional software or system dependencies necessary to run the software you intend to create.

B.4 Describe the processes you will use for development, documentation, and for maintaining and updating documentation for users of the software.

B.5 Provide the name(s) and URL(s) for examples of any previous software your organization has created.

C. Access and Use

C.1 We expect applicants seeking federal funds for software to develop and release these products under open-source licenses to maximize access and promote reuse. What ownership rights will your organization assert over the software you intend to create, and what conditions will you impose on its access and use? Identify and explain the license under which you will release source code for the software you develop (e.g., BSD, GNU, or MIT software licenses). Explain and justify any prohibitive terms or conditions of use or access and detail how you will notify potential users about relevant terms and conditions.

C.2 Describe how you will make the software and source code available to the public and/or its intended users.

C.3 Identify where you will deposit the source code for the software you intend to develop:

Name of publicly accessible source code repository:

URL:

Part IV: Projects Creating Datasets

A.1 Identify the type of data you plan to collect or generate, and the purpose or intended use to which you expect it to be put. Describe the method(s) you will use and the approximate dates or intervals at which you will collect or generate it.

DATA TYPE	PURPOSE	DATES OF COLLECTION
Pre-assessment (Qualtrics surveys)	Assess participant knowledge of available services from the library	August-December 2018
Customer Journey Map (Word documents)	Documentation of the researcher's experience and goals at each phase of the process will help determine improvements to services	August-December 2018
Design thinking brainstorming sessions (self-stick easel pad pages)	Identify overarching challenges identified in the journey maps and generate solutions	February-March 2019
Post-assessment (Qualtrics surveys)	Reassess participant knowledge of available services from the library	February-March 2019

A.2 Does the proposed data collection or research activity require approval by any internal review panel or institutional review board (IRB)? If so, has the proposed research activity been approved? If not, what is your plan for securing approval?

The proposal has been submitted to the OSU Institutional Review Board, and we plan to comply with all requirements. Since the project will focus on "normal educational practices", we do not anticipate any problems with institutional review.

A.3 Will you collect any personally identifiable information (PII), confidential information (e.g., trade secrets), or proprietary information? If so, detail the specific steps you will take to protect such information while you prepare the data files for public release (e.g., data anonymization, data suppression PII, or synthetic data).

Identifying information will not be recorded during this survey, unless voluntarily disclosed by subjects. Subjects will be informed of the anonymous nature of this survey prior to self-selecting. Online survey data will be collected and temporarily stored via Qualtrics online survey software with access limited to the investigators and necessary IT professionals. Paper responses will be recorded in an electronic file and stored on the OSU Library's project drive server, with access provided only to investigators and necessary IT professionals. Paper copies will be stored in a filing cabinet in the PI's office. Raw data, both electronic and paper, will be destroyed within two years upon completion of findings.

A.4 If you will collect additional documentation, such as consent agreements, along with the data, describe plans for preserving the documentation and ensuring that its relationship to the collected data is maintained.

Consent will be collected as part of the surveys and maintained in Qualtrics online survey software as well as in an electronic file stored on the OSU Library's project drive server, with access to both provided only to investigators and necessary IT professionals.

A.5 What methods will you use to collect or generate the data? Provide details about any technical requirements or dependencies that would be necessary for understanding, retrieving, displaying, or processing the dataset(s).

We will use, interviews, focus groups and participatory design. Technical requirements will be limited as data collected can be displayed in common or open formats via widely available software platforms like typical web browsers, Adobe Reader, and Microsoft Word.

A.6 What documentation (e.g., data documentation, codebooks) will you capture or create along with the dataset(s)? Where will the documentation be stored and in what format(s)? How will you permanently associate and manage the documentation with the dataset(s) it describes?

All datasets will be assigned descriptive metadata that will be stored alongside the datasets in CSV files using standard practices outlined in the OSU Library Metadata: A Guideline to Best Practices, https://info.library.okstate.edu/metadata/home. All storage locations will include README files that document practices and describe the dataset. Datasets and metadata will be preserved in the OSU Library's digital management and preservation system, Rosetta.

A.7 What is your plan for archiving, managing, and disseminating data after the completion of the award-funded project?

Data will be shared via the OSF and SHAREOK platforms as described above. Data will also be managed and archived, with regular checks of file integrity, in Rosetta and backed up on a local library project drive. Data will be accessible online via SHAREOK and stored locally for a minimum of ten years after completion of the project.

A.8 Identify where you will deposit the dataset(s):

Name of repository: SHAREOK.org

URL: https://shareok.org/

A.9 When and how frequently will you review this data management plan? How will the implementation be monitored?

The principle investigators will review the data management plan on a quarterly basis to determine if revisions to the plan are necessary and to ensure compliance. Investigators will track any revisions, including who made the revisions and the implementation of those revisions, in the data management plan's documentation (README) stored with the data.