



Academic Librarians and Research Data Services: Preparation and Attitudes

Carol Tenopir

Chancellor's Professor, School of Information Sciences
Director of Research and Director of the
Center for Information and Communication Studies
College of Communication and Information
University of Tennessee
Knoxville, TN, USA
E-mail: ctenopir@utk.edu

Robert J. Sandusky

Assistant University Librarian for Information Technology
Clinical Associate Professor
Richard J. Daley Library
University of Illinois at Chicago
Chicago, IL, USA
E-mail: sandusky@uic.edu

Suzie Allard

Associate Professor & Assistant Director
School of Information Sciences
University of Tennessee
Knoxville, TN, USA
E-mail: sallard@utk.edu

Ben Birch

Graduate Research Assistant
School of Information Sciences
University of Tennessee
Knoxville, TN, USA
E-mail: wbirch@utk.edu

Session:

***116 — The role of libraries in data curation, access and
preservation: an international perspective — Science and
Technology Libraries***

Abstract:

Research funding bodies recognize the importance of infrastructure and services to organize and preserve research data, and academic research libraries have been identified as locations in which to base these research data services (RDS). Research data services include data management planning, digital curation (selection, preservation, maintenance, and archiving), and metadata creation and conversion. We report the results of an empirical investigation into the RDS practices of librarians in US and Canadian academic research libraries, establishing a baseline of the engagement of librarians at this early stage of widespread service development. Specifically, this paper examines the opinions of the surveyed librarians regarding their preparedness to provide RDS (background, skills, and education), their attitudes regarding the importance of RDS for their libraries and institutions, and the factors that contribute to or inhibit librarian engagement in RDS.

Introduction

Many research funding bodies (in the US these include the National Science Foundation, the National Endowment for the Humanities, and the National Institutes of Health) recognize the importance of providing services and infrastructure to organize and preserve research data, and academic research libraries have been identified as locations in which to base these research data services (RDS) (Association of Research Libraries, 2006; National Science Foundation, 2008). The academic research library community is currently working to develop RDS as a new set of strategic services (ARL, 2010).

Research data services are defined here as services that address the full data lifecycle, including the data management plan, digital curation (selection, preservation, maintenance, and archiving), and metadata creation and conversion.

It is important to understand at this early stage the degree to which individual librarians working in academic research libraries actually engage in providing research data services (RDS), and the frequency with which they engage in particular research data services. The results of an empirical investigation into the RDS practices of librarians in US and Canadian academic research libraries establish a baseline of the engagement of librarians in RDS and provide LIS practitioners, administrators, and educators with data to inform strategic or tactical planning in academic research libraries. This paper reports results that address the following research questions:

- RQ1: Do academic librarians have the background, skills, and education to provide library-based research data services (RDS)?
- RQ2: What are librarian attitudes regarding the importance of RDS for their libraries and their institutions?

RQ3: What are the factors that contribute to or inhibit engagement of librarians in RDS?

Related Research

Librarians have discussed their possible roles regarding research data services now and into the future (Council on Library and Information Resources, 2008; ARL 2006; Hey & Hey 2006; Gold 2007.) The focus of these discussions is generally on the library's role in data curation, rather than the preparedness and attitudes of individual librarians.

Libraries were the object of study in ARL's 2009 e-science survey in North America (ARL, 2010). Only half of ARL libraries responded; of those about half (45%) had units to provide support for scientific research data on their campuses. An environmental scan by the Data Working Group at Cornell University Library found that a few university libraries were then involved in curation of research data, including John Hopkins, Purdue University, the University of Washington, and Cornell (Steinhart et al. 2008).

In the UK, a 2007 study found little awareness by librarians of whether research data services were in development at their institutions (Martinez, 2007). Another UK survey found a third of respondents believed that in five years time "manager of datasets from e-science/grid projects" would be a core role of librarians, with another third designating it an ancillary role (RIN / CURL, 2007).

Three key roles for data librarians were proposed by Swan and Brown (2008), including: "increasing data awareness amongst researchers; providing archiving and preservation services within the institution and through institutional repositories; and developing a new professional strand of practice in the form of data librarianship."

Seventy-three percent of the data managers surveyed as part of the PARSE.Insight project in 2009 were employed in libraries (Kuipers & Van der Hoeven, 2009). Among these respondents, the three most highly rated reasons to preserve research data included preservation of publicly funded research, stimulation of the advancement of science, and reanalysis of existing data.

Methods

This study surveyed librarians employed by Association of Research Libraries (ARL) member libraries, whose areas of responsibility seemed to make it likely that they would either be engaged in providing RDS, preparing to become engaged in RDS, or sensitive to the issues around data management, data curation, and / or e-research. "ARL is a nonprofit organization of 126 research libraries at comprehensive, research-extensive institutions in the US and Canada" (<http://www.arl.org/arl/index.shtml>). Most of these (116) are libraries in universities.

The librarians invited to participate in the survey were identified by examining the Web sites of ARL academic libraries, locating staff directories, and compiling contact information for librarians involved in specific roles in those libraries (See Table 1). Based upon the information available from their library’s staff directory, librarians who seemed most likely to be associated with a function or responsibility with a relationship to scientific data curation or data management were selected to be invited to complete the survey. Librarians who had responsibilities for selected disciplines, such as life or physical sciences, were also included in the survey population. Librarians specializing in cataloging, reference, instruction, or special collections, for example, were not included. Contact information could be found on the Web for 111 ARL libraries, and a total of 948 invitations to participate were distributed. A total of 223 librarians responded, for a response rate of 23.5%.

Table 1: Types of librarians invited to participate in the survey

Type	Count
Metadata	141
Digital collections	97
Life sciences	85
Physical sciences	81
Geographic information systems	71
Chemistry	70
Scholarly communications	69
Biomedical / health	68
E-science / ...	66
Electronic resources	62
Institutional repository	46
Data	44
Health / medicine	38
Other	10
Total	948

Table 1: Count of librarians by type invited to participate in the survey.

Limitations

The intent of the sampling process was to select librarians working in areas likely to be involved in an aspect of research data services. Provision of RDS is still an emerging area of responsibility, so we were interested in obtaining responses from, for example, life sciences librarians who were either deeply or uninvolved in providing RDS at the time of the survey. It is possible that the librarians who responded to the survey represent some self-selection bias: the responses may over-represent librarians who are relatively deeply engaged or interested in RDS and under-represent librarians who are uninvolved or uninterested in RDS.

Results

We asked respondents “Do you interact with faculty, students, or staff in support of their research data services (RDS) as part of your regular job responsibilities?” More than two-thirds of the 223 respondents have provision of research data services as an occasional or integral part of their job responsibilities (Figure 1).

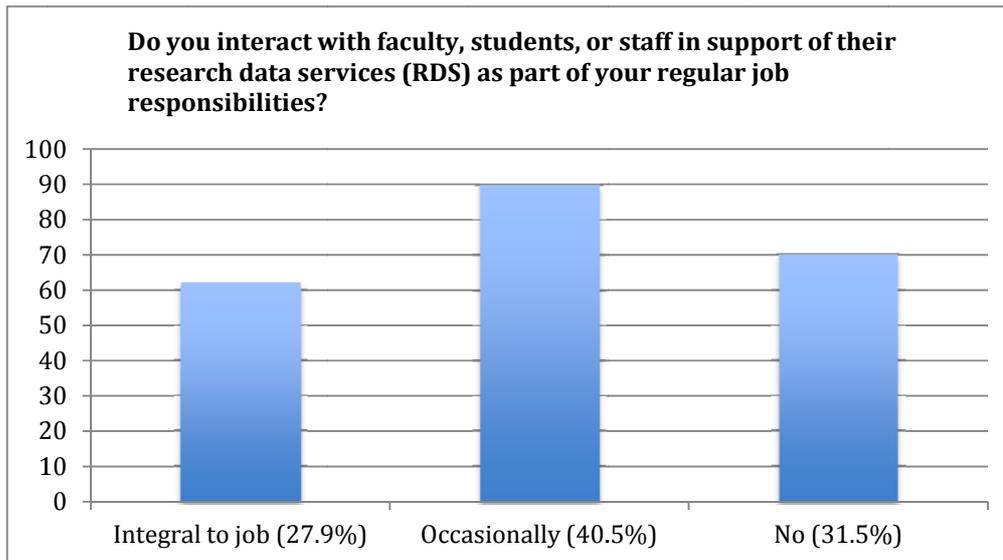


Figure 1: Frequency of responses to “Do you interact with faculty, students, or staff in support of their research data services (RDS) as part of your regular job responsibilities?”

We then used these three groups of respondents, which we label the “integral”, “occasional”, and “no” groups, to cross-tabulate responses to other questions that address our three research questions.

RQ1: Do academic librarians have the background, skills, and education to provide library-based research data services (RDS)?

Respondents were asked to agree or disagree with a set of seven statements related to skills, knowledge, and training related to RDS and their library’s support for their professional development as it relates to RDS. The responses to these seven questions were cross-tabulated with the degree to which RDS are integral to their job responsibilities (Table 2).

Table 2: Librarians' skills, knowledge, and training necessary to provide RDS

As a librarian...	Interaction Frequency	Agree Strongly	Agree Somewhat	Neither Agree Nor Disagree	Disagree Somewhat	Disagree Strongly
(2.1) ... I have the skills, knowledge, and training necessary to provide RDS.	Integral	12 (26.7%)	23 (51.1%)	2 (4.4%)	7 (15.6%)	1 (2.2%)
	Occasional	5 (7.2%)	27 (39.1%)	8 (11.6%)	21 (30.4%)	8 (11.6%)
	No	2 (4.4%)	8 (17.8%)	8 (17.8%)	15 (33.3%)	12 (26.7%)
(2.2) ... I have sufficient subject expertise to provide RDS to my patrons	Integral	18 (40.0%)	13 (28.9%)	6 (13.3%)	7 (15.6%)	1 (2.2%)
	Occasional	15 (21.7%)	24 (34.8%)	9 (13.0%)	15 (21.7%)	6 (8.7%)
	No	5 (11.1%)	16 (35.6%)	10 (22.2%)	11 (24.4%)	3 (6.7%)
(2.3) ... my job allows me sufficient time to provide RDS to my patrons	Integral	12 (26.7%)	16 (35.6%)	6 (13.3%)	7 (15.6%)	4 (8.9%)
	Occasional	2 (2.9%)	16 (23.5%)	14 (20.6%)	23 (33.8%)	13 (19.1%)
	No	1 (2.2%)	3 (6.7%)	15 (33.3%)	14 (31.1%)	12 (26.7%)
(2.4) ... I have access to training in RDS to help me meet my patrons' needs	Integral	8 (17.8%)	21 (46.7%)	5 (11.1%)	9 (20.0%)	2 (4.4%)
	Occasional	2 (2.9%)	15 (22.1%)	20 (29.4%)	21 (30.9%)	10 (14.7%)
	No	0 (0.0%)	4 (8.9%)	12 (26.7%)	13 (28.9%)	16 (35.6%)
(2.5) ... my library provides opportunities to develop skills related to RDS.	Integral	12 (26.7%)	20 (44.4%)	5 (11.1%)	5 (11.1%)	3 (6.7%)
	Occasional	6 (8.7%)	28 (40.6%)	17 (24.6%)	11 (15.9%)	7 (10.1%)
	No	1 (2.2%)	14 (31.1%)	7 (15.6%)	10 (22.2%)	13 (28.9%)
(2.6) ... my library supports me to take courses related to RDS.	Integral	19 (43.2%)	17 (38.6%)	3 (6.8%)	4 (9.1%)	1 (2.3%)
	Occasional	14 (20.3%)	22 (31.9%)	24 (34.8%)	7 (10.1%)	2 (2.9%)
	No	4 (9.3%)	11 (25.6%)	18 (41.9%)	6 (14.0%)	4 (9.3%)
(2.7) ... my library supports me to attend conferences or workshops elsewhere related to RDS.	Integral	24 (53.3%)	15 (33.3%)	2 (4.4%)	4 (8.9%)	0 (0.0%)
	Occasional	24 (34.8%)	29 (42.0%)	10 (14.5%)	6 (8.7%)	0 (.0%)
	No	5 (11.6%)	12 (27.9%)	18 (41.9%)	6 (14.0%)	2 (4.7%)

More than three-quarters of respondents (78%) for whom RDS are an integral part of their job responsibilities somewhat or strongly agreed that they have the necessary skills, knowledge, and training to provide RDS (row 2.1 in Table 2). About 46% of those who have occasional responsibilities for RDS agreed that they have the necessary skills, knowledge and training. For respondents who don't have RDS as part of their job responsibilities 60% feel they do not have the skill, knowledge and training necessary to provide RDS. This pronounced pattern of high agreement from the "integral" group, moderate agreement from the "occasional" group, and low agreement from the "no" group shown here is typical of the responses to most of these seven statements.

The responses to the second statement (Table 2, row 2.2) about librarians' subject expertise were more evenly distributed: 69% of the "integral" group strongly or somewhat agreed that they had sufficient subject expertise; about 57% of the "occasional" group and 47% of the "no" group somewhat or strongly agreed. At this

early stage this might be considered a position of strength for the future of library involvement with RDS—almost half to two-thirds of respondents feel they have the subject expertise necessary to provide these services to their patrons.

Responses to the statement that their jobs allow them sufficient time to provide RDS to their patrons (Table 2, row 2.3) shows a pronounced difference in the level of agreement / disagreement between the three groups. For the “integral” group, about 62% somewhat or strongly agree that their job allows sufficient time to provide RDS. Only a quarter (26%) of the “occasional” group somewhat or strongly agrees, and only 9% of the “no” group agrees with the statement. If RDS services are to be expanded at ARL libraries, RDS need to be made a priority in the responsibilities of the librarians who will be providing these services. This will require a reassessment of priority of all library services and a reallocation of librarian responsibilities.

When asked to agree or disagree with whether they have access to training in RDS to help them meet their patrons’ needs (Table 2, row 2.4), respondents for whom RDS are integral are much more likely to agree (about 65%) compared to the other two groups. Access to training seems to track with current responsibilities.

Similarly, the responses to the statement that their library provides opportunities to develop skills related to RDS (Table 2, row 2.5) shows the same pattern of higher agreement from people in the “integral” group (71% agree somewhat or strongly). We cannot tell from the answers if the librarians in the occasional or no groups work at libraries that do not provide opportunities to develop RDS skills or if these librarians are simply unaware of opportunities that exist. However, even if not provided in their library, a majority of librarians agree that they are provided with support to take courses to develop skills related to RDS (Table 2, row 2.6). Again, respondents in the integral group are much more likely to agree.

Most respondents in the “integral” and “occasional” groups agree with the statement “my library supports me to attend conferences or workshops elsewhere related to RDS,” (Table 2, row 2.7), When the three groups are combined, about 70% agree strongly or somewhat with the statement.

RQ2: What are librarian attitudes regarding the importance of RDS for their libraries and their institutions?

Respondents were asked to agree or disagree with a set of six statements related to the importance of RDS for libraries and the institutions within which libraries are a part. The responses to these six questions were cross-tabulated with the degree to which RDS are integral to their job responsibilities (Table 3).

Table 3: Librarians' attitudes regarding the importance of RDS for libraries and institutions

	Interaction Frequency	Agree Strongly	Agree Somewhat	Neither Agree Nor Disagree	Disagree Somewhat	Disagree Strongly
(3.1) RDS are just as important as other activities that I provide for my patrons.	Integral	26 (57.8%)	11 (24.4%)	4 (8.9%)	4 (8.9%)	0 (0.0%)
	Occasional	19 (27.5%)	28 (40.6%)	11 (15.9%)	8 (11.6%)	3 (4.3%)
	No	5 (11.9%)	10 (23.8%)	15 (35.7%)	8 (19.0%)	4 (9.5%)
(3.2) RDS are unnecessary for librarians to provide to their patrons	Integral	1 (2.3%)	0 (0.0%)	4 (9.3%)	10 (23.3%)	28 (65.1%)
	Occasional	0 (0.0%)	2 (2.9%)	6 (8.8%)	25 (36.8%)	35 (51.5%)
	No	0 (0.0%)	2 (4.7%)	5 (11.6%)	24 (55.8%)	12 (27.9%)
(3.3) RDS are a priority at my library	Integral	15 (33.3%)	15 (33.3%)	6 (13.3%)	6 (13.3%)	3 (6.7%)
	Occasional	7 (10.4%)	20 (29.9%)	24 (35.8%)	14 (20.9%)	2 (3.0%)
	No	1 (2.3%)	7 (16.3%)	14 (32.6%)	10 (23.3%)	11 (25.6%)
(3.4) Providing RDS will increase the visibility and impact of our institutional research	Integral	29 (69.0%)	10 (23.8%)	3 (7.1%)	0 (0.0%)	0 (0.0%)
	Occasional	29 (42.6%)	34 (50.0%)	5 (7.4%)	0 (0.0%)	0 (0.0%)
	No	14 (32.6%)	19 (44.2%)	9 (20.9%)	1 (2.3%)	0 (0.0%)
(3.5) RDS are a distraction from the library's core mission.	Integral	1 (2.4%)	0 (0.0%)	2 (4.8%)	9 (21.4%)	30 (71.4%)
	Occasional	1 (1.5%)	0 (0.0%)	5 (7.4%)	24 (35.3%)	38 (55.9%)
	No	3 (7.0%)	1 (2.3%)	8 (18.6%)	17 (39.5%)	14 (32.6%)
(3.6) The library is the best-suited entity at my institution to provide RDS	Integral	17 (39.5%)	19 (44.2%)	6 (14.0%)	0 (0.0%)	1 (2.3%)
	Occasional	20 (29.4%)	15 (22.1%)	21 (30.9%)	10 (14.7%)	2 (2.9%)
	No	7 (16.3%)	17 (39.5%)	13 (30.2%)	4 (9.3%)	2 (4.7%)

When asked to agree or disagree with whether RDS are just as important as other activities that they provide for their patrons (Table 3, row 3.1), respondents in the “integral” group are more likely to agree (about 82%) than those in the “occasional” (68%) and “no” groups (36%). When the three groups are combined, about two-thirds (63%) agree strongly or somewhat.

Looking at the issue from another and more general perspective, we asked respondents to agree or disagree with the statement “RDS are *unnecessary* for librarians to provide to their patrons” (Table 3, row 3.2). A vast majority of respondents in all three groups disagreed with this statement, indicating by implication that they feel RDS are *necessary* services.

Asked to agree or disagree with the statement “RDS are a priority at my library” (Table 3, row 3.3), the responses form the pattern of much higher agreement from the “integral” group than the “occasional” group and the lowest level of agreement from the “no” group. This represents the opinion of these individual librarians and does not necessarily reflect the official priorities of their institutions. A second survey of academic library policies was sent to directors of academic libraries that

are members of the Association of College and Research Libraries (ACRL) in 2011-2012. These results are forthcoming.

There is overwhelming agreement among librarians in all three groups that “providing RDS will increase the visibility and impact of our institutional research” (Table 3, row 3.4). Clearly, these librarians agree that RDS have value to the research mission of their institutions. Looking at it from the perspective of the library, respondents in the “integral” and “occasional” groups overwhelmingly disagree with the statement “RDS are a distraction from the library’s core mission” (Table 3, row 3.5). Thus, by implication, RDS are considered consistent with the core mission of an academic research library.

Asked to agree or disagree with the statement “the library is the best-suited entity at my institution to provide RDS” (Table 3, row 3.6), more respondents from the “integral” group agree than those in the “occasional” group (about 84% to 52%). Surprisingly, 56% of respondents in the “no” group agree with this statement and the responses from the “occasional” group show the highest level of disagreement. When the three groups are combined, about 62% agree somewhat or strongly that the library is the best-suited entity to provide RDS. It would be interesting to identify, from the perspective of librarians employed by research libraries, which other entities on campuses might be better suited to offer RDS.

RQ3: What are the factors that contribute to or inhibit engagement of librarians in RDS?

Respondents were asked to identify what motivates their involvement in the provision of library-based RDS. We asked those already involved in providing RDS to identify the single most important motivation for their involvement and to also identify other factors that motivate their participation. We asked those librarians who are not yet involved in providing RDS to identify the factors that would most motivate them to become involved.

When librarians who are already involved in providing RDS, either as an integral or occasional part of their job responsibilities, were asked what is *the single most important motivation* for their involvement, they indicated a range of professional responsibilities or professional interest (Figure 2).

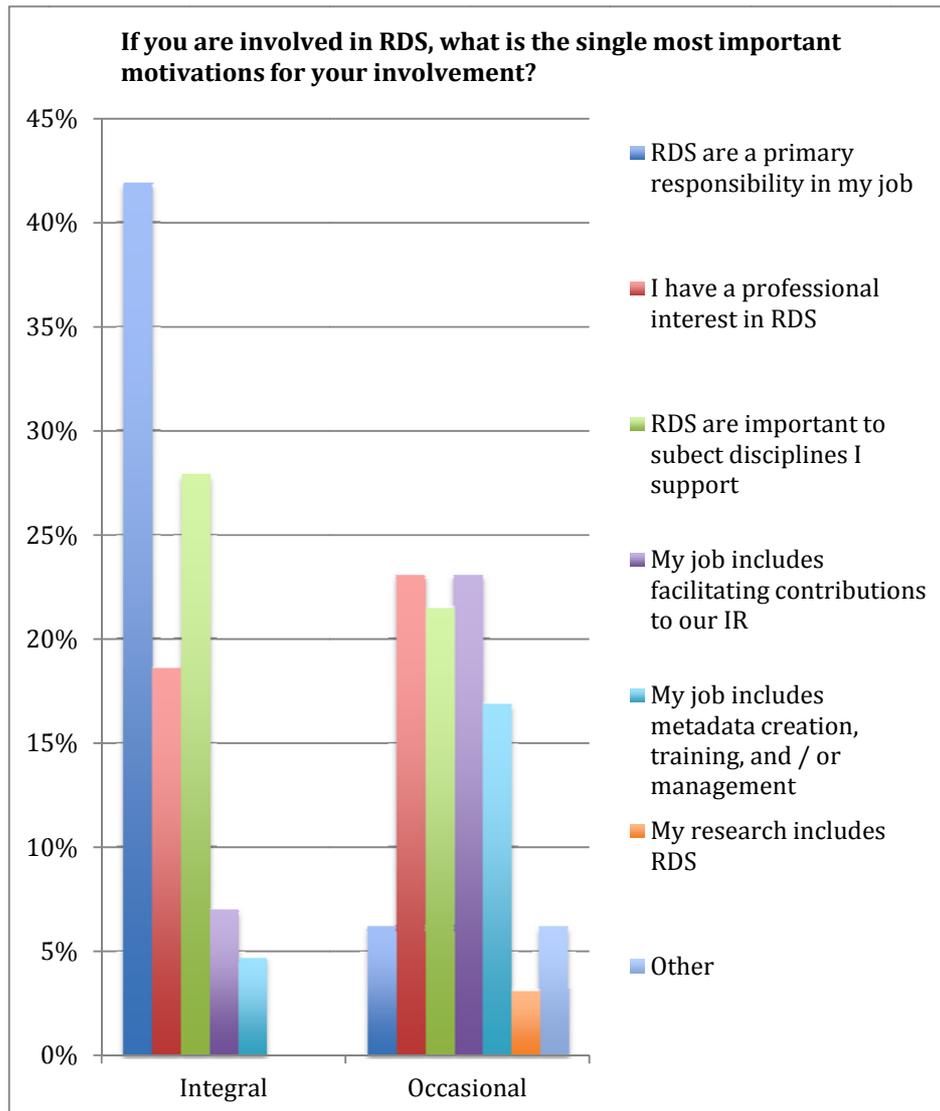


Figure 2: If you are currently involved in RDS, what is the single most important motivation for your involvement?

Librarians who are already involved in providing RDS, either as an integral or occasional part of their job responsibilities, were also asked to identify other motivations for their involvement (Figure 3.) Respondents were allowed to select more than one response to this question. Again, a range of professional responsibilities motivate involvement with RDS, with professional interest the most frequently selected answer in both groups, followed by the importance of RDS to the subject disciplines they support.

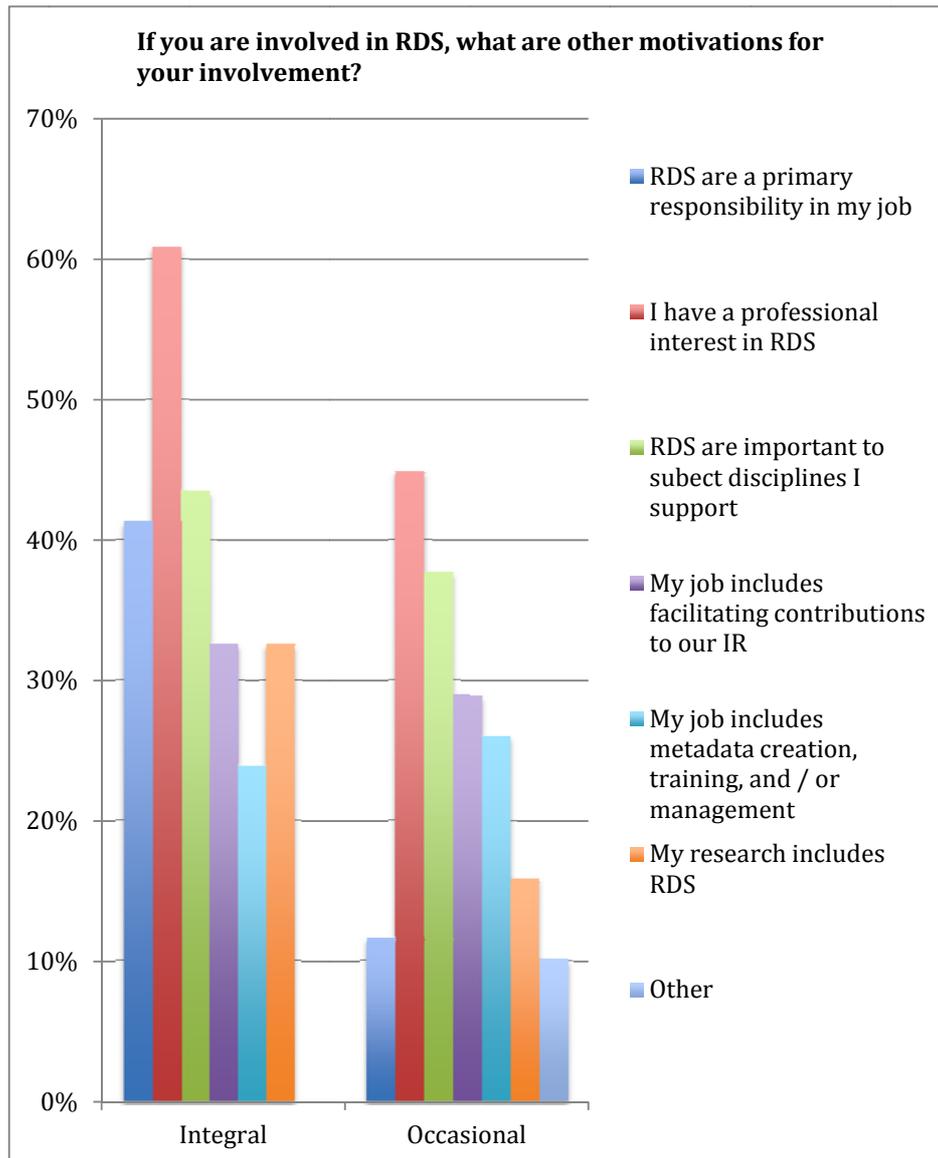


Figure 3: If you are currently involved in RDS, what are other motivations for your involvement? Respondents were able to select more than one response.

When librarians who are not involved in providing RDS were asked to identify what would most motivate them to participate, being asked to do so by their patrons was the primary motivation. Increased institutional involvement in RDS, addition of RDS job responsibilities, and development of an institutional repository for data were all mentioned by fifty percent or more of the librarians in this group (Figure 4).

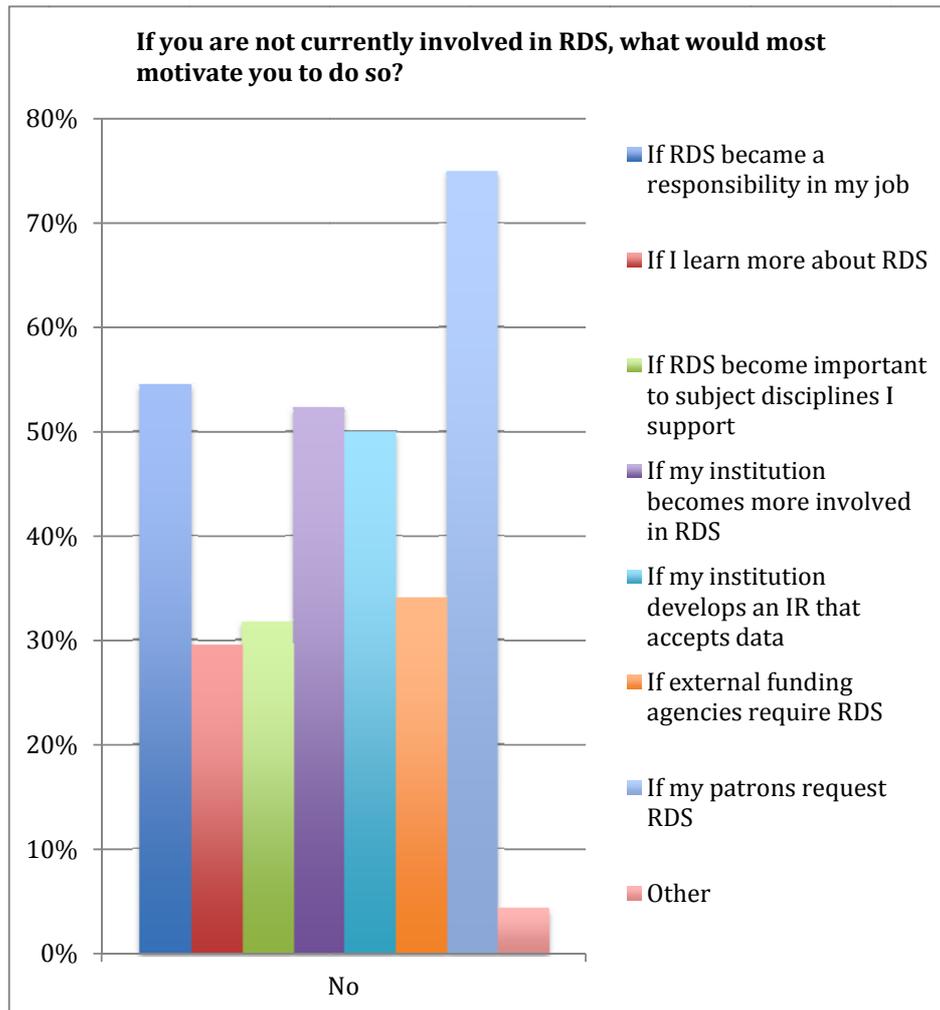


Figure 4: If you are not currently involved in RDS, what would most motivate you to do so? Respondents were able to select more than one response.

Conclusions

Nearly three-quarters of the ARL librarians who responded to this survey do not have research data services as an integral part of their job responsibilities at this time, yet we found evidence that many ARL librarians believe they have the knowledge and skills and opportunities to provide RDS in the future and are motivated by professional interests, patron demand, and job responsibilities. Their attitudes show they believe that RDS are important services for academic research libraries to provide and RDS are consistent with the library mission and role.

These librarians believe that research data services will increase the visibility and impact of institutional research. An implication is that library-based RDS are important opportunities for increased alignment between library services and the university research mission.

Libraries are now at an early point in a transition from collection-based services to RDS, requiring resetting of priorities, realignment of responsibilities, and provision of opportunities for librarians to develop skills related to RDS.

Further Research

This survey was conducted as part of the NSF-funded DataONE (Data Observation Network for Earth) project. It is just one in a series of baseline assessments of DataONE stakeholders. The baseline assessment of scientists was completed in 2011 (Tenopir et al, 2011). Baseline assessments of US and Canadian academic library policies, US federal librarians and library policies, data managers, and environmental science college teachers were conducted in 2011-2012 and will be published soon. Future baseline assessments to be conducted in 2012-2013 will include assessments of institutional policy makers, publishers, and post-graduate and undergraduate students. Follow-up surveys of all these groups are planned for the future.

Acknowledgements

Betsy Gunia and Christina Murray contributed to this project in its early stages, as part of a DataONE summer internship. Members of the DataONE Usability & Assessment Working Group helped revise and refine the survey instrument. DataONE is funded by the US National Science Foundation Division of Cyberinfrastructure, William Michener, P.I.

References

- Association of Research Libraries (2006). To stand the test of time: Long-term stewardship of digital data sets in science and engineering. Washington, DC: Association of Research Libraries. Retrieved May 4, 2012 from <http://www.arl.org/bm~doc/digdatarpt.pdf>.
- Association of Research Libraries (2010). E-Science and data support services: A study of ARL member institutions. Washington, DC: Association of Research Libraries. Retrieved May 4, 2012, from http://www.arl.org/bm~doc/escience_report2010.pdf.
- Council on Library and Information Resources (2008). No brief candle: Reconceiving research libraries for the 21st century. Washington DC: Council on Library and Information Resources. Retrieved May 4, 2012 from <http://www.clir.org/pubs/reports/pub142/pub142.pdf>.
- Gold, A. (2007). Cyberinfrastructure, data, and libraries, part 2: Libraries and the data challenge: Roles and actions for libraries. D-Lib Magazine 13, (9/10). Retrieved June 21, 2010 from <http://www.dlib.org/dlib/september07/gold/09gold-pt2.html>.
- Hey, T. & Hey, J. (2006). e-science and its implications for the library community. Library Hi Tech, Vol. 24 Iss: 4, pp.515 – 528.

- Kuipers, T. & Van der Hoeven, J. (2009). Survey report (D3.4). Didcot, UK: PARSE.Insight. Retrieved June 23, 2010 from <http://www.parse-insight.eu/publications.php#d3-4>.
- Martinez, L. (2007). The e-research needs analysis survey report. London: CURL/SCONUL Joint Task Force on e-Research. Retrieved June 23, 2010 from <http://www.rluk.ac.uk/files/E-ResearchNeedsAnalysisRevised.pdf>.
- National Science Foundation, Office of Cyberinfrastructure Directorate for Computer & Information Science & Engineering (2008) Sustainable digital data preservation and access network partners (DataNet) program solicitation - NSF 07-601. Retrieved September 22, 2010 from http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503141.
- RIN / CURL. (2007). Researchers' use of academic libraries and their services: A report commissioned by the Research Information Network and the Consortium of Research Libraries. London: Research Information Network and Consortium of Research Libraries in the British Isles. Retrieved June 23, 2010 from <http://www.rin.ac.uk/our-work/using-and-accessing-information-resources/researchers-use-academic-libraries-and-their-serv>
- Steinhart, G. et al. (2008). Digital research data curation: Overview of issues, current activities, and opportunities for the Cornell University Library. Retrieved June 14, 2010 from <http://hdl.handle.net/1813/10903>
- Swan, A. & Brown, S. (2008). The skills, role and career structure of data scientists and curators: An assessment of current practice and future needs. Truro, UK: Key Perspectives Ltd. Retrieved June 23, 2010 from <http://eprints.ecs.soton.ac.uk/16675/>.
- Tenopir C, Allard S, Douglass K, Aydinoglu AU, Wu L, et al. (2011) Data sharing by scientists: Practices and perceptions. PLoS ONE 6(6): e21101. doi:10.1371/journal.pone.0021101