Journal of eScience Librarianship
putting the pieces together: theory and practice

Opportunities and Barriers for Librarians in Exploring Data: Observations from the Data Curation Profile Workshops

Jake R. Carlson

Purdue University, West Lafayette, IN, USA

Abstract

Setting and Objective: The Purdue University Libraries offered a series of professional development workshops in 2011-2012 to provide librarians with a broad understanding of issues in data curation and to train them in the use of the Data Curation Profiles Toolkit (DCP Toolkit). An additional goal of the workshop was to develop a better understanding of the experiences, attitudes, and needs of librarians as they explore new roles.

Design and Methods: Workshop participants were asked to complete three surveys: one before the workshop, one right afterwards, and one delivered three months after they had attended the workshop. Participants' responses to the survey questions that pertained to demographic information, confidence levels, and levels of engagement before and after the workshop were reviewed and analyzed.

Results: The results of the survey indicated that participants' confidence levels in multiple areas increased after the workshop, but that their levels of engagement remained relatively stagnant. An analysis of the free text comments made in the survey revealed multiple issues in librarians' efforts to increase their engagement in working with data including time, staffing, and organizational support from their library.

Conclusions: The challenges encountered by librarians seeking to engage in data management and curation issues are found at the individual level (acquiring skills and confidence) and at the organizational level (creating a supportive environment). Both levels will need to be addressed by libraries seeking to develop data services.

Introduction

In recent years there has been a surge of interest in librarians taking on roles and responsibilities in providing data services to their constituencies. Many in the library field believe that librarians are a natural partner in addressing issues in research data management and curation given our knowledge and skills in organizing, disseminating, and preserving diverse sets of materials. Working with research data may provide an opportunity to reinvigorate libraries' roles within the research operations of the institutions they serve.

The perceived potential of librarians to take on roles and responsibilities in data has led to a proliferation of articles and reports on the responsibilities that librarians could take on and the skills that they would need to acquire to do so. In response, professional development opportunities are being developed and offered by library schools and organizations. However, it is not yet clear to what extent librarians who participate in

these professional development programs are becoming more active and engaged in working with data. Is the training that librarians have received from these programs making a difference in getting more librarians actively engaged in data management and curation issues?

This article reports on the levels of engagement seen from librarians who have attended the Data Curation Profile Toolkit Workshop (DCP Workshop). The DCP workshop is a professional development program offered by the Purdue University Libraries, and is designed to increase levels of engagement by librarians with faculty at their respective institutions on data management and curation issues. Workshop participants were invited to complete three surveys – given before and after the workshop – to measure the change in their level of engagement with data producers at their respective institutions. The results of the surveys provide some insight into the opportunities that librarians see for engagement as well as barriers they have encountered that limit what they are able to do.

Literature Review

There is no shortage of articles and reports that describe potential roles for libraries in managing and curating research data, and the skills that librarians may need to acquire to do so. In the middle of the last decade, the Association for Research Libraries (ARL) produced two reports that set the stage for discussing potential roles for libraries, as organizations, in supporting the emerging research paradigm of e-Science. Both reports advocated for libraries to develop infrastructure and services for the curation and preservation of research data, and to develop partnerships with faculty, research communities, and other entities to develop and maintain these services (Friedlander and Adler, 2006; ARL, 2007). At around the same time, Anna Gold surveyed data librarianship and recommended that libraries expand into new roles. She suggested that

libraries continue to build on efforts to develop roles in the post-publication or "downstream" side of the research cycle which would include crafting services towards supporting data publication as well as building collections of published data sets. Gold also focused on roles for libraries in the "upstream" phase of the research lifecycle, prior to publication where data are still being generated and analyzed. She believes that librarians could forge active collaborations with researchers and contribute their expertise towards the adoption of standards, the development of documentation, and the creation of workflows that would support the eventual dissemination and curation of the data being produced (Gold, 2007). Other potential roles for libraries put forth in the literature include raising awareness and providing training in good data practices (Swan and Brown, 2008); developing data repositories (Hey and Hey, 2008); and helping researchers craft data management plans (Steinhart, et.al., 2010).

Conceptualizing new roles and responsibilities for libraries naturally leads to questions about what knowledge and skill sets librarians would need to acquire to carry them out. These conceptions are typically based on the perception that librarians possess a foundational base that can be leveraged beyond the traditional formats of books and journals and applied towards addressing issues in data management and curation (Mullins, 2007). However, there is a wide variation in the literature as to how the knowledge and skills of a librarian could be applied as well as on what additional skills they need to develop. One possible example is that librarians could apply their communications and negation skills towards coordinating data practices across their institutions (Pryor and Donnelly, 2009). Some see the need for librarians to be able to acquire a deep understanding of the repository system and other technologies being used to curate data, even if they are not developing or managing these systems themselves, and for librarians to be much more aware of and at-

tuned to the research process (Newton, Miller and Bracke, 2010). Others see a need for librarians to be able to acquire strong analytical, project management, and problem-solving skills (Gabridge, 2009). Still others see a need for librarians to possess knowledge of selection and appraisal techniques, metadata standards and schema, ontologies, data formats, identifiers, citation practices, and existing data centers (Lyon, 2012). Given the variety of roles discussed and the myriad of skill sets needed to perform these roles, it is not terribly surprising that e-science and data librarianship has been described as nebulous and undefined (Alvaro, et. al., 2011).

There are now multiple libraries that are offering services and support for data management and curation, and many other libraries are taking actions to explore how data services could be incorporated at their respective organizations. Many of these libraries have written up and published their approaches and experiences in developing data services, including the Georgia Institute of Technology (Walters, 2009), Purdue University (Brandt, 2007), and Cornell University (Steinhart, Saylor, et.al, 2008). However, what tends to be documented in these publications are the nature of the services and the needs that they are intended to address. Much less has been written on the experiences and pathways of the librarians who have made the transition into incorporating data services as a part of their job responsibilities.

As the success of a library's initiatives developing data services will depend upon the ability of librarians to successfully adopt new roles and responsibilities, it is worth conducting additional investigations into how librarians have responded to opportunities to engage in data management and curation issues as well as the barriers that they have encountered. Garritano and Carlson described the steps they took and the challenges they faced in working with a research center at Purdue to develop a data workflow

to capture and curate data (Garritano and Carlson, 2009). More recently, a group of librarians at the University of Oklahoma conducted a survey of science librarians employed at institutions affiliated with the ARL that included a section on their preparedness to assume responsibilities for providing data services. They found that the surveyed librarians had some mixed emotions in taking on these roles. Some viewed working with data as a natural evolution in the responsibilities of a science librarian, while others saw data as a separate and distinct role. Overall, the librarians interviewed expressed uncertainty about what data initiatives and resources were available at their institution and about what skills they would need to acquire in order to work with data (Antell, Foote, Turner and Shults, 2014). Another survey conducted on the preparedness of librarians to assume responsibilities with data, their attitudes towards doing so, and their levels of engagement found that professional motivations to become involved were high overall and attitudes towards librarians taking on responsibilities with data were generally positive. Survey responses were categorized according to how integral data services were to the librarian's job responsibilities. Not surprisingly, respondents who indicated that data was integral to their job gave more positive responses to questions about possessing the necessary skills, knowledge, and training to work with data and towards the amount of support received from their library than respondents who indicated only some or no involvement with data services (Tenopir, Sandusky, Allard and Birch, 2013).

Background

The DCP Toolkit was developed from research conducted by the Purdue University Libraries and the Graduate School of Library and Information Science at the University of Illinois at Urbana-Champaign with support from the Institute of Museum and Library Services (IMLS) (Witt, et al., 2009). It is designed to help librarians begin discussions

with faculty and students about their research data and needs in managing and curating their data. By conducting data interviews as described in the DCP Toolkit, librarians are able to gather information about a particular data set being developed or used, how that data set is being managed currently, and about the faculty or student's needs in managing or curating the data set. Information from the interview is then used to complete a Data Curation Profile (DCP). A DCP is a structured framework for organizing and sharing the information gathered in the data interview which enables this information to be used as a foundation for developing initiatives or services to address the needs expressed by the interviewee. The shared structure of DCPs is intended to make the content easy to read for any potential service provider and to enable DCPs to be compared across disciplines and fields of study to gain a larger understanding of practices and needs relating to data. Published DCPs are freely available from the Data Curation Profiles Directory: http://docs.lib.purdue.edu/dcp. More information about DCPs and access to the DCP Toolkit can be found on the Data Curation Profiles Toolkit website: http://datacurationprofiles.org.

Methodology

The Purdue University Libraries received a grant from the Institute of Museum and Library Services (IMLS) to offer 12 workshops to train librarians how to use the DCP Toolkit. Developed by the author and D. Scott Brandt, Associate Dean for Research at the Purdue Libraries, the DCP Workshop was specifically designed for practicing librarians seeking to become more involved in working with research data. No previous knowledge or experience in working with data was required. We offered these workshops at academic libraries and conference venues across the United States in 2011 and 2012. The DCP Workshops were co-sponsored by regional library associations and others to promote the workshops to their

members and manage registration.

Our objectives in developing the DCP Workshop were:

- To provide librarians with a broad understanding of data curation issues, especially as they relate to Libraries.

- To introduce the Data Curation Profile Toolkit as a tool for librarians/archivists to use at their host institutions, and train them in how to use the DCP Toolkit.

- To better understand the experiences, attitudes, and needs of librarians as they explore new roles relating to working with research data.

In order to gauge our success in achieving these objectives we developed three surveys which were created using Qualitrics survey software and delivered electronically to the email account provided to us by the workshop participant. The surveys were reviewed and approved by Purdue's IRB and participants were asked to read and agree to a consent form before taking each of the surveys.

The first survey was delivered to workshop participants approximately two weeks before their attendance at the DCP Workshop. This survey contained questions about the background of the participant, including their education and position in the library. We also included questions regarding confidence levels when engaging with faculty on data issues. We sought to ascertain a participant's current level of engagement in working with research data by asking them what professional responsibilities they had pertaining to data and if they had ever participated in certain activities, and if so to what extent. The activities pertaining to data that we asked about included reading articles or other literature, attending presentations or training events, engaging in discussions, and working with faculty or other university personnel. This information served as a baseline from

Figure 1: A count of the professional position types listed by DCP Workshop partici-
pants.
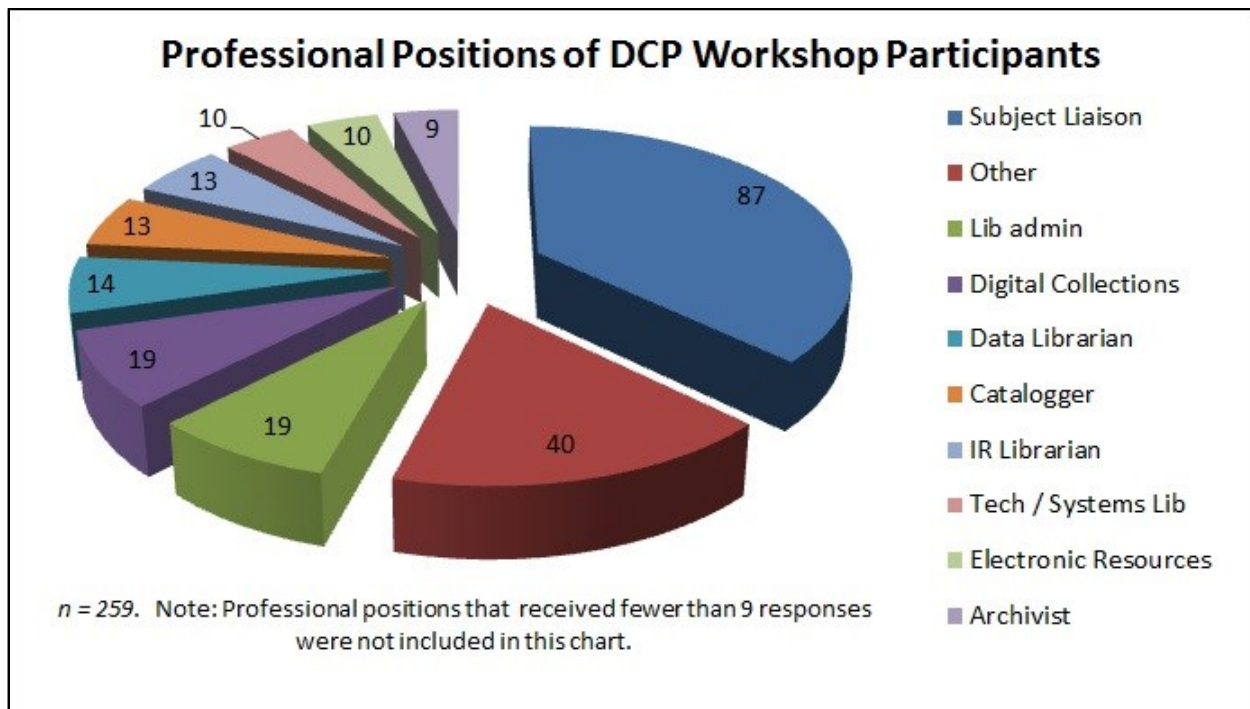
which we measured the effect of attending the workshop on participants.

The second survey was delivered immediately after the workshop was offered. The purpose of this survey was to collect feedback about the effectiveness of the workshop as well as to gauge the levels of interest of attendees in pursuing action in working with data. To this end we repeated our questions about confidence levels of attendees from the first survey. We also asked participants how likely they were to put their training into practice by conducting a data interview of their own and then developing a DCP from the interview.

The third survey was delivered to participants three months after they had attended the workshop. We designed this survey to see to what extent participants had made use of the information they learned from the DCP workshop. We asked participants to
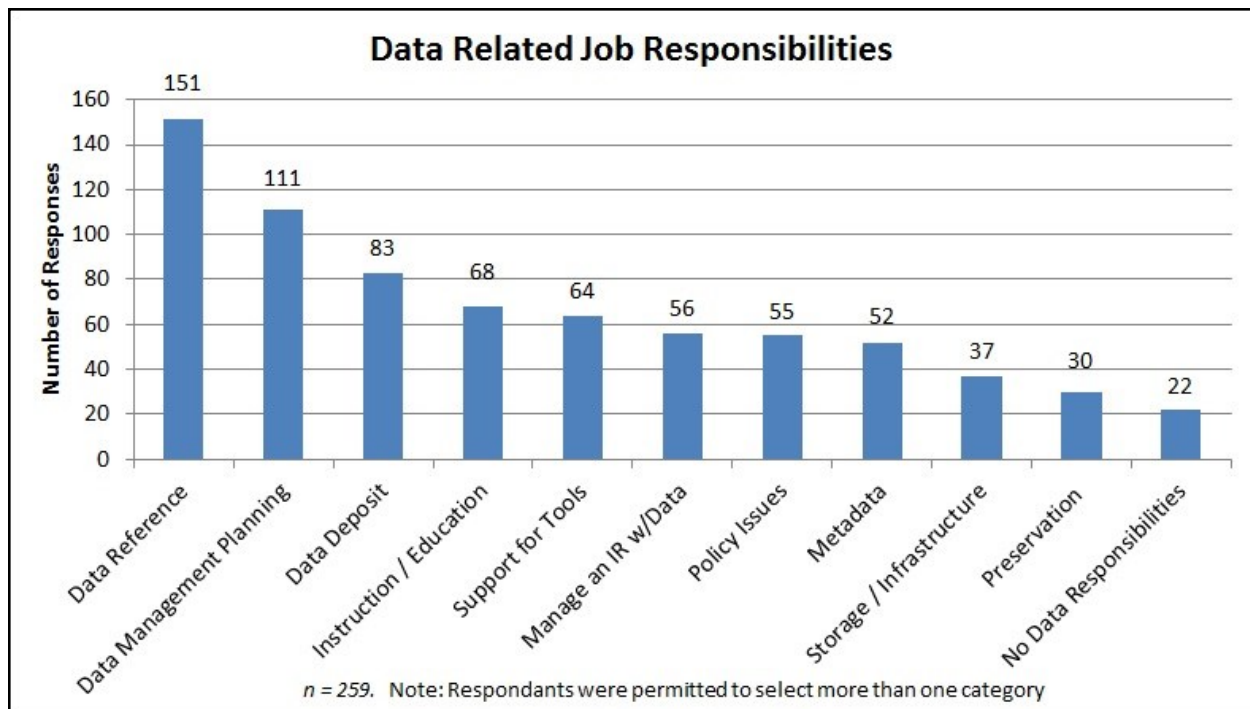
characterize their levels of engagement in data since attending the DCP Workshop, as well as the engagement levels of their library more generally. We also once again returned to questions that were asked in the first survey about confidence levels in engaging faculty to be able to speak to the lasting impact of the DCP workshop.

Participants in this study were self-selected based on their attendance of the DCP Workshop, so the results of the surveys cannot be generalized to the larger population of academic librarians. Nevertheless, we feel that the results present an interesting portrait of librarians who are seeking to explore roles or acquire responsibilities in working with data.

Survey Results

The data collected from the three surveys were exported from Qualtrics into Excel spreadsheets for analysis by the author. Ex-

Figure 2: Assigned responsibilities pertaining to data currently performed by DCP Workshop participants.

Data Related Job Responsibilities

n = 259.   Note: Respondants were permitted to select more than one category

cel was used to quantify the results from questions where participants were asked to select a response (or responses) from a list of possible responses and from questions where participants selected a response on a likert scale. A total of 324 people attended one of the twelve DCP workshops. Of these attendees, 263 responded to the first survey delivered before the DCP workshop for a response rate of 81%. The second survey delivered immediately after the workshop was completed by 206 attendees for a response rate of 64%. Finally, 147 participants filled out the third survey given three months after the DCP workshop for a response rate of 45%. Out of all of the workshop participants, 119 of them, or 37%, completed all three surveys. It should be noted that not everyone who completed a survey answered every question.

Demographic Information

As reported in the first survey, a wide variety of librarians enrolled in the workshop. The

responses to the question "Please select a title from the list below that best describes your professional position" are shown in Figure 1.

The largest contingent of librarians at the DCP Workshop was subject liaisons. The next largest category was "other," indicating the difficulty of capturing the proliferation of job titles and responsibilities in surveys like this. Participants who selected "other" were asked to provide their job title. Although job titles were varied, many of them implied responsibilities relating to some aspect of public service, such as "Education Librarian," "Instruction & Outreach Librarian," "Information Services Librarian," and "Public Services Librarian." Interestingly, "Data Librarian" was only selected by 14 participants.

We also asked participants if they had any responsibilities in providing a data service of some kind. Figure 2 displays the responses from this question.

Figure 3: Changes in the Average Confidence Levels of DCP Workshop Participants.
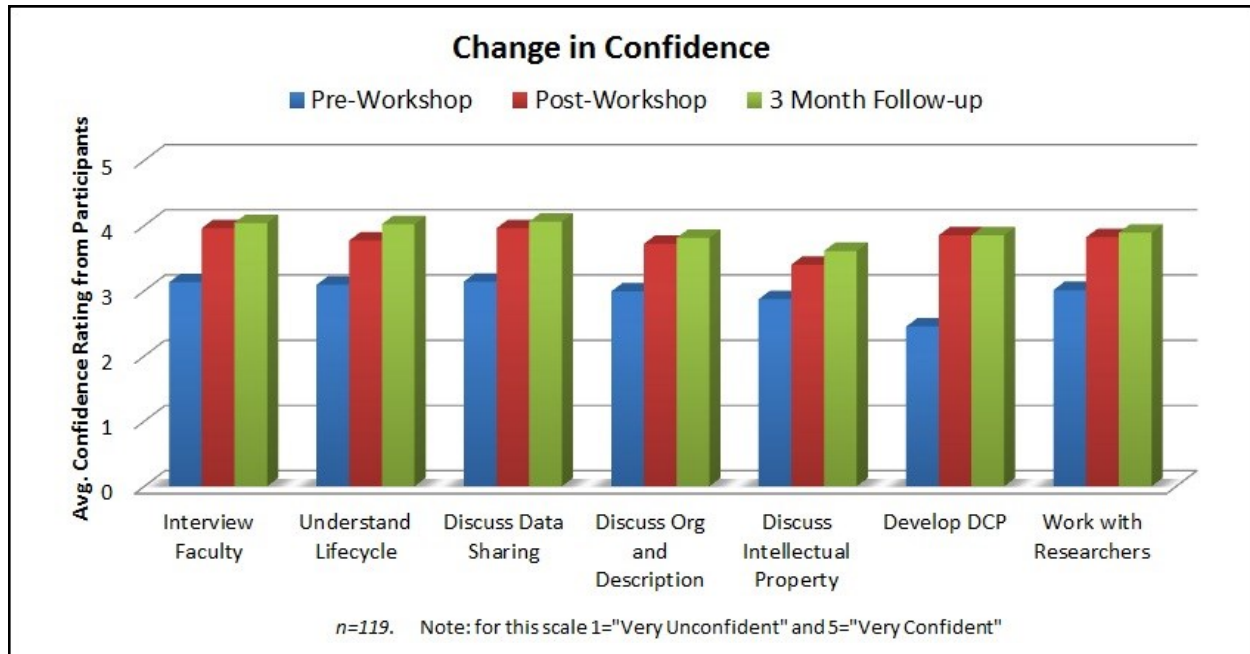
Although Figure 1 demonstrates that very few of the DCP workshop participants were in positions that were dedicated to working with research data, Figure 2 demonstrates that the vast majority of participants are tasked with responsibilities for data in one or more areas (participants were allowed to select multiple options). Only 22 of the 259 participants did not select any of the options in the survey, indicating that they did not have any data responsibilities. All of the other participants selected at least one area of current responsibility. Not surprisingly, given the high proportion of subject liaisons and other public service-driven librarians attending the DCP Workshop, the responsibilities pertaining to public service: data reference, data management planning, instruction / education, were among the top responses. It is not clear from the survey if participants were formally assigned these responsibilities with data or if they occurred as a natural component of fulfilling broader responsibilities such as reference or instruction more generally. Several participants did note that they had not yet started performing these responsibilities though they were planned, or that they

were just getting started in taking them on.

Impact of the Data Curation Profiles Workshop

We had asked a series of questions about the confidence levels of participants' abilities in each of the three surveys as a means to measure the lasting impact of the DCP Workshop. For example, we asked participants: "How confident are you in your ability to distinguish between stages in a data life cycle?" We used a likert scale for this question with a response of 5 to mean "Very Confident" and a response of 1 to mean "Very Unconfident." The average results from all three surveys are depicted in Figure 3.

To reduce potential noise in the data, the sample size for this question was filtered down to the 119 participants who answered all three surveys. The results show the average reported confidence levels of participants increased from before they attended the DCP Workshop to immediately afterwards. Furthermore, the increased confidence levels were still evident from respons-

Figure 4: Responses of DCP Workshop participants to the question: "I have attended conferences or events that were focused on data curation" in percentages.
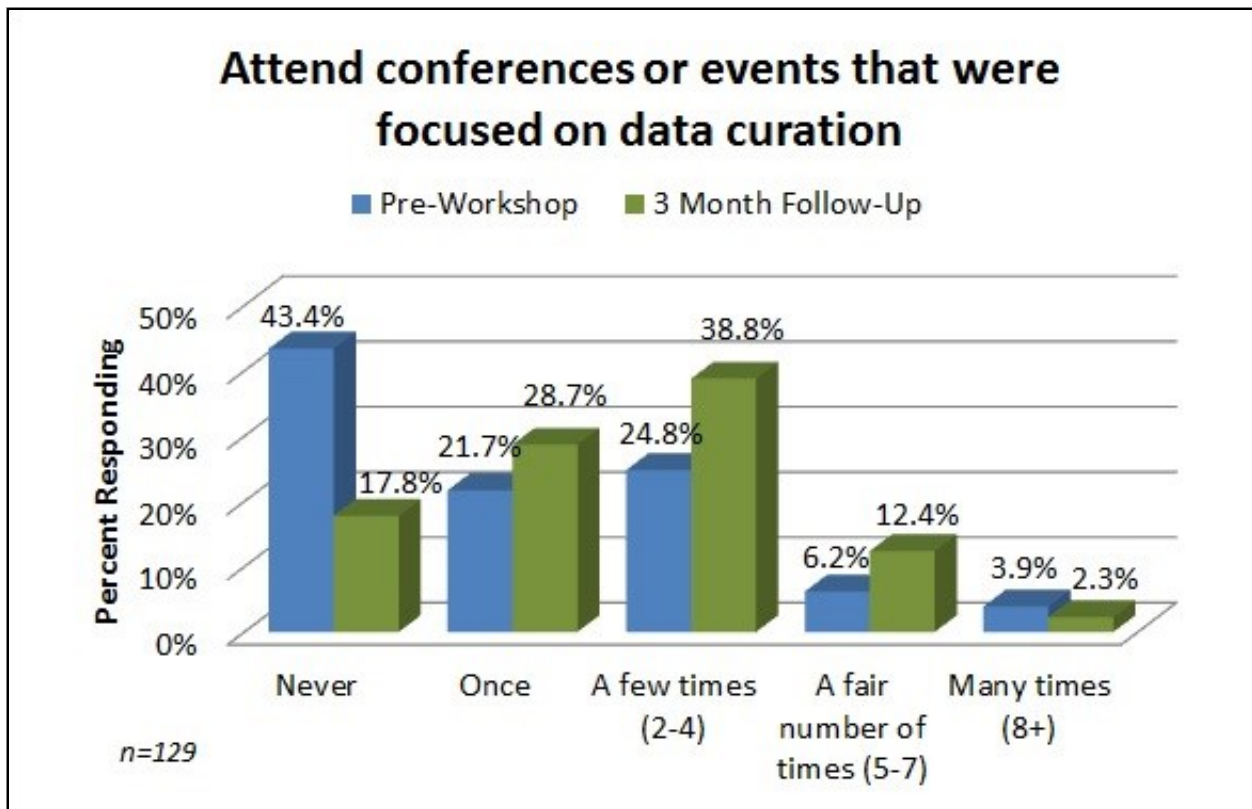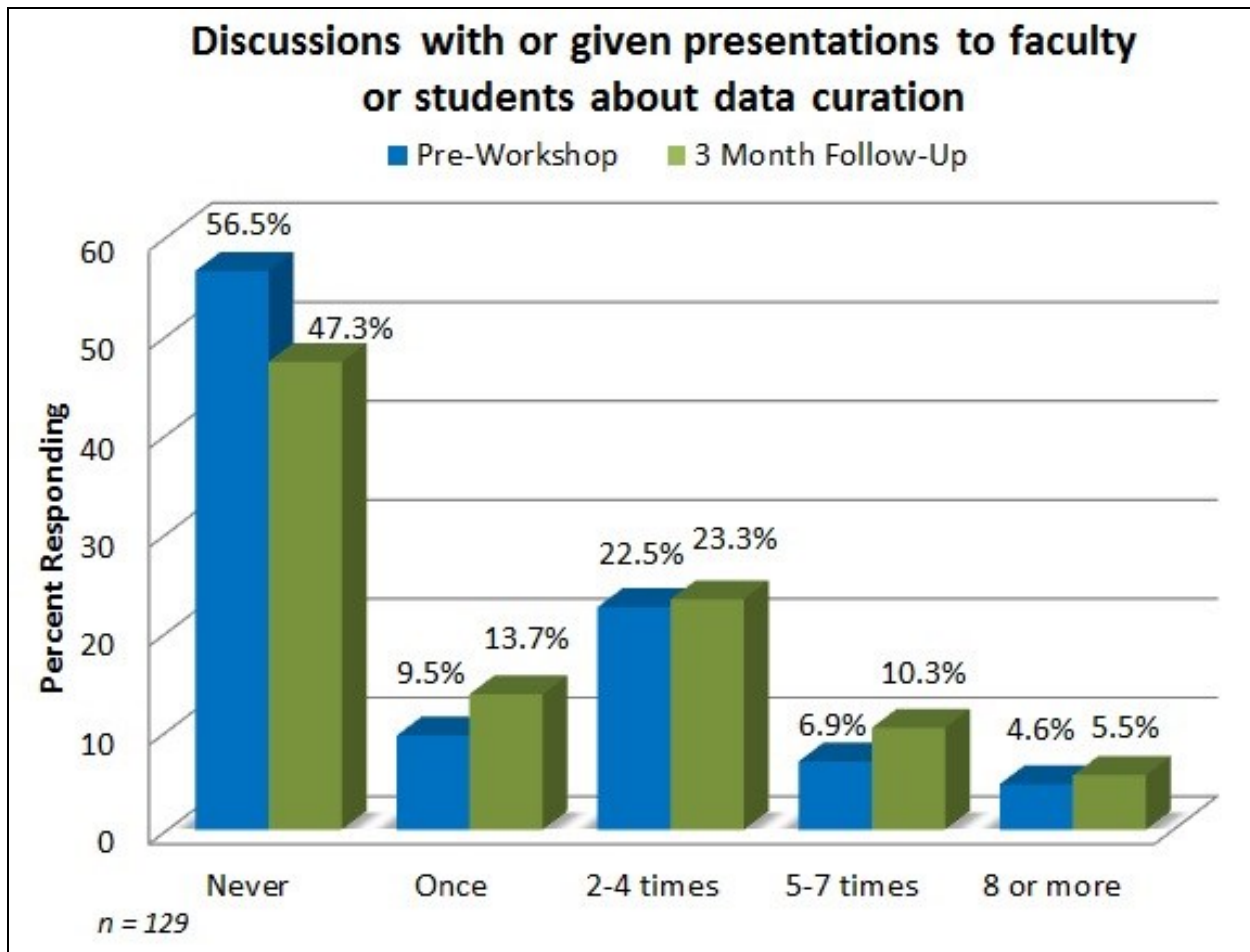

es given by participants three months later. To determine the statistical significance on the results in each of the seven questions on confidence levels across the surveys, data was exported from Excel into SPSS statistical software. Using SPSS, a repeated measures ANOVA test was run on the data. The results demonstrated a significant difference between the answers on all seven of the questions on confidence levels given by participants in the pre-workshop survey and the post-workshop surveys ($p < .001$ for all).

Finally, in the surveys we asked a series of questions to determine the participants' level of engagement before and after DCP the workshop. Here again, the sample size was filtered down to just the participants who responded to both the pre-workshop survey and the three month follow-up survey which worked out to be 129 people. Unfortunately, the low number of responses in some of the categories prevented the author from conducting statistical analyses on the results. Nevertheless, the results provide an interesting glimpse into how engagement by librarians in data is unfolding in academic libraries. An example from this series of questions on the engagement of librarians in data curation is displayed in Figure 4. The results indicate that participants have increased their engagement in professional development activities in data management and curation issues in the three months since attending the DCP Workshops.

Other responses from this series are represented in Figures 5 and 6. In contrast to the increases observed from the question on professional development activities in Figure 4, Figures 5 and 6 convey that levels of direct engagement with faculty and students by librarians on data curation issues have stayed relatively stagnant amongst

Figure 5: Responses of DCP Workshop participants to the question: "I have had discussions with or given presentations to faculty or students about data curation" in percentages.

Discussions with or given presentations to faculty or students about data curation

Pre-Workshop   3 Month Follow-Up

Never: 56.5%, 47.3%
Once: 9.5%, 13.7%
2-4 times: 22.5%, 23.3%
5-7 times: 6.9%, 10.3%
8 or more: 4.6%, 5.5%

n = 129

participants after attending the DCP Workshop.

Analysis of Free Text Comments

The results of the survey indicate that, although participants in the DCP Workshop gained confidence in their knowledge and abilities to interact with researchers and have been more active in professional development activities on average, they have not yet fully translated their increased confidence and interest into action. In the three month follow-up survey, we asked an open-ended question to participants on where they saw opportunities and barriers for their engagement with data as well as the engagement of their libraries. The responses we received provided some possible explanations for our findings on the engagement activities of the participants in the DCP Workshop.

The 107 responses received to this question were analyzed, and identified the major thematic categories that were being expressed by the participants in their responses. The author then analyzed the responses once again and assigned each of them into the categories that were identified. Some of the responses from participants contained multiple issues which were broken down into dis-

Figure 6: Responses of DCP Workshop participants to the question: "I have worked directly with faculty or students on addressing data curation issues" in percentages.
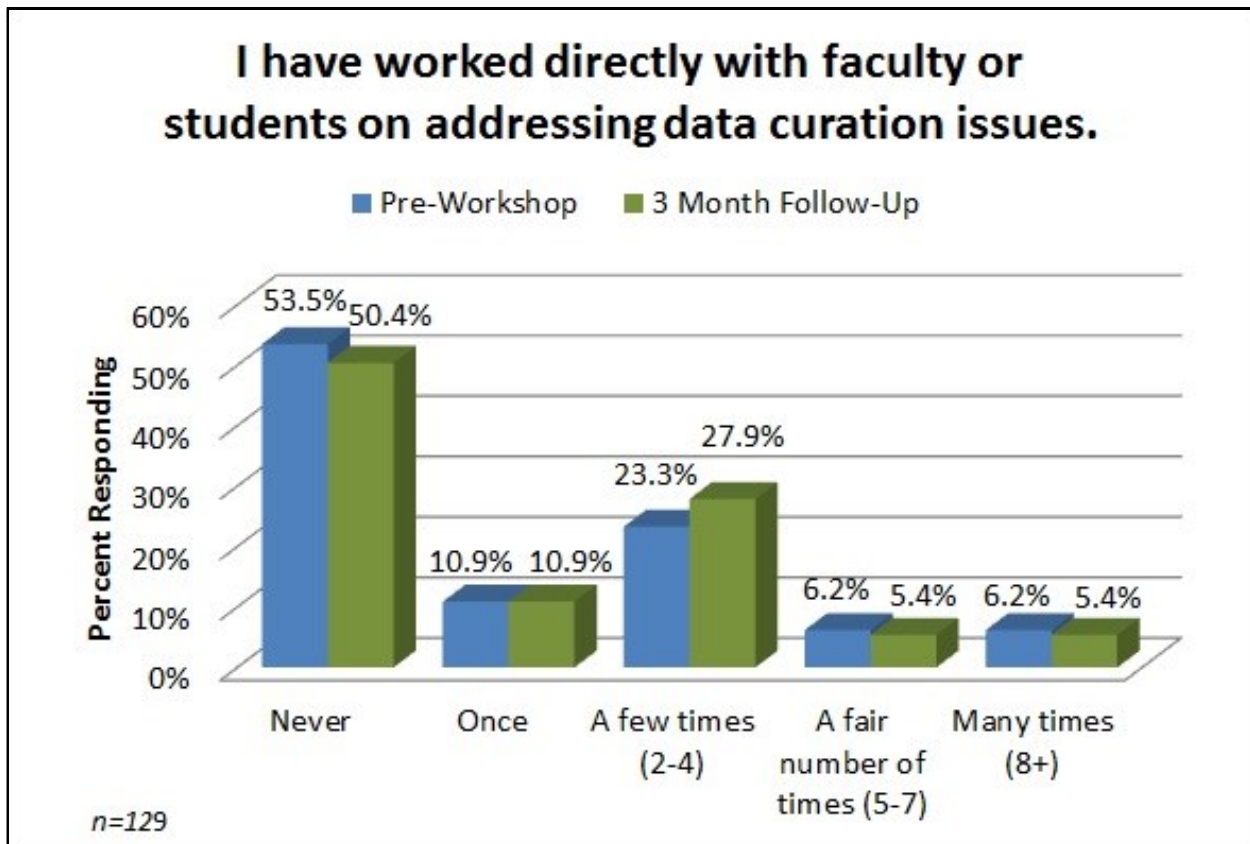
crete statements, and each statement was then assigned into a thematic category independently from each other. A total of 140 distinct statements were identified and categorized. Finally, each statement was reviewed once more to determine if it had been expressed as an "opportunity" or as a positive sentiment, a "barrier" or as a negative sentiment, or if the expression made expressed neither a positive nor a negative sentiment but was "neutral" in its tone. Occasionally, an individual statement contained both positive and negative sentiments in which case it was listed as a "barrier and opportunity."

In analyzing these comments, six overarching themes on barriers and opportunities became apparent: organizational support, time, connections with other units in the institution, staffing, resources/money, and participation in the ARL e-Science Institute. Out of these six, four of these themes were expressed largely as barrier issues for participants: organizational support, time, staffing, and resources / money. A fifth theme, connections with other units at the institution, leaned slightly towards the barrier sentiment but contained a number of expressions of opportunity as well. A sixth theme, participation in the ARL e-Science Institute, was expressed largely as an opportunity. Other statements made by participants did not put forth a type of opportunity or a barrier so much as indicate the respondent was taking action, considering taking action, or planning to take action. Finally, a few participants indicated that it was not a part of their job to engage with researchers or that engagement with data was outside of the scope of their library. The results of the analysis are represented in Table 1.

Table 1: Results of an analysis on the survey question: "Please comment briefly on the opportunities or barriers to your own engagement or to your libraries engagement with data curation issues."

Response Category	Number of Statements Made by Sentiment				Total # of Statements
	Barrier	Opportunity	Barrier & Opportunity	Neutral	
Organizational Support	20	7	1	3	31
Time	24	0	0	1	25
Connections with Other Units at the Institution	12	9	1	2	24
Staffing	11	4	0	3	18
Resources / Money	9	0	0	0	9
Participation in the ARL e-Science Institute	0	6	0	1	7
Other Categories(# of Responses): Taking Action (10), Not my Job (6), Out of Scope (4), Considering Action (3), Planning Action (3)					

Issues with Organizational Support from the Library

The most common theme that arose from the analysis of the open-ended question on opportunities and barriers to engagement was that of organizational support from the library, which was mentioned 31 times. Twenty participants, or roughly two thirds of those who made a statement in this area, expressed that the current state of their respective library organizations presented barriers to taking action. Data curation has been a popular topic of discussion in the library field and there is some pressure to plug into the conversation and stay informed of developments. However, library administrations must juggle keeping up with the latest developments in the field with determinations of how much to invest in any given area. Several participants indicated that their libraries' initial interest had not yet translated into devoting staff time and efforts towards taking action, as evidenced by the following quotes:

"Our library has been working on other topics and projects. There is not as much interest in the library as I had first thought, for this topic."

"There are a lot of competing priorities and data
curation has not been given a high enough priority for us to move forward with it so far."

Other participants in the DCP Workshop expressed a desire to move ahead in acting on their interests in working with data, but felt that they could not yet pursue their interests as their libraries had not yet decided on how to proceed in this area. If their library, as an organization, was not prepared or ready to move forward, the ability of individual librarians to act on what they have learned from the DCP Workshop is limited.

"… I feel better equipped to work with [faculty], (based on the workshop) and reading more, but my library is not ready to move ahead. I cannot at this point, move ahead of that."

"…there's only so much planning we can do without the support and leadership of our wider library system."

A few of the participants expressed some frustration with their libraries' administration, stating that the organizations themselves were creating barriers towards engagement.
"Disorganization on some of the assignment of tasks. Being able to contact faculty and start work on a data curation project without intervention from administration."

On a more positive note, other participants indicated that their libraries were creating opportunities by incorporating data into its planning process or other actions.

"Exploration of data curation issues by way of a pilot project is included now in our Library's Strategic Plan for 2012-2014."

Issues of Time

The second theme that emerged from the analysis was centered on time. When expressed in participant responses, time and timing-centered issues were almost universally seen as barriers to their engagement with faculty or students on data issues. Several participants mentioned that their day-to-day duties or other responsibilities preclude acting on their interest in data. The following quotes exemplify a common response:

"Time, time, time; the spirit's willing, but the calendar isn't cooperating."

"There is a great deal of energy surrounding this, but no concurrent increase in available time…"

Other participants expressed more nuanced concerns relating to time. For some, it was not clear where they should be directing their efforts to achieve a clear return on investment to justify the use of their time.

"It's mostly just finding the time to work on this, and the right places to focus energy. The library, and university, just got more budget cuts, so that's not helping."

Still other participants expressed some doubt that they had the knowledge or ability to respond to the needs of the researchers that they would be engaging. They believed that initiating discussions with researchers could result in their being overwhelmed; especially as they have other responsibilities to fulfill in the library. Furthermore, the time it would take to feel fully prepared to engage researchers is prohibitive.

"A bit afraid to open the floodgates."

"Barriers to my own engagement include my own feelings that I lack sufficient expertise, and sufficient time to develop expertise, given my other duties not related to data curation."

There was also mention of the amount of time that was required of faculty. The DCP Toolkit is an interview-based tool and interviews, especially ones that are in-depth like the DCP, can take a fair amount of time. Although the DCP Toolkit can be modified and shortened as needed to suit the situation, it is still an instrument designed for depth more than breadth. Some participants felt that the DCP Toolkit required a large investment for an organization that was just getting started and trying to find its way.

"I have been conducting interviews with researchers about their data services needs and knowledge. I will be using this information to determine how the library can help researchers with their data needs. The biggest barrier is the amount of time it takes to gather the information since the interviews are best done on a one-to-one basis."

"I think the tool has been well thought out for your needs. We are not at a point where we need such an in-depth tool. I do foresee using some of the toolkit to do something a bit more low key and less time intensive to get us started."

On a similar note there was some trepidation about approaching researchers "too soon." Some participants felt that the services and resources to respond to researcher needs had to be in place or well developed before they could readily engage with researchers.

"*We're just gearing up our repository, so it's premature to approach researchers."*

Issues with Staffing

A sizable number of participants brought up staffing levels or support when responding to the question of opportunities and barriers in engaging researchers. Several participants brought up the lack of a sufficient number of staff as an impediment to their own engagement as demonstrated in the following

response:

"We are incredibly short staffed… and there is no time for deep engagement in anything."

However, some participants mentioned that positions were being created that centered on e-science and/or data curation librarianship. These positions will likely address the lack of capacity in engagement on data issues, but some participants noted that hiring data librarians may have the effect of getting themselves or others in the library more involved.

"…We were also holding interviews for the e-science librarian position recently and they person will report to me/my unit and so in order to prepare for our hire, I've been more involved."

In contrast, others described situations in which they were "lone wolves" of sorts. They, as an individual or as a part of a small group, are interested in engaging more with researchers on data, but they have other responsibilities and a limited amount of support from other staff in the library. Given their situations, they expressed hesitation in doing more than they were currently doing.

"I still have a lot of my pre-existing duties to deal with. And although my supervisor is very interested in data curation, most of my coworkers are not."

"There are only a few of us engaged in data curation issues, and we all have additional responsibilities. More people in the library need to be involved."

Issues with Resources and/or Money

A number of participants also mentioned resources or money as an issue for themselves or their library in their further engagement. The statements on resources or money made by participants were uniformly negative in their sentiment with participants replying that the lack of resources or money presented a barrier to their efforts. Statements on the lack of resources or money were often made in conjunction with the other identified themes: staffing, time, and connections with other units in particular.

"The primary challenge is lack of available resources: especially staff time and server/storage source."

"…A barrier to our libraries' engagement with data curation issues is the lack of IT infrastructure on our campus, and the insufficient funding in the libraries to provide the tech infrastructure and quantity/quality of personnel expertise."

Issues with Connecting to Other Organizational Units in the University

Addressing issues in managing and curating research data is a complex task that will almost certainly involve multiple actors in any one institution. Some of the workshop participants made note of these challenges in reflecting upon their opportunities and barriers to engagement. There were different interpretations of how the need for partnerships across the institution affected libraries. Some focused on the time and effort it would take to forge these relationships and produce results as barriers.

"We are actively searching for an administrator who will help shape our library's engagement with campus data curation issues…it will likely take a very long time to have solutions to many data curation needs on campus so we are caught in many cart before the horse scenarios."

Others saw the other actors as potential competition to the library (or believed that others at the institution would see things this way).

"[There is] the perception [at the institution] that data curation is strictly an IT thing, and the IT folks are so busy dealing with the day to day challenge of keeping the network running, that they are not thinking at all about possible solutions for curation."

"The library identifies a need for centralized data curation but we are but one option for taking care of data. There are plenty of competing parties interested in doing the same…"

However, the views of participants in connecting to other units were not always negative. Some librarians saw their libraries creating opportunities for themselves through making connections to other units on campus.

"Since the workshop, the library with other partners developed a report on the infrastructure needed to better support data curation and related data issues at our institution."

"New center for data management, curation, and dissemination is opening up at our institution. Director is considering a possible role for the library, armed with a DCP Toolkit."

Discussion

The results of the surveys of participants at the DCP Workshops suggest that there is a disconnect between the levels of interest expressed by librarians, public service librarians in particular, in getting more involved in data issues and their perceived ability to act on their interest. Furthermore, the results suggest that a lack of engagement is less likely due to a librarian's confidence in her ability to engage with researchers about data and more likely due to perceived barriers: available time to invest in this area, sufficient staffing levels in the library, support from other staff and their libraries' administration, a lack of readiness or leadership in their library to move forward, a perceived lack of resources, or weak connections between the libraries and other units on campus. The convenience sampling used in these surveys prevents any firm conclusions from being drawn from the results. However the analysis presented here does raise questions, not just in how we educate librarians to engage in data management and curation issues, but in how we educate our libraries as organizations to support these types of engagements.

The need to change traditional models of library organizations is not a new argument. Just as there has been a multitude of articles and reports on roles for librarians in managing and curating data, there has been a steady stream of articles and reports that articulate how libraries as organizations need to evolve in order to meet the demands research practices in the 21st century. For example, findings in a recent study on how future leaders in the library field view the organizational culture of their libraries revealed that future library leaders felt limited by the current culture of their organization and would prefer a culture that was more externally focused and flexible (Maloney, Antelman, Arlitsch and Butler, 2010). Others argue that libraries seeking to provide e-science-based services ought to mirror their organizational structure on the multidisciplinary environments of research centers and should strive to embed librarians within the research teams that they support (Luce, 2008). The underlying theme in these and other publications is that libraries must reconsider their organizational structures and cultures to able to take on data management and other innovative service areas successfully.

However, much of the current literature tends to speak at the level of ideas for change and less on how change actually occurs in libraries. In considering roles and responsibilities for themselves, many individual librarians who attended the DCP Workshop expressed a desire to understand how to engage in data initiatives and services at a very practical level. In other words, how does one actually "do data?" The same questions should be asked in reshaping library organizations, environments, and cultures. How does a library organization recast itself to support the development of data services and the librarians tasked with offering these services as a normative part of day-to-day operations? What organizational structure does a library need to develop to provide support for embedding librarians into research projects, or even just for supporting librarians in launching discussions with faculty about their needs for their data? Are there particular approaches, strategies or criteria that could be applied to rethinking libraries

as organizations to support active engagement on data issues at all levels?

One recent exploration towards developing answers to these questions that was brought up in the survey results was the e-Science Institute. The e-Science Institute was developed by the ARL and the Digital Library Federation (DLF), then continued through a partnership between the DLF and DuraSpace. Its purpose is to help research libraries explore the landscape of e-Science as practiced at their respective institutions and to develop a strategic agenda that would drive their organization's response. The primary intent of the strategic agenda is to encourage the library as an organization to consider how it could respond to the needs of e-Science practitioners at their institution and articulate how the organization would need to change to accommodate and support these responses. This includes issues in data management and curation in addition to other aspects of e-Science such as supporting virtual research environments, or data visualization services by the library. Most participating institutions send a team of three to the institute: a data librarian, a library administrator, and a member from their institution from outside of the libraries, such as a faculty member, university administrator, or an IT professional. The composition of these teams, and the nature of the institute itself, is specifically designed to encourage thinking beyond the current structure of the traditional library organization.

The DCP Workshop participants who mentioned the e-Science Institute in their responses largely saw the experience as an opportunity for themselves and for their libraries. One participant identified connections between preparing individual librarians to engage with researchers on data through the workshop and addressing the structure of the library organization through the e-Science Institute:

"The ARL is currently running an E-science Institute… My supervisor is involved in the institute, and our University Librarian is supportive of it.

This has produced a "perfect storm" of interest in the Libraries, in the V.P. Research's office, and in our Sponsored Programs office, concerning data management issues. The Data Curation Profiles workshop was extremely valuable and works great in conjunction with the E-science institute."

Conclusion

Launching data services is a complex endeavor. Librarians seeking to respond to the needs of their constituencies must develop an understanding of the approaches and practices employed by researchers in conducting research, learn researcher's needs for their data, and understand how their skills and perspectives as librarians can be applied to address these needs. Libraries as organizations must create an environment that supports the roles and responsibilities of librarians working with data and ensure that librarians receive the training they need to develop relevant knowledge and skill sets. Library organizations must also find ways to align and partner with other agencies of the university effectively. Most of all there needs to be a clear understanding between librarians and the organizations that they work for regarding intentions, initiatives, and expectations. Success will depend on developing librarians who are prepared to engage their constituencies to discover, understand, and react to the data management and curation issues, and on developing library organizations that are able to provide the time, resources, and support for librarians who are initiating these engagements.

Developing successful librarians and library organizations are not separate paths of consideration, although they are often treated as such. The opportunities and barriers librarians encounter when engaging in data management and curation issues stem from both their own abilities and the environment in which they work. If we are to develop a more complete understanding of how initiatives and programs in data services succeed or fail, we will need to acknowledge the connection between individual and organization and address it in a more holistic fashion.

References

Alvaro, Elsa, Heather Brooks, Monica Ham, Stephanie Poegel, and Sarah Rosencrans. "E-science Librarianship: Field Undefined." *Issues in Science and Technology Librarianship*, 66 (2011): 1-16, http://dx.doi.org/10.5062/F46Q1V55

Antell, Karen, Brian Shults, Jody Bales Foote, and Jaymie Turner. "Dealing with Data: Science Librarians' Participation in Data Management at Association of Research Libraries Institutions." *College & Research Libraries*. Anticipated Publication Date: July, 2014. http://crl.acrl.org/content/early/2013/04/05/crl13-464.full.pdf+html.

Brandt, D. Scott. "Librarians as partners in e-research: Purdue University Libraries promote collaboration." *College & Research Libraries News* 68, no.6 (2007): 365-367, 396. http://crln.acrl.org/content/68/6/365.full.pdf.

Friedlander, Amy. and Adler, Prudence. 2006. "To Stand the Test of Time: Long-term Stewardship of Digital Data Sets in Science and Engineering." *Association of Research Libraries (ARL).* http://www.eric.ed.gov/PDFS/ED528649.pdf

Hey, Tony and Jessie Hey. "e-Science and its implications for the library community." *Library Hi Tech* 24, no. 4 (2006): 515-528, http://dx.doi.org/10.1108/07378830610715383

Gabridge, Tracy. "The Last Mile: Liaison Roles in Curating Science and Engineering Research Data." *Research Library Issues*, no. 265 (2009): 15-21. http://old.arl.org/bm~doc/rli-265-gabridge.pdf

Garritano, Jeremy R, and Jake R Carlson. "A Subject Librarian's Guide to Collaborating on e-Science Projects." *Issues in Science Technology Librarianship* 57, no. 57 (2009): 5, http://dx.doi.org/10.5062/F42B8VZ3

Gold, Anna. "Cyberinfrastructure, Data, and Libraries, Part 2: Libraries and the Data Challenge: Roles and Actions for Libraries." *DLib Magazine* 13, no. 9/10 (2007): 10. http://www.dlib.org/dlib/september07/gold/09gold-pt2.html.

Joint Task Force On Library Support For E-Science. "Agenda for Developing E-Science in Research Libraries." *Association of Research Libraries (ARL).* http://old.arl.org/bm~doc/ARL_EScience_final.pdf

Lyon, Liz. "The Informatics Transform: Re-Engineering Libraries for the Data Decade." *International Journal of Digital Curation* 7, no. 1 (2012): 126-138, http://dx.doi.org/10.2218/ijdc.v7i1.220

Luce, Rick. "A New Value Equation Challenge: The Emergence of eResearch and Roles for Research Libraries." *No brief candle reconceiving research libraries for the 21st century* 142 (2008): 13-20. http://www.clir.org/pubs/reports/pub142/luce.html

Maloney, Krisellen, Kristin Antelman, Kenning Arlitsch, and John Butler. "Future Leaders' Views on Organizational Culture." *College & Research Libraries* 71, no. 4 (2010): 322-347. http://crl.acrl.org/content/71/4/322.short.

Mullins, James L. "Enabling International Access to Scientific Data Sets: Creation of the Distributed Data Curation Center (D2C2)." IATUL Annual Conference Proceedings 17 (2007): 1-36. http://www.lib.kth.se/IATUL2007/assets/files/fulltext/Mullins_J_full.pdf.

Newton, Mark P, C.C. Miller, and Marianne Stowell Bracke. "Librarian Roles in Institutional Repository Data Set Collecting: Outcomes of a Research Library Task Force." *Collection Management* 36, no. 1 (2012): 53-67, http://dx.doi.org/10.1080/01462679.2011.530546

Steinhart, Gail; Chen, Eric; Arguillas, Florio; Dietrich, Dianne; and Kramer, Stefan.

"Prepared to Plan? A Snapshot of Researcher Readiness to Address Data Management Planning Requirements." Journal of eScience Librarianship 1(2) (2012): 63-78. http://dx.doi.org/10.7191/jeslib.2012.1008.

Steinhart, Gail, John Saylor, Paul Albert, Kristine Alpi, Pam Baxter, Eli Brown, Kathy Chiang, Jon Corson-Rikert, Peter Hirtle, Keith Jenkins, Brian Lowe, Janet McCue, David Ruddy, Rick Silterra, Leah Solla, Zoe Stewart-Marshall and Elaine L. Westbrooks. 2008 "Digital Research Data Curation: Overview of Issues, Current Activities, and Opportunities for the Cornell University Library." *e-Commons @ Cornell* http://hdl.handle.net/1813/10903.

Swan, Alma, and Brown, Sheridan. 2008. "Skills, Role & Career Structure of Data Scientists & Curators: Assessment of Current Practice & Future Needs." *Key Perspectives*. http://www.jisc.ac.uk/media/documents/programmes/digitalrepositories/dataskillscareersfinalreport.pdf.

Tenopir, Carol, Robert J. Sandusky, Suzie Allard and Ben Birch. "Academic librarians and research data services: preparation and attitudes." *IFLA Journal* 39, no.1 (2012): 70–78, http://dx.doi.org/10.1177/0340035212473089

Walters, Tyler O. "Data Curation Program Development in U.S. Universities: The Georgia Institute of Technology Example." *The International Journal of Digital Curation* 4, no. 3 (2009): 83-92, http://dx.doi.org/10.2218/ijdc.v4i3.116

Witt, Michael, Jacob Carlson, D Scott Brandt, and Melissa H Cragin. "Constructing Data Curation Profiles." *The International Journal of Digital Curation* 4, no. 3 (2009): 93-103, http://dx.doi.org/10.2218/ijdc.v4i3.117

Note

Parts of this paper were presented by the author at the Data Curation Profiles Symposium that took place at Purdue University on September 24, 2012. A video of the presen-tation can be found at: http://docs.lib.purdue.edu/dcpsymposium/2012/pres/5/.

Disclosure: The author reports no conflicts of interest.