JLSC

ISSN 2162-3309 | JLSC is published by the Pacific University Libraries | http://jlsc-pub.org

Volume 6, General Issue (2018)

How Are We Measuring Up? Evaluating Research Data Services in Academic Libraries

Heather L. Coates, Jake Carlson, Ryan Clement, Margaret Henderson, Lisa R. Johnston, & Yasmeen Shorish

Coates, H., et al. (2018). How Are We Measuring Up? Evaluating Research Data Services in Academic Libraries. *Journal of Librarianship and Scholarly Communication,* 6(General Issue), eP2226. https://doi.org/10.7710/2162-3309.2226

This article underwent semi-anonymous peer review in accordance with JLSC's review policy.

ISSN 2162-3309 10.7710/2162-3309.2226

PRACTICE

How Are We Measuring Up? Evaluating Research Data Services in Academic Libraries

Heather L. Coates
Digital Scholarship and Data Management Librarian, Indiana University-Purdue University Indianapolis

Jake Carlson
Director of Research Data Services, University of Michigan

Ryan Clement
Data Services Librarian, Middlebury College

Margaret Henderson
Health Sciences Librarian, San Diego State University

Lisa R Johnston
Research Data Management/Curation Lead and Co-Director of the University Digital Conservancy, University of Minnesota

Yasmeen Shorish
Data Services Coordinator, James Madison University

INTRODUCTION In the years since the emergence of federal funding agency data management and sharing requirements (http://datasharing.sparcopen.org/data), research data services (RDS) have expanded to dozens of academic libraries in the United States. As these services have matured, service providers have begun to assess them. Given a lack of practical guidance in the literature, we seek to begin the discussion with several case studies and an exploration of four approaches suitable to assessing these emerging services. **DESCRIPTION OF PROGRAM** This article examines five case studies that vary by staffing, drivers, and institutional context in order to begin a practice-oriented conversation about how to evaluate and assess research data services in academic libraries. The case studies highlight some commonly discussed challenges, including insufficient training and resources, competing demands for evaluation efforts, and the tension between evidence that can be easily gathered and that which addresses our most important questions. We explore reflective practice, formative evaluation, developmental evaluation, and evidence-based library and information practice for ideas to advance practice. **NEXT STEPS** Data specialists engaged in providing research data services need strategies and tools with which to make decisions about their services. These range from identifying stakeholder needs to refining existing services to determining when to extend and discontinue declining services. While the landscape of research data services is broad and diverse, there are common needs that we can address as a community. To that end, we have created a community-owned space to facilitate the exchange of knowledge and existing resources.

Received: 11/15/2017 Accepted: 06/08/2018

Correspondence: Heather L. Coates, 755 W. Michigan Street, Indianapolis, IN 46202, hcoates@iupui.edu

INTRODUCTION

As federal funding agency policies on data management and sharing have been implemented, the prevalence of research data services (RDS) in academic libraries has exploded. With this growth comes an expansion of what the term *research data services* encompasses. Widespread use of the term implies a common set of services (e.g., data management plan consultations, data discovery and citation, and data information literacy instruction), which is rather misleading. Libraries are still in the process of defining the boundaries of research data services (Cox, Kennan, Lyon, & Pinfield, 2017; Bryant, Lavoie, & Malpas, 2017). The context for RDS programs—their organizational placement, degree of centralization, and which library staff provide them—is highly heterogeneous (Fearon, Gunia, Lake, Pralle, & Sallans, 2013). These services may also be composed of traditional library functions paired with new services (Cox et al., 2017). Often, RDS are not feasible without strong partnerships between offices of research and IT services. In other cases, existing services outside the library and available expertise necessitate a customized approach to developing new services. As RDS have matured, data specialists have begun to consider how to assess ongoing needs and measure and demonstrate value and impact for these services.

The literature on library-based research data services is still relatively new, with much of it published in the last 10 years. Descriptive studies on service development and composition, workforce capacity and capability, and perceptions about the role of the library in research data management make up a significant proportion of this literature (Cox et al., 2017; Lyon, 2007; Lyon, 2012; Tenopir, Sandusky, Allard, & Birch, 2014). In contrast, there is little discussion, practical or otherwise, of evaluating RDS as a coherent set of services. The lack of guidance for developing, maintaining, expanding, and discontinuing RDS contributes to the uncertainty expressed by many practitioners about how to proceed. Though RDS share common elements, there is great heterogeneity in how they are situated within the library and institution, the resources allocated (e.g., staff time and expertise), and the relationships with researchers, administration, and other support units. While this requires that service providers customize evaluation to their own circumstances, we believe there are common goals and needs that merit a community-based approach to developing evaluation strategies.

Building on a panel presentation at the Research Data Access and Preservation Summit in 2016, this article outlines current practice, identifies needs, and describes areas for improvement, as experienced by five librarians engaged in providing RDS. Using a case study approach, we offer a pragmatic discussion of strategies and challenges for evaluation and assessment. Drawing on the evidence base for existing library assessment and program evaluation approaches, we identify ways to move the practice forward as a community. Each of the five case studies describe evaluating RDS to meet one or more of the following needs:

- To determine whether the program or service has met its stated objectives;
- To describe ways to improve a program or service;
- To demonstrate the value of the program or services by gathering compelling evidence of their contribution to library and institutional goals; and
- To support other service-related decisions.

In the following five case studies, each author will describe the assessment goals, strategies, and outcomes for research data services at their institution. In each case, we see how assessment efforts adapt as services grow and change over time. At Virginia Commonwealth University (VCU), the data librarian developed an evidence-based practice approach to identify needs and improve services. At the University of Michigan, the RDS team analyzed data management plans, conducted interviews, and convened data summits to inform their services for subject liaisons and researchers. At James Madison University and Middlebury College, data specialists used Springshare tools to more strategically gather information about service usage and needs. The team at Minnesota developed an internal dashboard to more easily address common requests from administrators.

LITERATURE REVIEW

For this discussion, assessment and evaluation refers to the process of generating and using data to make decisions about services—from needs assessment to usage analysis to outcome and impact assessment. But first, we must start with the why. Service evaluation and assessment in libraries are conducted for many reasons. Nitecki (2004) describes several purposes for evaluating usage and usability, such as adjusting services in response to changes in available resources, responding to the needs of library users, improving service delivery, and conducting external reporting. However, these are not the only elements that shape an evaluation effort. The evaluator's role within the library and the relative emphasis placed on operations or clients, or a combination of both, shapes the purpose, methodology, and results of an evaluation effort (Nitecki, 2004).

In some cases, we can borrow from existing library assessment practice to inform the evaluation of RDS. Data information literacy instruction can be assessed like other types of teaching in libraries (Cahoy, 2014), whether the purpose is examining learning outcomes or promoting program or teacher efficacy. Assessing the outreach and engagement activities of data specialists can be modeled after the assessment of liaison librarians; indeed, many who provide RDS are also subject liaisons. Murphy and Gibson (2014) describe many ways to examine relationships, the quantity and quality of engagement, and the outcomes of liaison services. Similarly,

data reference questions can be evaluated as a component of existing reference services. Usage of a data management website or guide can produce real-time analytics on page views, referral sources, links followed, or search terms successfully used to discover content.

Services such as data management plan (DMP) consultations, ongoing consultations, and embedded support for data management activities are similar to traditional library services, but differ enough to pose challenges in adapting existing evaluation strategies and tools. The creation and curation of digital collections, preservation and curation services, infrastructure development, and maintenance to support such curation, discovery, and reuse efforts require new approaches, both for formative assessment and for determining long-term outcomes. Chapman, DeRidder, and Thompson (2015) note that generalizability for evaluating digital libraries across institutions is difficult due to wide differences in software and platforms. In digital library evaluation, there has been a heavy focus on analytics, costs, user studies, the user interface, software, and citations (Hufford, 2013; Saracevic, 2000; Zhang, 2010). Monitoring usage trends for complex services and infrastructure is more difficult. These efforts may include metrics such as the frequency and source of submissions, characteristics of submission such as file formats, metadata, and curation steps taken by the library upon receipt of data. Since groups may use a service for different reasons, capturing stories, anecdotes, and use cases can be a powerful way to connect evaluation efforts to user needs and experiences to demonstrate impact. While we are good at measuring quantifiable outputs—content created, classes taught, questions answered—some of the most important work we do is indirect. We support faculty in achieving their research goals. We enable students to be successful in their coursework and research. We contribute to reaccreditation efforts. How do we measure and evaluate our contributions to less tangible goals of creating knowledge, building networks of support, identifying unmet needs, and developing communities?

Evaluation offers a set of tools for accomplishing specific goals—to identify needs, to gather input regarding the effectiveness or perception of a service, or to demonstrate that a set of activities are affecting outcomes. In order to be useful, evaluation should be conducted with a well-defined question in mind. In practice, librarians are often pressured to collect such data with little encouragement or support for developing an effective strategy. While there has been some effort to support library and museum professionals in outcome-based evaluation (Institute of Museum and Library Services, 2000), many librarians are not trained in this skill set. Furthermore, the nascent form of research data services coupled with the rapidly shifting research and higher education environment suggest that process or formative approaches to assessment may be better suited to the task.

Here, we examine four approaches to RDS evaluation—reflective practice, formative assessment, developmental evaluation, and evidence-based library and information practice

(EBLIP). We do not explore summative evaluation, which typically addresses things like outcome evaluations, impact evaluations, and cost-effectiveness/benefit analysis. In the case of RDS, we are trying to understand and affect researcher behaviors that are shaped by complex and dynamic systems. These systems are themselves created by multiple and various stakeholders in research, including academic institutions, research communities and communities of practice, professional organizations, funding agencies, national policies and legislation, and publishers, among others. Given the tendency for summative evaluation to rely on models of causal mechanisms for program outcomes (Patton, 2011, p. 23), it is not a useful approach for evaluating RDS at this time.

Evaluation often begins with the personal, when professionals identify their own areas for improvement and seek information to help them make decisions about *how* to improve. Reflective practice can be a remarkably useful way to begin thinking about *how* we do our work, an important precursor to more formal evaluation. Practitioners rely on practical knowledge to do their work, which is the integration of content knowledge, process knowledge, tacit knowledge, and our beliefs and values (Jarvis, 1999, 49). Tacit knowledge is the type of knowledge that is difficult to articulate linguistically though it can be demonstrated (i.e., not all knowledge is narrative). A reflective practice approach can help practitioners to more clearly articulate the knowledge and questions we have to fill our knowledge gaps. Jarvis (1999, p. 68) describes reflective practitioners as professionals who are "not just responsive to the changing conditions of their practice; they are proactively asking questions about it." Generally, reflective practice encompasses reflective planning, reflection in action, and retrospective reflection (Jarvis, 1999, pp. 61–72). Mason (2006, pp. 89–96) frames it as an intentional merging of three worlds: one's own experience; one's colleague's world of experience; and one's world of observations, accounts, and theories. Though not all novices will develop into experts, reflective practice can support the transition to competency and beyond (Jarvis, 1999, pp. 51–60). Each person's practice is unique to their situation and ever-changing performance, though practitioners can attain validation by developing clear descriptions of phenomena that are easily recognized by fellow practitioners. Reflective practice is one strategy for RDS providers to identify service needs, particularly when used in support of attaining a specific goal. For solo data specialists especially, reflective practice may be the most natural place to begin. However, practices oriented to the individual rather than the service or program are not sufficient to effectively describe whether a program has met its stated objectives, or how to improve a service or demonstrate impact upon stakeholders. While personal improvement approaches can support individual practitioners in improving their own practice, evaluation necessarily separates the service from the individual(s) providing it.

Developmental evaluation (DE) is a relatively new addition to the program evaluation toolbox. Rather than evaluating to improve a model (formative evaluation, discussed in the next

paragraph) or judging whether something worked or not (summative evaluation), the purpose of developmental evaluation is to help people figure out what to do when things are uncertain or unknown. It was created to support the development of social change initiatives in complex or uncertain environments (Better Evaluation, n.d.). Patton (2011) states that DE is particularly appropriate when the knowledge base is too small to suggest effective practices. Gamble (2008) notes that "not all problems are bounded, have optimal solutions, or occur within stable parameters." A developmental evaluation approach is particularly well-suited for this stage of RDS, which often involves investigating an identified problem or opportunity and creating a service to address it. DE offers strategies to gain a better understanding of the problem, identify available and needed resources, identify key stakeholders, and describe the broader context. This evaluation approach is particularly helpful for surfacing and testing assumptions about the problem and cultivating a culture that supports learning, both of which are crucial to the success of RDS. For emerging service models such as ongoing or embedded consultation, data publishing, or data curation, developmental evaluation may be the best option. As services mature, other approaches can be incorporated into an evaluation plan.

In contrast to DE, the purpose of formative evaluation is to improve an existing model, service, or program. It includes activities like needs assessment, evaluability assessment, and structured conceptualization. However, formative evaluation is distinct from process and implementation evaluation. When used for instruction, formative assessment is used to provide immediate feedback for growth rather than for grading purposes (i.e., summative assessment). Similarly, when developing or adapting a service for a new population or setting, formative evaluation is a strategy for learning from stakeholders, determining whether the program or service model meets their needs, and identifying ways to improve it. Like summative evaluation, formative evaluation assumes that the purpose of evaluating is to test the program or service model, rather than informing the development of a model (Patton, 2011, pp. 36–40). Services such as data reference or data information literacy instruction may be excellent candidates for formative assessment, as the service models are often based on existing models of reference and information literacy.

Terminology and unfamiliar concepts can pose significant barriers to applying program evaluation approaches to library services. RDS providers may feel more comfortable with an approach that is rooted in LIS, such as the principles of evidence-based library and information practice (EBLIP). In their book *Being Evidence Based in Library and Information Practice*, Koufogiannakis and Brettle (2016) describe a new framework for EBLIP that explicitly acknowledges the role of local evidence and professional knowledge in combination with research evidence, and offers a structured approach for decision making that has great potential for evaluating RDS. The proposed framework describes five phases of EBLIP:

1. Articulate - What do I already know?
2. Assemble - What are the best evidence sources to answer this question?
3. Assess - How does the evidence I have apply in my context?
4. Agree - What is the best decision based on the available evidence?
5. Adapt - What worked? What didn't? What can be improved?

Hernon, Dugan, and Matthews (2014) describe how EBLIP principles can be applied to library service evaluation, providing comprehensive lists of evidence types organized by service. Rather than relying on easily available evidence, evidence-based library and information practice encourages professionals to start in familiar territory—by crafting a well formulated question. Whether used for service evaluation or other decisions, EBLIP guides practitioners through a structured process—articulating a clear and answerable question, assembling evidence from a variety of sources, assessing the quality and relevance of the evidence, agreeing as a group or organization how to proceed and begin implementation, and adapting based on further evidence and reflection. Questions are the primary tool of evidence-based practice in librarianship; thus, most of the tools provide structure to guide inquiry. Brettle and Koufogiannakis (2016) offer a series of prompts for each phase of the revised framework, while Booth (2006) developed a tool (Setting-Perspective-Intervention-Comparison-Evaluation (SPICE)) to develop clear and answerable research questions. Others have proposed domains of librarianship to help focus questions and situate them within a functional area of library practice (Crumley & Koufogiannakis, 2002; Koufogiannakis, Slater, and Crumley, 2004). Perhaps the most widely used EBLIP tool is the critical appraisal. Critical appraisal checklists such as those developed by Glynn (2006) and Koufogiannakis, Booth, and Brettle (2006) offer practitioners explicit guidelines with which to evaluate research evidence presented in the literature. Evaluating emerging services as they respond to rapidly shifting dynamics in higher education and research requires data specialists to be fluent with a diverse and flexible array of tools. We do not always explicitly detail these evaluation approaches within the case studies, but instead describe them as they were developed. We will, however, revisit these approaches in the discussion and next steps.

Case Study: Virginia Commonwealth University

The research data management program at Virginia Commonwealth University (VCU) began in August 2013 (Henderson & Knott, 2015; Henderson, Raboin, Shorish, & Van Tuyl, 2014). The program was developed as part of a larger effort to enhance scholarly communications services in the library. Due to budget cuts, we were unable to fill other positions, including one for an additional data librarian. The program became a solo librarian effort to launch a new service, offer consultations and teaching, and perform outreach and relationship building

tasks. This early uncertainty made it challenging to determine where to start and how to assess the services provided.

Deciding how to evaluate and assess services requires data specialists to make a lot of choices. At the individual level, it is possible to evaluate whether annual goals are met. Depending on how these goals are written, they may be too broad or too limiting for assessing the program. Individual performance goals are also personal; thus, they do not encompass the service as a whole. Such goals are also subject to change as professional skills grow and connections across the institution are established. While an individual librarian might reach a goal of consulting with five faculty members about data management plans each month, it is unclear what that accomplishment means for the RDM program. Individual goals contribute to but cannot substitute for program-specific goals.

Like other librarians, we track events and consultations such as training sessions, talks around campus, consultations, and reference questions. This data can help to identify trends, but the differences make it difficult to extrapolate or make comparisons between the various activities. More simply, counts can only tell us *how much*. How can we compare a quick student question to a workflow analysis for a laboratory experiment, or a lunchtime session on data organization for a few attendees, to a review of data compliance issues for grant compliance officers spanning the university? Can we capture meaningful differences by tracking the duration of the interaction, characteristics of attendees such as departmental affiliation or rank, and the format of the interaction? These questions reflect the daily decisions faced by data specialists in evaluating RDS. It is also difficult to compare a new department to an established department, or personal statistics to those of a liaison with a specific group of faculty and students to target. This approach fails to guide service-improvement efforts.

Drawing on expertise in evidence-based medicine, the solo data librarian began to think of assessment in terms of evidence-based practice. As described in the literature review, the first step is to ask an answerable question. Instead of attempting to assess all areas, we can ask "What is the best way to assess a new program?" This approach proved helpful for finding information that framed the options for assessment strategies appropriate to our reasons for assessment. As well as showing usage of services and resources for internal administration, statistics are also collected for reporting to accrediting agencies. But there are other reasons for assessment. Kloda (2015) notes that assessment has two (not mutually exclusive) goals: to inform decision-making and to demonstrate impact or value. Each of these goals requires quantitative and qualitative evidence. Lakos and Phipps (2004, p. 345) also point out the necessity of considering library performance from the customer's point of view: "In the rapidly changing information environment, libraries have to demonstrate that their services have relevance, value, and impact for stakeholders and customers. To deliver effective and high-

quality services, libraries have to assess their performance from the customer point of view."

Though there is little evidence available that describes effective assessment in RDS, we can borrow from other areas of librarianship and even other professions. Follow-up surveys are one way to learn about the customer's point of view. Marshall et al. (2013) surveyed physicians, residents, and nurses at 118 hospitals to find out how they used information sources in patient care. The survey examined many aspects of information searching and use including the type of work participants were doing, the resources they used, whether they found information, how they used the information they found, and the importance of that information to the final outcome. A survey sent to participants after an Open Science Framework (OSF) instruction session at VCU included similar questions to assess the usefulness of the training and the need for further sessions:

- Did you learn enough about OSF to start your own project?
- Is there any further help that would make you more likely to use OSF?
- How likely are you to use OSF in the future?
- When you signed up for the training session, what did you hope to learn?
- Did it meet your expectations?

These questions reflect some of the information needs experienced by data specialists and can be adapted for RDS trainings. They also speak to the adoption of strategies or tools into existing research practices, which is vital evidence for demonstrating the contribution of library services to the institutional mission.

Rather than viewing assessment as an additional step, it can also fulfill the research requirements for data librarians on the tenure track. It is important to conduct assessment in a systematic way so that it can meet all the requirements for informing or enhancing practice, as well as ethical and legal requirements, such as obtaining institutional review board (IRB) approval and abiding by Family Educational Rights and Privacy Act (FERPA) guidelines. In all cases, developing well-crafted research questions and carefully collecting evidence for reporting is crucial. While assessment as research may have a different scope than assessment for operational use only, it can improve the practice of individual librarians while contributing to the evidence base. We conducted a baseline survey that examined the current status and future needs of faculty with research data using questions adapted from existing surveys. Using a common set of questions enabled us to compare the data formats used by faculty at VCU to those used by faculty at Northwestern (Buys & Shaw, 2015, Figures 1 and 2). Similarly, we reused the reasons for not sharing data from a survey at the NIH Library by Federer, Lu, and Joubert (2015); the questions pertaining to storage amounts, data storage, and backup meth-

ods came from Akers and Doty (2013) at Emory University. By making surveys comparable to others, it becomes possible to perform a meta-analysis of institutional surveys to provide deeper insight into the data behaviors of researchers.

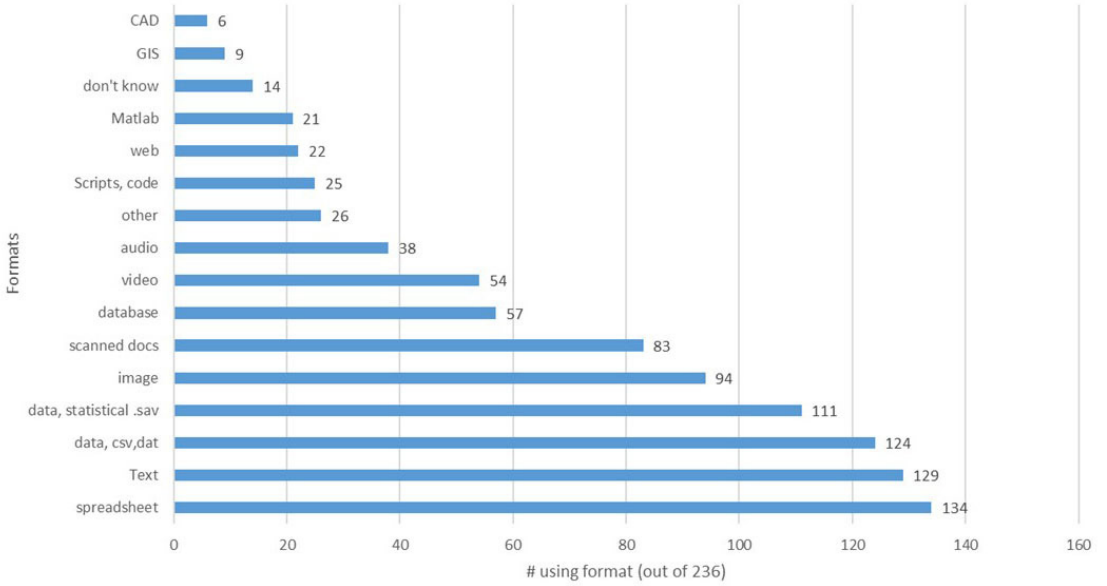

Figure 1. Faculty data formats from VCU survey

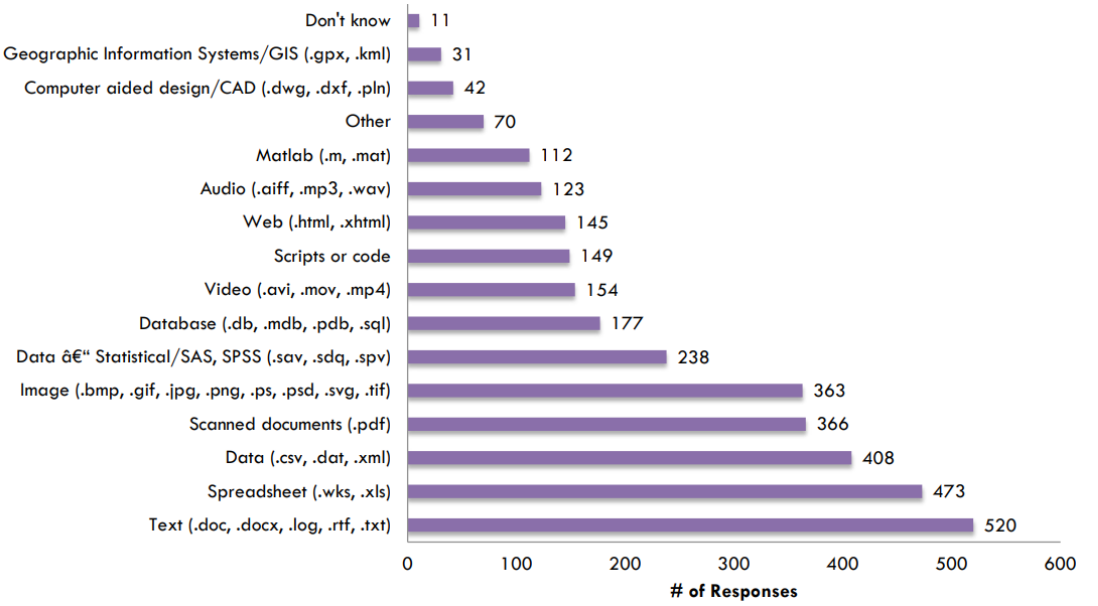

Figure 2. Type/format of data from Northwestern study (Buys & Shaw, 2015)

Extending beyond counts and surveys, the relationships we build are an important indicator that our services are having an impact. Academic institutions rely heavily on group decision-making, particularly as they relate to faculty governance. Goals that require consensus and collaboration with many groups take time to develop. For instance, it took two and a half years to develop the connections necessary to assemble a university-wide data summit that led to an institutional research data committee that will aid policy, funding, and faculty support (Figure 3). In situations where a more formal assessment is needed, the record of emails, phone calls, and meetings could be used to track these relationships over time, and a relationship map could be created to show the strength and connections between contacts. The map could be used to plan where better connections are needed, or to identify gaps in the relationships needed to support RDM services.

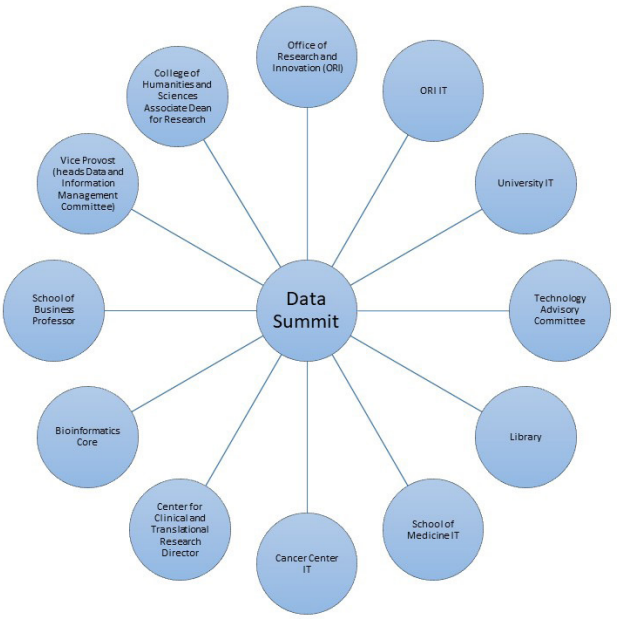

Figure 3. Members of the Institutional Research Data Committee

Case Study: University of Michigan

In 2014, the University of Michigan (U-M) began an ambitious initiative to develop and implement research data services to assist researchers in responding to new pressures and expectations to do more with their research data. U-M's approach to developing and implementing RDS has been to integrate these services into the work of librarians across the university. In considering how to structure our RDS, the library recognized that the data needs of researchers vary across discipline, data type, and access to resources and local practices, among other factors. Given this variation, library liaisons are the best suited to engage researchers about

their data and their needs. Liaisons are already responsible for building and maintaining relationships with faculty and students in the departments they serve. To support this initiative, the library created the RDS unit, a small team which serves as a core resource to librarians by developing service frameworks and providing the resources needed to carry them out. In situations where liaisons encounter service requests that they feel are beyond their ability to address effectively, they can bring in RDS personnel or other librarians with relevant expertise to work with them in addressing the data need.

As many libraries have done, the U-M Library has been actively conducting formative needs assessments to better understand the current environment for researchers and inform how the library could best support data management, sharing, and preservation through services. The specific types of needs assessments we conduct depend on the purpose, scope, target population(s), and intended use of the information we collect. For example, U-M's Engineering Librarians conducted an analysis of data management plans submitted by researchers from the College of Engineering to the National Science Foundation (NSF) as a means to inform the data management plan review service (Samuel, Grochowski, Lalwani, & Carlson, 2015). This analysis highlighted specific areas researchers appeared to be having some difficulty addressing such as metadata standards, intellectual property, and preservation storage. In response the library's data education team developed training sessions on these subjects to better prepare librarians to provide assistance to researchers. We also shared our results with the IT unit and grants office in the College of Engineering to strengthen our partnerships with them through illustrating types of questions on DMPs that would benefit from their expertise and input.

In 2015, librarians conducted interviews with 36 researchers over a two-month period to better understand the types of data that researchers are generating or collecting and how their data are currently being managed and curated, as well as to identify the challenges researchers are facing in working with their data. We then transcribed these interviews and identified the needs expressed by the interviewee and, where possible, what satisfying the need would enable him or her to accomplish. Each need was then translated into a statement structured as follows: Researcher "A" from department "B" has a need "C" to accomplish "D." Fitting needs into a common structure helped us facilitate comparisons between needs and to group like needs together. We then generated a report listing researcher needs and how they were grouped, allowing us to see the depth and breadth of particular needs. This report was used in part as the basis for the "minimum viable product" (MVP) document that served as the foundation of Deep Blue Data, the library's data repository. The MVP document served as a shared agreement between repository developers and stakeholders in defining the base functionality of Deep Blue Data that had to be in place before it could be considered ready for release. We also generated customized reports of our findings according to the discipline of the researcher

(Social Science, Arts & Humanities, Engineering, etc.) and distributed them to the library liaisons serving these disciplines so they would have a better idea of the specific data needs of the researchers they serve.

These reports have served as the foundation for discussions within each of the liaison teams as well as the library as a whole in developing our RDS program. The RDS team held two library-wide "data summits" over the summer of 2015 for librarians to discuss the needs assessment reports, share their experiences in working with researchers on their data management and curation needs thus far, and to consider what the library could offer as ongoing research data services. The outcome of the data summit discussions was the formation of six areas of focus for the RDS program loosely structured along the research data life cycle: data management planning, discovery, management and organization, documentation and description, data sharing, and data preservation. From there, the RDS team surveyed library liaisons to ascertain their comfort with and readiness to provide services in each of the six areas. Each liaison was asked if they felt that they were able to provide the service themselves ("do"), provide the service working alongside someone else ("collaborate"), or want to refer a service request to someone else ("refer"). The results of the "do, collaborate, refer" survey were then used to further inform training offered by the library's data education team and served as the basis for continuing discussions about what each liaison team would provide in terms of RDS, recognizing that different disciplinary researchers would have different needs.

The University of Michigan formally launched the RDS program and Deep Blue Data repository in September 2016. Needs assessment continues to be a means of connecting to researchers and informing the growth of our services. The challenge now lies in expanding assessment efforts to determine the impact of the services offered to researchers as well as the effectiveness of the service model for delivering these services. Given that the U-M model for offering research data services is to utilize the knowledge and skill sets available across the library, the RDS team will need to assess the work being done by librarians in engaging and delivering services, as well as the support mechanisms put in place to facilitate their work. For example, the RDS team will want to know how confident librarians are in engaging and carrying out our research data services. Are they getting the help they need from others in the library they have identified as collaborators, when they need it? Do they have access to the tools and resources they need to be successful? Answers to these and similar questions will help to continue to refine the service model and improve the support mechanisms needed for the library to offer research data services.

Many of the challenges the RDS team faces in assessing RDS are not all that different from assessing other library services. However, as a newly created service that spans across the library, it has been difficult to connect effectively with existing assessment instruments. Each unit in

the library has defined its own needs in tracking the services it provides and has developed ways to collect and analyze this information. Even for library units providing RDS-related services, such as reference, the tools used by library staff to document service transactions are not uniform throughout the library, nor do they align well with RDS assessment needs. The RDS team has considered requesting that additions be made to the assessment tools used by other units in order to capture more data about RDS-based service transactions. However, given the complexity of integrating RDS needs into existing approaches and the likely disruptions that it would cause, we are not actively pursuing integration at this time.

In the meantime, the RDS team has developed some customized tools for capturing the information necessary for delivering services and for documenting interactions with users. Many interactions with researchers involve multiple people. Establishing a baseline understanding of the situation and the work to be done in a way that can be shared and understood was of paramount importance. Our meeting notes template document was designed as a flexible means of capturing critical information through listing just a few broad-based fields for people to fill in. After entering information about the meeting itself (who was there, when it took place, etc.) the reporter is asked to fill in information about the data under discussion (including file formats, if available), the needs for the data, and possible next steps for addressing these needs. These meeting notes serve as an efficient communication tool to ensure that librarians have a shared understanding of the services to be delivered and as a baseline to measure our work against. The library has recently set up an account with Jira Service Desk for issue and project tracking, which has allowed the RDS team to move the meeting notes template into Jira to make tracking and documenting easier.

The RDS team also uses Jira to track and document the curation reviews performed on data deposited to Deep Blue Data. Once RDS knows a data set will be deposited, a ticket is created in Jira to capture information about the depositor and the data, including the category (experimental, observational, simulation, etc.), type (audio, image, tabular, etc.), and format of the data set. Jira is used as a centralized repository of information and communication about the data deposit. Any additional information or material about the data deposit that is subsequently created is attached to the ticket. Such information includes the results of the data curation review that is done as well as the email thread(s) between the depositor and librarians to share the results of the review and negotiate additions or edits to the deposit.

Though the RDS team has been collecting information and tracking the work done by the library, we have not yet conducted a full-scale assessment of our research data services. Improving assessment is one of the priorities of the 2018–2021 RDS Strategic Plan. Our assessment goals include continuing to develop our capacity to conduct needs assessments, building the necessary frameworks to learn from the experiences of library staff in offering RDS, and de-

veloping a better understanding of the nature and depth of the outcomes of providing RDS to U-M researchers.

Case Study: James Madison University

James Madison University (JMU) is a large comprehensive public university in Virginia, with approximately 21,000 students—90% of whom are undergraduate. While the National Science Foundation (NSF) mandated that a data management plan (DMP) be included in every grant proposal starting in January 2011, it wasn't until August 2011 that there was a librarian on staff with the content knowledge to address requests from departments on campus. As the new science librarian (and part-time data librarian), I took on the challenge of providing research services and DMP consultation to a population that had been without support for several months. In the beginning, these requests were focused on meeting the DMP requirement with minimal effort. Few faculty were inclined to discuss data management as a part of the research life cycle, or to discuss ways to leverage existing data in their project design. Initially, every interaction was in response to a user-initiated request. Due to staffing constraints and institutional will, the library was unable to provide infrastructure support to help strategically plan for growing data services, or to programmatically define a limited scope that would meet user needs. Moreover, as all data consultation occurred alongside existing liaison responsibilities, the ad hoc data services support begat ad hoc data services assessment.

Prior to the summer of 2014, there was no organizational analytic-capture software in place at the library. Email correspondence and instruction statistics were the only way to measure consultation requests, or "demand." There was no system in place to assess the outcome or impact of the data consultation. I often asked faculty to share the feedback they received on the DMP. Either they did not follow up, or the feedback itself was minimal and unhelpful. Moreover, untangling complex email threads—even when utilizing an organization structure within Outlook—was time consuming and provided inconsistent results.

With the implementation of LibAnswers in July 2014, the JMU Libraries now had a system to track data consultations. Using the internal note field, librarians could assign a data tag to track the question, answer, and other characteristics of the consultation (Figure 4). In August 2016, the transaction capture model was edited to allow for more granularity. At that point, "data services" became a topic designation, enabling more details about the transaction to be captured (Figure 5). This allowed me to pull up the statistics on various services (questions about data sources, requests for DMP consultations, data-focused reference questions, etc.), and to see what other liaisons were tracking as well. This was a great way to count things, but counting things is not the same as assessing things. Nonetheless, counting is an important first step toward assessing services.

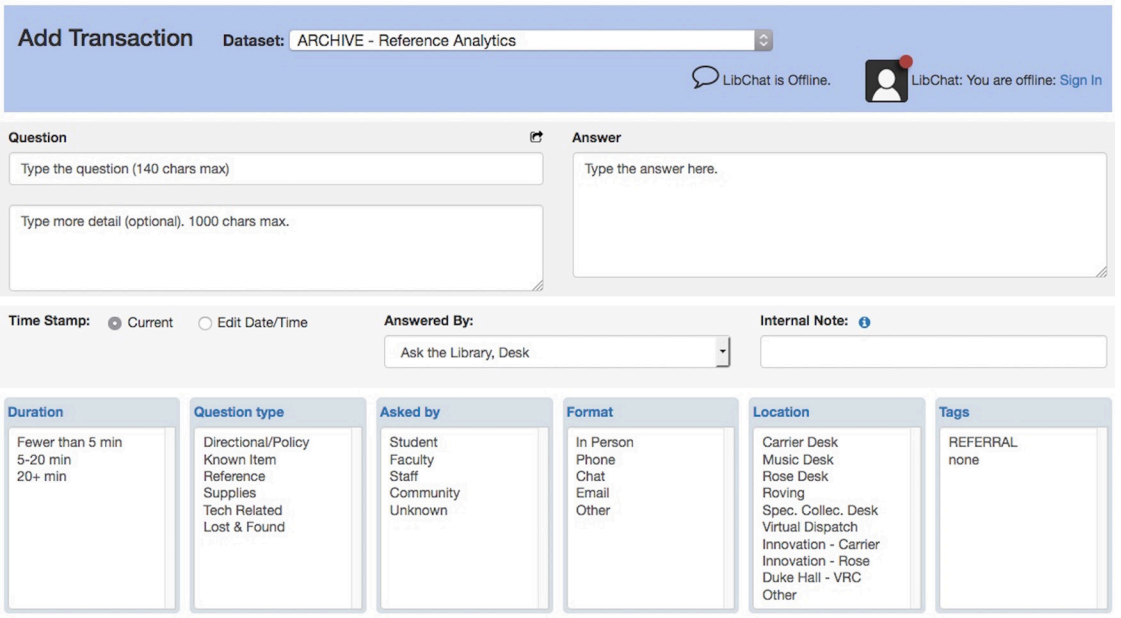

Figure 4. Screenshot of the initial LibAnswers configuration

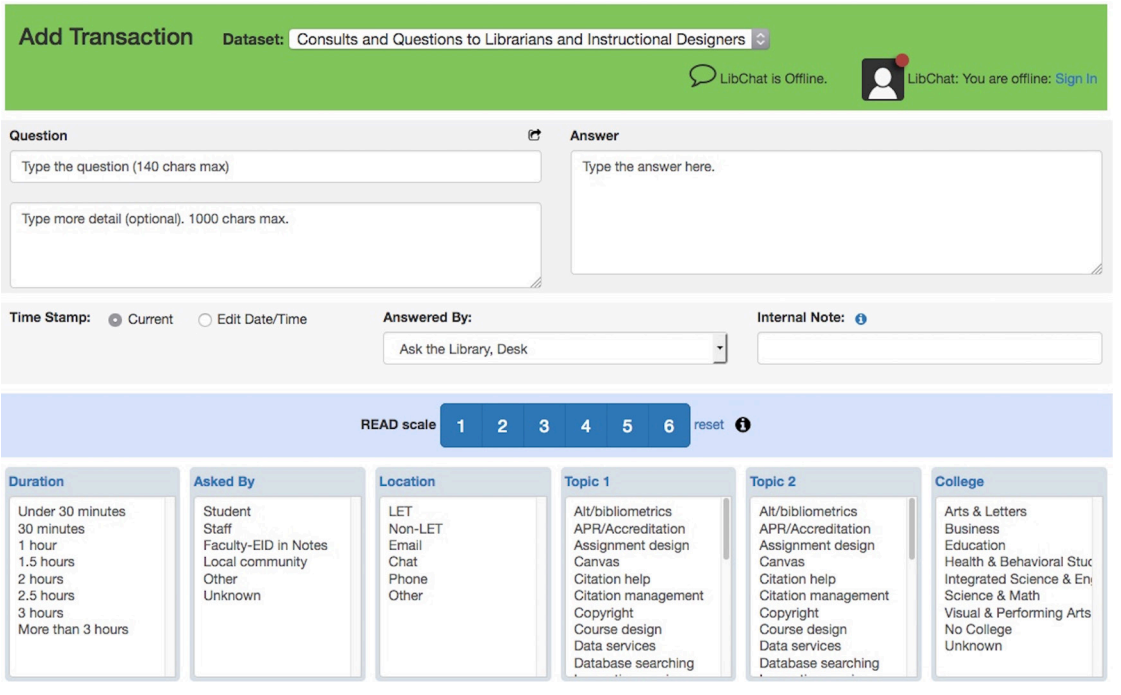

Figure 5. Screenshot of LibAnswers configuration with "data services" as a topic

As a liaison to multiple departments who also taught a semester-long research class, I was limited in my ability to be proactive with data services. In July 2016, the part-time data librarian transitioned to the role of full-time data services coordinator. In this role, it was possible to get in front of the speeding data bus and be a part of more conversations on campus. How much of an impact has this new dedicated data position had on users? The contact with humanities and social science faculty about data has increased (an increase from two over a two-year period to eight in an eight-month period). More inquiries and invitations to speak with lab groups, research classes, and institutes are being received, as tracked in LibAnswers (compare Figures 6 and 7). In four months, 23 data-related consultations were conducted, as compared to 24 consultations during the previous two years (2014–2016) with only part-time staffing of research data services. While these are imprecise metrics (not every transaction was logged), the evidence—both from LibAnswers and from conversations with faculty and students—demonstrates that full-time attention to this subject was a need on campus. This evidence was used to demonstrate to administrative stakeholders that faculty were responsive to and appreciative of expert support from the libraries (as indicated by repeat visits and referrals). Moreover, the resulting conversations from these interactions provided relevant information that was used to develop better research support services for the faculty.

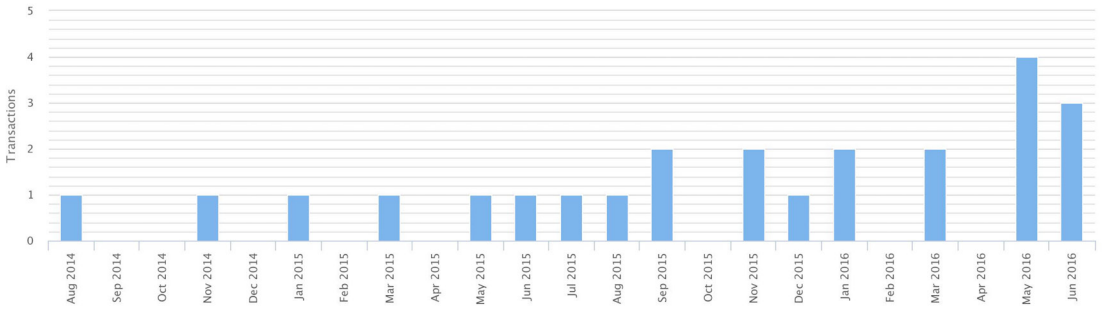

Figure 6. JMU Total data transactions = 24, over two years.

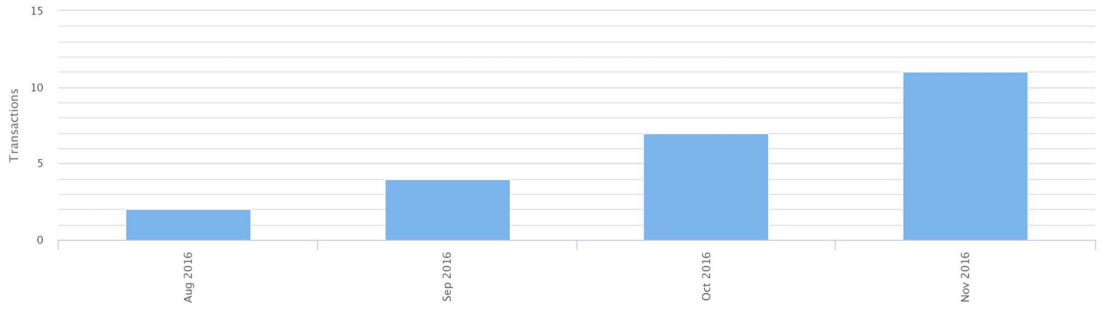

Figure 7. JMU Total data transactions = 23, over 4 months

Academic librarians sometimes find themselves in a reactive position rather than a proactive one, responding to campus needs instead of anticipating them. Without strategic and robust support, reacting to user needs can mean that the library operates from a place of continual catch-up. Through deliberate and consistent communication with the Office of Research and Scholarship, the JMU Libraries are attempting to change that dynamic. Supported by the quantitative data gathered in LibAnswers, which indicate a diverse need for data support as it relates to research and teaching, the libraries, campus IT, and the Office of Research and Scholarship have embarked on an in-depth, campus-wide information-gathering exercise. Through strategic meetings with faculty from across campus, the intention is to identify shared challenges and determine appropriate solutions at the enterprise level. Even for areas without enterprise-appropriate solutions, identifying areas of need can help departments allocate resourcing, or help faculty understand the limitations of the institution.

Qualitative assessment of data services at the individual interaction level is elusive, due to faculty disinclination to follow-up with DMP feedback and the lack of a postconsultation assessment tool. However, by leveraging increased awareness and use of library support in this area, new conversations have emerged that may help develop system-wide solutions to data management, use, and curation challenges. By the summer of 2017, I had met with 24 departments on campus to collect qualitative feedback. This was done using a semi-structured interview worksheet adapted from the Data Asset Framework documents (http://www.data-audit.eu/docs/DAF_Implementation_Guide.pdf).

While the scope of these meetings is broader than library-based research data services, they provide important data points in the data services assessment landscape. This effort is also an example of how important it is to build relationships between the library, research office, and campus IT. While the quantitative data on consultations with faculty helped motivate these partners to support this information-gathering exercise, it is mutual trust and respect for domain expertise that has been crucial to its success. As an example, while the final report is still in process, the most common feedback from nearly all departments was the need for a site license to NVivo software. I was able to provide this early-stage feedback to relevant administrative stakeholders at the midpoint of the interviews because it was so ubiquitous. As a result, the university has moved forward with securing the license before this project is completed. This nimble response to such low-hanging fruit provides both an easy "win" and a great example of a collaborative and responsive support environment.

Case Study: Middlebury College

Middlebury College is a top-tier liberal arts college in Middlebury, Vermont. About 2,450

undergraduate students attend Middlebury, along with approximately 200 graduate students in a variety of programs around the globe. Middlebury College founded the first environmental studies program in the United States, and maintains robust teaching and research programs in the natural, physical, and social sciences. The college currently has three librarians with data-related responsibilities: a data services librarian, a science data librarian, and a digital projects and archives librarian. The data-related responsibilities allocated to these librarians can be seen in Table 1.

Position	Data Discovery	Data Management & Curation	Software Support & Data Analysis	Data Visualization	Data Preservation
Data Services Librarian	X	secondary	X	X	
Science Data Librarian		X	X		X
Digital Projects & Archives Librarian					X
Student Tutors			X	X	

Table 1. Responsibilities of key data-related staff members

As at most small college libraries, librarians at Middlebury have a number of liaison departments, as well as responsibility for shared functions (e.g., user experience, scholarly communications, first-year outreach), in addition to their primary functional responsibilities as data specialists. The three data specialists share 13 subject areas. Students and faculty come to this environment with the expectation of close relationships, so a librarian at a liberal arts college must maintain extremely high-touch relationships with users in their areas of responsibility. Considering these demands, it makes a great deal of sense to track and understand these various demands as best as possible.

One of the responsibilities of the data services librarian is to assist the library in better understanding the available data on our services and collections. As part of this project, the library began to test the Springshare LibInsight product as part of our "Library Data Project." Prior to this testing, most of the library metrics collection and analysis was done at the work-group level (e.g., research and instruction, web services, collection management). It was more reactive than proactive, often with the goal of only collecting data that was necessary for external reporting purposes. Moving to the new system, the data services librarian worked with other librarians to reevaluate the forms for research desk interactions (Figure

8) as well as scheduled consultations and workshops (Figure 9). As data-related questions and topics can include a wide variety of terms (e.g., "stats/statistics," "census," "DMP," "Stata"), all librarians were asked to use the tag "data" in the "other notes" field of the collection forms. Since this tagging is often forgotten in the heat of a reference interaction, the process has required oversight and quality-control efforts. So far it has yielded information useful for improving data services outreach. For example, librarians have detected patterns in how data-related questions are asked—by whom, through which channel, and at what time during the semester. They have also identified patterns in which topics are asked about through particular channels. For instance, if many questions about finding and using census data are coming to librarians working at the research desk, the data services librarian can run a workshop for librarians on finding and using census data to increase their comfort in answering these questions. If many students from a particular class are scheduling one-on-one consultations due to trouble understanding codebooks, librarians can suggest a workshop on that topic to the faculty member teaching that class. This assessment has allowed the librarians to back up anecdotal experience with data.

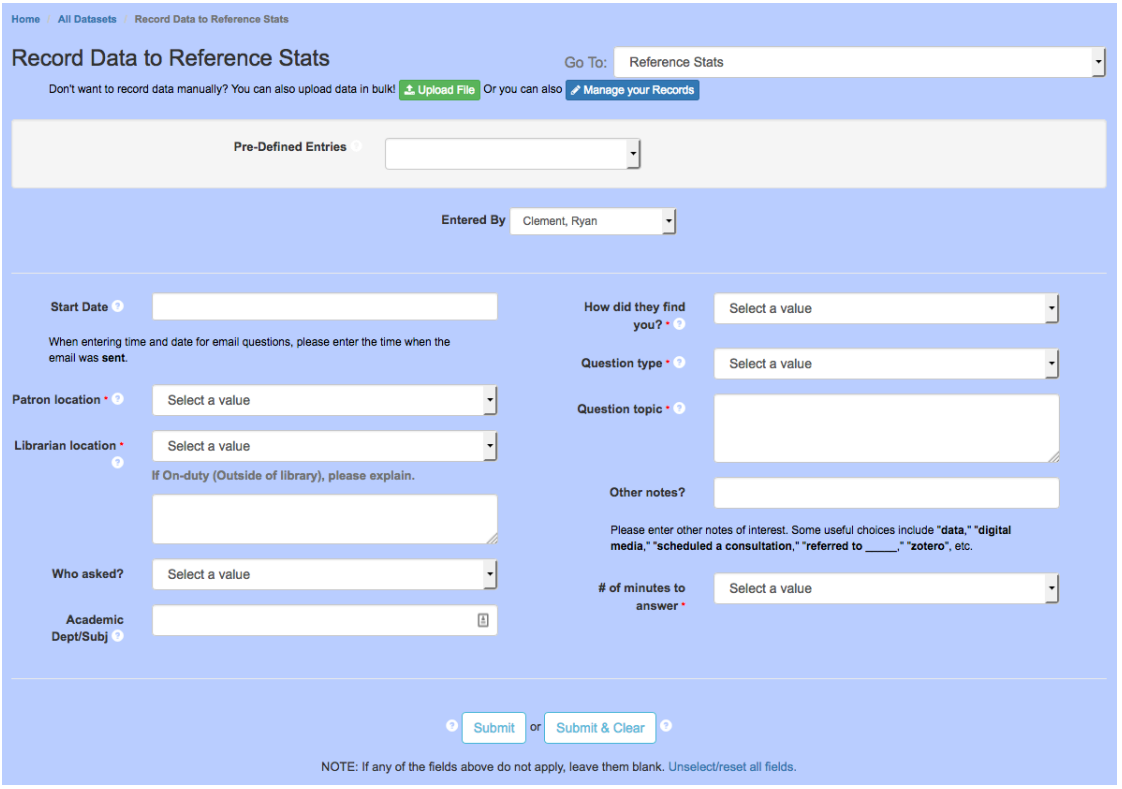

Figure 8. Screenshot of Middlebury's form for capturing research questions on and off the desk

Figure 9. Screenshot of Middlebury's form for capturing scheduled assistance such as workshops or consultations with students, faculty, or staff

An area for strong potential development of data services at small liberal arts colleges centers on how to reach the undergraduate population. At Middlebury, the data reveals that the library is reaching students with data questions. Both faculty referrals and workshop follow-ups bring a higher proportion of students to the research desk than other types of direction (see Figure 10). Often, though, these interactions are still centered on finding or citing data. Using the new data collection efforts in the library, we can begin looking for better opportunities to inject data management and other data literacy topics into these interactions with students. Being able to more easily analyze data on workshops, one-on-one consultations, and unscheduled reference interactions *together* can help to identify areas of need for students and librarians. For example, faculty teaching certain classes have expressed that doing workshops on finding and cleaning data for their students would be a waste of time because "every student is searching for different data." However, the data collected has allowed me to find patterns in the question topics and departments, identifying much commonality in the work the students are doing. This data lends support to the argument that library workshops on finding and cleaning data would be beneficial for these students. Following up on this data, I have been able to teach workshops on finding and interpreting data for certain economics classes, increasing the value to students as well as the scalability of the consultation services.

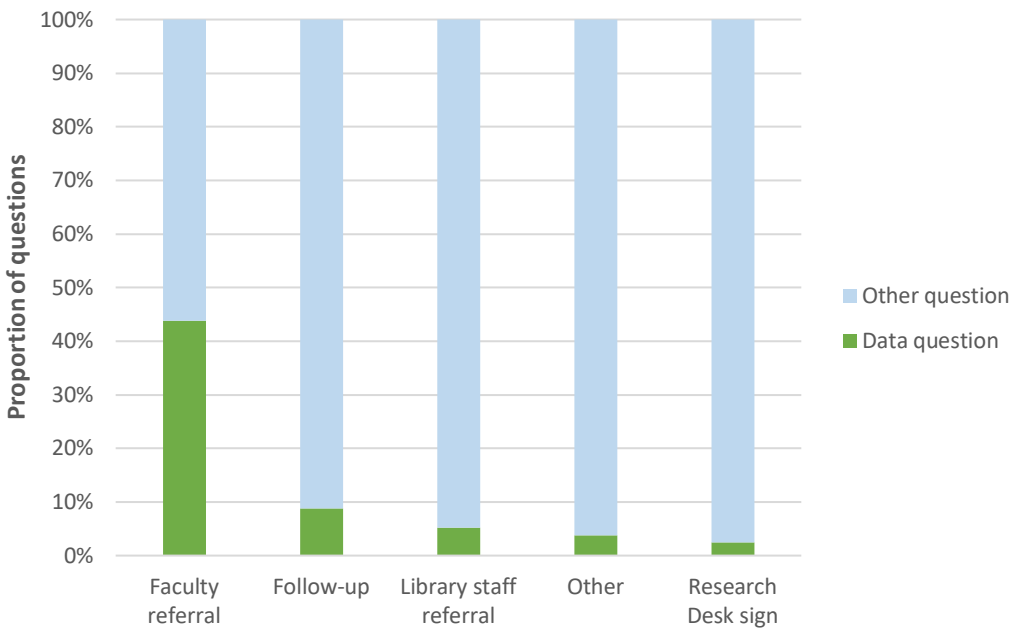

Figure 10. How patrons found their way to the research desk for questions related to data as a proportion of all questions

Another area for improvement identified in the newly collected data is the need to cross-train *all* librarians in data-related topics. As noted above, in a smaller, high-touch institution such as Middlebury, it is difficult for any one or two librarians to cover all inquiries on a particular topic. Without more long-term data available, it is difficult to describe the trend. Initial analysis shows that at least 4% of all research questions and 15% of all scheduled consultations and workshops cover data-related topics. The questions that come to the research desk are spread throughout all librarians who staff this desk, with every librarian answering at least *some* of these questions. The ability to subset these questions allows the data services librarian to find patterns in the themes and skills needed. This information is then used to identify topics for cross training, thus enabling all librarians to answer data-related questions at the point of inquiry.

Despite the small size of the student body, the demand for data services at Middlebury is high. With the development of an institutional open access and data repository, as well as increased interest in specialized data instruction, it is important to ensure that the data specialists are able to use their time wisely and receive the support they need. By more systematically collecting metrics on our data-related activities, as well as working to make this data more accessible to directors and administrators, we can make data services more proactive and sustainable in the long term.

Case Study: University of Minnesota

One of the main challenges faced by academic libraries with a growing data management program is how to quantify, measure, and keep track all of the hard work and success. But simply tracking metrics is not enough. Data specialists can store a mountain of metrics, but numbers alone do not convince stakeholders. To succeed, an RDS needs to have an effective approach to *presenting* the results of assessment to stakeholders in a meaningful, easy-to-digest way.

As a large land-grant research institution, the University of Minnesota serves roughly 50,000 students and 20,000 faculty and staff who brought in $754 million in sponsored research awards in 2015. To support the growing research data needs of this large, multidisciplinary university, the University of Minnesota Libraries have provided research data management services since 2008. These services include data management training and workshops, DMP consultations, regular needs assessments, and data repository and curation services through the Data Repository for the University of Minnesota (DRUM), which launched in March 2015.

Over the years, assessment of our services has occurred in a wide variety of ways and with

irregular timing. Most of the data was tracked in Google forms that were available to the individuals entering the data. To increase awareness and visibility of the statistics, the research data management/curation lead launched an internal library staff dashboard called the "DRUM Statistics & Visualization" page. It has grown to incorporate a range of data management and curation service assessments that are shared with stakeholders in real time. Using Google Sites as the platform, data from a number of different Google forms is synchronously pulled and visualized as live graphs and charts. Selection of data for the portal is based on past requests from the administration, such as the counts of consultation requests each year and the number of library staff involved in providing data-related services. This information can be hard to quantify and is not captured in other library reporting tools, such as our implementation of Desk Tracker (for reference desk statistics). Data is manually entered into the Google spreadsheets by the DRUM coordinating staff and the research data management/curation lead throughout the year.

The dashboard presents the following information distributed across four tabs:

- DRUM statistics. Each data set submitted to our data repository is reviewed. Information such as author name, department, status, and data-set type, file format, curation assignment (made to one of our six data curators), time to completion for curation tasks (Johnston et al., 2016), and the names of liaisons involved are tracked on ingest. This information allows us to chart the number of DRUM submissions over time, overall growth in data types (e.g., breakdown of file formats, disciplines represented, etc.), and the number of data sets originating from a particular department or group and the library liaison involved (Figure 11).
- DRUM satisfaction survey results. Satisfaction surveys are sent to DRUM users to gather feedback on particular features of DRUM that are valued and how users heard about the service (Figure 12).
- Overall RDS statistics include web page hits, the number of new consultation requests, DMPs reviewed, grant partnership requests, and the number of referrals to other units.

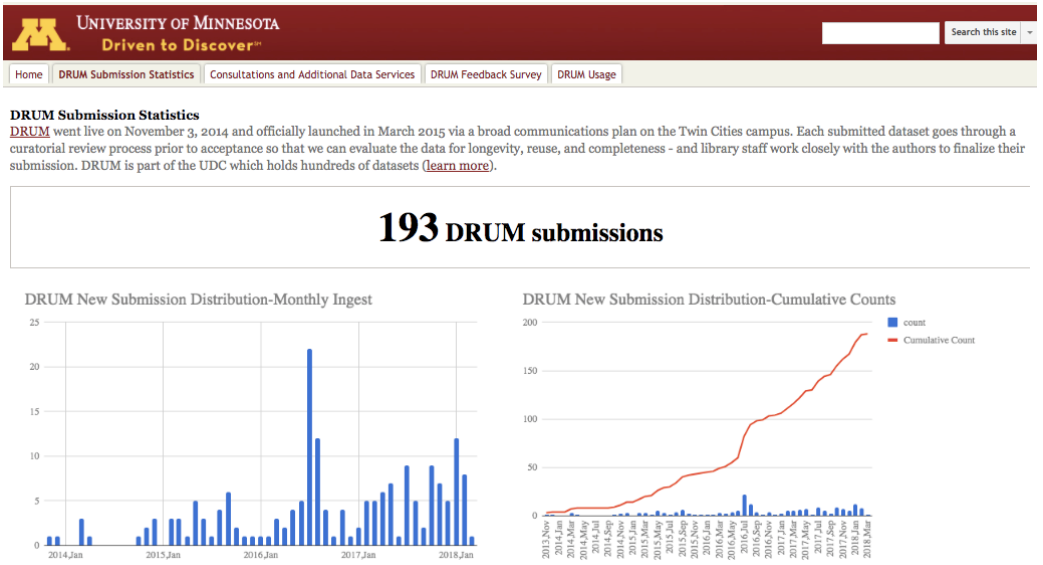

Figure 11. Metrics tracked by the University of Minnesota Library "DRUM Statistics and Visualization" dashboard – DRUM submission statistics

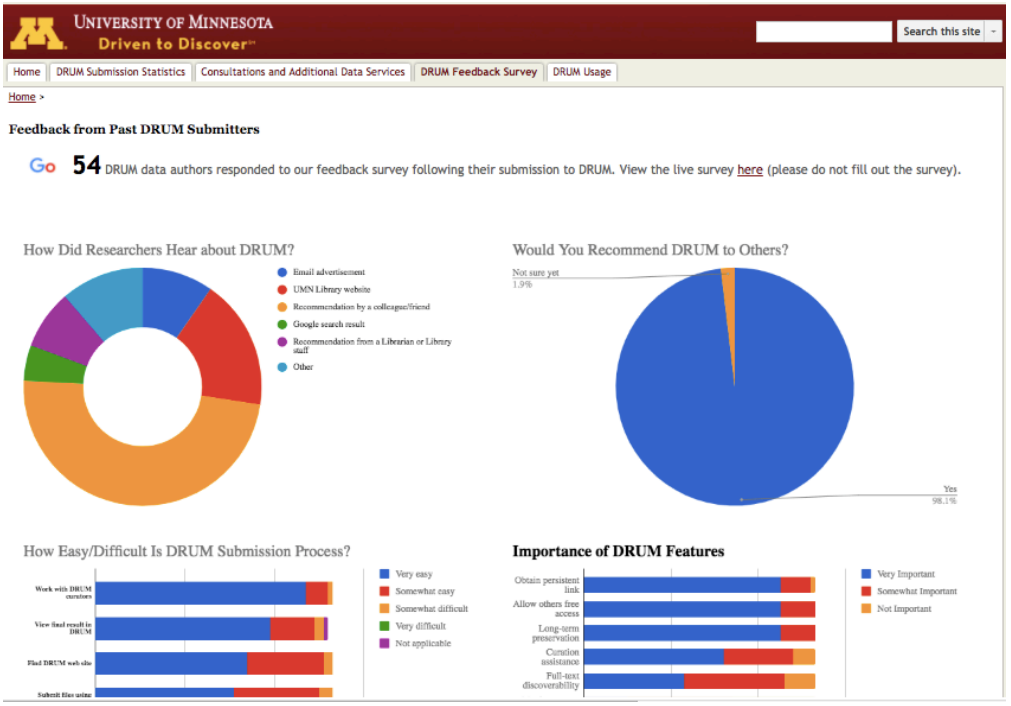

Figure 12. Metrics tracked by the University of Minnesota Library "DRUM Statistics and Visualization" dashboard – DRUM Feedback Survey

The dashboard approach has provided a number of lessons over the past year. On the positive side, this approach provided one clear view of what the data management services initiative is accomplishing. Responding to requests from library administration has been quick and easy, and the data can be used to compile end-of-the-year reports. Transparency is also key. It is important to share metrics on a regular (in this case continual) basis and to have people hold you accountable to growing the service appropriately. On the other hand, not all data is currently tracked in this way. Many liaisons have data consultations and requests that are not captured because they use Desk Tracker, Question Point, and other ad hoc methods. Also, the library uses many other systems for library assessment (including Tableau and SciVal), with which we have no integration at this time. Once a service is able to provide this information, the requests may increase! For better or for worse, the assessment tool itself becomes yet another "service" to continually improve and update.

DISCUSSION

These five case studies represent a range of institutional contexts, service composition, staffing, and assessment needs and approaches. One thing they all share is that each is still developing assessment approaches that suit their particular mix of services and make sense within the institutional context. At VCU, RDS providers used an evidence-based practice approach to identify needs and improve services. The RDS team at Michigan has conducted a detailed needs assessment to tailor services to local needs by analyzing data management plans, conducting interviews, and convening data summits with liaisons. They are now working to capture information on service delivery and beginning to examine the impact that these services have had on users. At JMU and Middlebury, RDS providers began using tools offered by Springshare to more strategically gather information about service usage and needs. The team at Minnesota developed an internal dashboard to answer common requests from administrators and provide real-time visualizations. As shown in Table 2, none of the case study institutions rely upon a single evaluation approach. Formative assessment was used by many institutions, while reflective practice proved useful for both solo and team-based needs. Unsurprisingly, all of the case studies described are using an evidence-based approach to making service decisions. Rather than offering one-size-fits-all assessment solutions, these case studies illustrate a variety of approaches for evaluating and assessing still-emerging services.

Reflective Practice *Appropriate for identifying user needs and concerns, trends in service usage, and areas for service improvement or individual training.*	In the context of preparing annual reports, the data librarian at VCU identified questions about what information should be tracked, whether it should be compared to data from other service areas, and how to use evidence to set goals and evaluate whether they were accomplished. Recognizing that they were answering similar questions about metrics again and again, a data librarian at Minnesota developed a real-time data visualization dashboard as a transparent reporting tool.
Formative Assessment *Appropriate for refining goals and adapting strategies for achieving goals.*	The data librarian at VCU began by gathering information about events, consultations, presentations, and workshops, such as date, location, duration, number of attendees, role, department, preparation time, and email address for follow up evaluation or outreach. The data librarian at JMU conducted interviews with faculty to determine data and computing needs. This information informed site license needs and will generate a recommendation document to guide enterprise-level solutions to shared challenges. The team at Michigan held a summit to gather needs from librarians to shape core services. The RDS team conducted interviews to describe needs and analyze patterns across disciplines. This information was transformed into a specification document that guided the development of the data repository. A group of engineering librarians conducted assessment of DMPs to inform the creation of a new service.
Developmental Evaluation *Appropriate for informing the development of services or programs in a poorly described and rapidly changing environment.*	*None of the case studies implemented this strategy. This is not surprising, since it is not commonly used in libraries. However, by adapting the Michigan team's approach, we have an example of developmental evaluation for RDS. By conducting regular interviews with researchers to describe needs and analyze patterns at regular, defined intervals, the team at Michigan could obtain a steady stream of data to describe changing needs to inform services.*
Evidence Based Library and Information Practice (EBLIP) *Appropriate for all evaluation needs as it is simply a structured approach for making evidence based decisions.*	To identify needs and contribute to the evidence base, the data librarian at VCU conducted a survey to describe common file formats and reasons for not sharing research data. At JMU and Middlebury, data librarians used the Springshare LibAnswers platform to systematically gather information related to data reference questions and consultations.This information was used to shape services, such as making the case for classroom instruction and increased service. The example from Minnesota enables continuous improvement and data-driven decision making by maintaining a real-time data visualization dashboard that tracks the type, frequency, and discipline of data sets deposited and curated by the data repository.

Table 2. Evaluation approaches described in the case studies

Data services frequently span traditional library service models (e.g., reference and collection development) and emerging models, such as collaborative service provision with other units. Done well, needs assessment allows RDS providers to act proactively rather than reactively to shifting demands and new opportunities. It can also be a motivator for sustaining conversations and building lasting relationships. Outcome-based evaluation is not well-suited to the task of measuring the use and impact of emerging service models or the value of collaborative partnerships across campus. However, in the area of instructional assessment, RDS providers have a strong base of assessment practice and evidence on which to build. Though many faculty and students acknowledge that the library supports their learning and research, not all libraries have the resources to generate data and conduct analyses suitable for identifying correlations between the use of library services and positive outcomes. Obtaining sufficient expertise and tools, as well as securing the time to carry out quality service evaluation is a persistent challenge. This can be compounded by the need to situate RDS evaluation within existing library assessment and evaluation efforts, as described at the University of Michigan, VCU, and JMU. Providing research data services is a shared effort, both within the library and without. The ability to conduct needs assessments, modify existing tools, or ask library staff to take on additional tasks requires buy-in from colleagues who are often unfamiliar with data services. Collaboration requires mutual understanding, which is dependent on effective communication, the time to develop and maintain relationships, and support from leadership. Supporting good research practices, and good data practices in particular, is a team effort demanding expertise from across an institution. Unfortunately, this reality is frequently not reflected in our planning. Effectively communicating evaluation results to a variety of audiences is also crucial for these efforts to be perceived as worthwhile and effective. This is especially true for gaining administrative buy-in to act upon evaluation results. These case studies echo many of the challenges related to the assessment of digital library collections as described by Troll Covey (2002), particularly the struggle to situate assessment as a core activity and obtain sufficient time and training to do it well.

These case studies highlight a truth of library service. Program evaluation does not simply happen as a by-product of the service—it must be planned (Nitecki, 2004). Conversely, planning for evaluation shapes the program and becomes a part of the program itself (Nitecki, 2004). Service evaluation can be viewed as a form of sense-making that begins as an informal activity rooted in the personal, such as reflective practice. As a professional develops a mental model for the services they provide and the ways they can be evaluated, the evaluation effort becomes more formal and service oriented, rather than focused on the person providing the service. Only when evaluation has become service oriented can the evidence and results be used for comparison or benchmarking. Our literature review suggests that the evaluation strategies most appropriate for research data services may include

reflective practice, evidence-based library and information practice, developmental evaluation, and formative evaluation. Each of these allows data specialists to begin where they are and emphasizes the importance of asking the right questions.

A common purpose for evaluation is internal or external reporting, particularly for justifying resources allocated. In these situations, evaluation questions should be aligned with institutional or library goals so that the evidence addresses locally relevant questions and outcomes that matter. One of those questions is often "what impact does X service have on its users?" Ascertaining impact is a difficult prospect. The needs vary from researcher to researcher, as does the definition of success. Evaluating impact is also influenced by the scope of the services; it may not be possible to address the full range of needs or provide the depth of engagement and support that they desire. Furthermore, the impact of delivering a service, such as assigning a digital object identifier (DOI) to a data set or preservation, may not manifest for years, or even decades. Demonstrating impact is a long-term endeavor. It is vital that we begin to gather this evidence to enable research data services to successfully transition from an emerging service area to a core function of the library.

Since the value of RDS may be unclear to faculty and administrators, we must be prepared to clearly describe how these services support institutional goals and demonstrate the value of library expertise in stewarding the scholarly record. As libraries make this case, we should be careful to describe the full range of resources provided: research support services, instruction, collections and resources developed by the library, access to external resources, staff expertise, consortial and collaborative relationships, facilities, and the library budget. In an environment of uncertainty about the funding of higher education and government research, providing counts of available resources is no longer sufficient for improvement or demonstrating value. Libraries, like the institutions in which we are situated, must be ready to tell a compelling story about how we contribute to society and improve lives. The emergent state of RDS, combined with uncertainty and resource constraints in higher education, demands that we be deliberate in assessing and evaluating RDS so as not to squander available resources—time, money, and the scarce attention of our patrons. We may not have time to wait for the long-term benefits of our services to become apparent.

NEXT STEPS

Given the scarcity of literature that describes current practices for evaluating research support services in academic libraries, an important next step is to conduct an environmental scan of library practice. As we have shown in the case studies above, research data services are not limited to large research institutions, so any scan should be representative of the broad range of institutions offering such services.

All library staff engaged in RDS provision need support from library administration in allocating sufficient time and resources to service evaluation. In conjunction with this need, many librarians will require further training and educational opportunities to conduct, use, and report service evaluation effectively. This is an opportunity for professional organizations to tap into existing expertise to develop educational opportunities such as online courses and in-person workshops.

In many cases, including the case studies reported here, the assessors, decision makers, and users of the evaluation are also the librarians and staff providing research data services. Since our services are intimately tied to the ever-changing research practices of our users (and nonusers), assessment efforts would benefit from the involvement of other stakeholders in the analysis, reporting, and decision-making phases of the evaluation process.

Though metrics for research data services come in a wide variety of formats and are captured across multiple systems, we can mitigate these challenges by applying research data management expertise to our own practice, particularly with regards to documenting our evaluation questions and the processes by which we go about answering and reporting them. We recommend that RDS providers begin to share successful strategies and resources to learn from each other.

The authors believe a community approach is the next step in advancing current practice. To foster connections between existing communities (Research, Data, Access & Preservation Summit (RDAP) community, Research Data Alliance: Libraries for Research Data Interest Group and ACRL Digital Scholarship Section, among others), we propose creating a community of practice using a study group framework, similar to those run by Mozilla Science (https://science.mozilla.org/programs/studygroups). The study group would be held regularly to discuss strategies, challenges, and to share success stories. Paired with a community-owned, shared repository of knowledge and resources in the Open Science Framework, the study group offers a sustainable and convenient venue to advance practice. See the OSF project space (OSF; https://osf.io/3hjeg) for examples described here, proposed goals to start, or to indicate interest in participating.

REFERENCES

Akers, K. G., & Doty, J. (2013). Disciplinary differences in faculty research data management practices and perspectives. International Journal of Digital Curation, 8(2), 5–26. https://doi.org/10.2218/ijdc.v8i2.263

Better Evaluation. (n.d.). Developmental evaluation. Retrieved September 11, 2017, from http://www.betterevaluation.org/en/plan/approach/developmental_evaluation

Booth, A. (2006). Clear and present questions: Formulating questions for evidence based practice. Library Hi Tech, 24(3), 355–68. https://doi.org/10.1108/07378830610692127

Bryant, R., Lavoie, B., & Malpas, C. (2017). A Tour of the Research Data Management (RDM) Service Space (The Realities of Research Data Management). Dublin, OH: OCLC. https://doi.org/10.25333/C3PG8J

Buys, C., & Shaw, P. (2015). Data management practices across an institution: Survey and report. Journal of Librarianship and Scholarly Communication, 3(2). https://doi.org/10.7710/2162-3309.1225

Cahoy, E. S. (2014). Assessment of teaching, learning, and literacies. In Assessing liaison librarians: Documenting impact for positive change (pp. 35–51). Chicago: American Library Association.

Chapman, J., DeRidder, J., & Thompson, S. (2015). Developing best practices in digital library assessment: Year one update. D-Lib Magazine, 21(11/12). https://doi.org/10.1045/november2015-chapman

Cox, A. M., Kennan, M. A., Lyon, L., & Pinfield, S. (2017). Developments in research data management in academic libraries: Towards an understanding of research data service maturity. Journal of the Association for Information Science and Technology, 68(9), 2182–2200. https://doi.org/10.1002/asi.23781

Crumley, E., & Koufogiannakis, D. (2002). Developing evidence-based librarianship: Practical steps for implementation. Health Information & Libraries Journal, 19(2), 61–70. https://doi.org/10.1046/j.1471-1842.2002.00372.x

Fearon, D., Gunia, B., Lake, S., Pralle, B. E., & Sallans, A. L. (2013). Research Data Management Services. SPEC Kit 334. Washington, DC: Association of Research Libraries. https://doi.org/10.29242/spec.334

Federer, L. M., Lu, Y.-L., & Joubert, D. J. (2016). Data literacy training needs of biomedical researchers. Journal of the Medical Library Association, 104(1), 52–57. https://doi.org/10.3163/1536-5050.104.1.008

Gamble, J. A. A. (2008). A developmental evaluation primer. The McConnell Foundation. Retrieved from http://mcconnellfoundation.ca/report/a-developmental-evaluation-primer/

Glynn, L. (2006). A critical appraisal tool for library and information research. Library Hi Tech, 24(3), 387-399. https://doi.org/10.1108/07378830610692154

Henderson, M. E., & Knott, T. L. (2015). Starting a research data management program based in a university library. Medical Reference Services Quarterly, 34(1), 47–59. https://doi.org/10.1080/02763869.2015.986783

Henderson, M., Raboin, R., Shorish, Y., & Van Tuyl, S. (2014). Research data management on a shoestring budget. Bulletin of the Association for Information Science and Technology, 40(6), 14–17. https://doi.org/10.1002/bult.2014.1720400606

Hernon, P., Dugan, R. E., & Matthews, J. R. (2014). Getting started with evaluation. Chicago: American Library Association.

Hufford, J. R. (2013). A review of the literature on assessment in academic and research libraries, 2005 to August 2011. portal: Libraries and the Academy, 13(1), 5-35. https://doi.org/10.1353/pla.2013.0005

Institute of Museum and Library Services. (2000). Perspectives on Outcome Based Evaluation for Libraries and Museums. Washington, DC: IMLS.

Jarvis, P. (1999). The practitioner-researcher: Developing theory from practice. Jossey-Bass Higher and Adult Education Series. San Francisco, CA: Jossey-Bass Publishers.

Johnston, Lisa R., Carlson, J., Hudson-Vitale, C., Imker, H., Kozlowski, W., Olendorf, R., & Stewart, C. (2016). Definitions of data curation activities used by the Data Curation Network. Retrieved June 1, 2018 from the University of Minnesota Digital Conservancy, http://hdl.handle.net/11299/188638

Kloda, L. (2015). Assessment and evidence based library and information practice. Brain-Work: The C-EBLIP Blog. Retrieved September 10, 2017, from https://words.usask.ca/ceblipblog/2015/06/16/assessment-and-evidence-based-library-and-information-practice/

Koufogiannakis, D., & Brettle, A. (2016). Being evidence based in library and information practice. Chicago, IL: Neal-Schuman Publishers.

Koufogiannakis, D., Booth, A., & Brettle, A. (2006). ReLIANT: Reader's guide to the literature on interventions addressing the need of education and training. Library and Information Research, 30(94), 44–51.

Koufogiannakis, D., Slater, L., & Crumley, E. (2004). A content analysis of librarianship Research. Journal of Information Science, 30(3), 227–39. https://doi.org/10.1177/0165551504044668

Lakos, A., & Phipps, S. E. (2004). Creating a culture of assessment: A catalyst for organizational change. Portal: Libraries and the Academy, 4(3), 345–61. https://doi.org/10.1353/pla.2004.0052

Lyon, L. (2007). Dealing with data: Roles, rights, responsibilities and relationships. Consultancy Report. UKOLN. Retrieved from http://opus.bath.ac.uk/412/

Lyon, L. (2012). The informatics transform: Re-engineering libraries for the data decade. International Journal of Digital Curation, 7(1), 126–38. https://doi.org/10.2218/ijdc.v7i1.220

Marshall, J. G., Sollenberger, J., Easterby-Gannett, S., Morgan, L. K., Klem, M. L., Cavanaugh, S. K., . . . Hunter, S. (2013). The value of library and information services in patient care: Results of a multisite study. Journal of the Medical Library Association, 101(1), 38–46. https://doi.org/10.3163/1536-5050.101.1.007

Mason, J. (2002). Researching your own practice: The discipline of noticing. London: Routledge Falmer.

Murphy, S. A., & Gibson, C. (2014). Programmatic assessment of research services: Informing the evolution of an engaged liaison librarian model. In Assessing liaison librarians: Documenting impact for positive change (pp. 105–19). Chicago: American Library Association.

Nitecki, D. A. (2004). Program evaluation in libraries: Relating operations and clients. Archival Science, 4(1–2), 17–44. https://doi.org/10.1007/s10502-005-6989-3

Patton, M. Q. (2011). Developmental evaluation: Applying complexity concepts to enhance innovation and use. New York: Guilford Press.

Samuel, S. M., Grochowski, P. F., Lalwani, L. N., & Carlson, J. R. (2015). The human element of librarianship: Analyzing data management plans: Where librarians can make a difference. Presented at the American Society for Engineering Education, Seattle, WA. Retrieved from https://www.asee.org/public/conferences/56/registration/view_session?session_id=4895

Saracevic, T. (2000). Digital library evaluation: Toward an evolution of concepts. Library Trends, 49(2), 350–69. Retrieved from http://hdl.handle.net/2142/8343

Tenopir, C., Sandusky, R. J., Allard, S., & Birch, B. (2014). Research data management services in academic research libraries and perceptions of librarians. Library & Information Science Research, 36(2), 84–90. https://doi.org/10.1016/j.lisr.2013.11.003

Troll Covey, D. (2002). Usage and usability assessment library practices and concerns. Washington, D.C.: Digital Library Federation and Council on Library and Information Resources.

Zhang, Y. (2010). Developing a holistic model for digital library evaluation. Journal of the Association for Information Science and Technology, 61(1), 88–110. https://doi.org/10.1002/asi.21220