SURVEY RESULTS

EXECUTIVE SUMMARY

Introduction

With an increased emphasis on open data, data management planning requirements, and potential "big data" research opportunities, research institutions are recognizing an emerging demand to provide a wider and more refined array of data services to meet needs at different points in the research process. Many research libraries are answering that demand by expanding or adopting new research data services, most within the last three years. The timing was appropriate, therefore, to survey ARL member libraries, assessing early endeavors, and benchmarking future growth as we anticipate demands for these services to expand and for libraries to find new opportunities for relevant services.

In this context, we identified two emerging areas of services that are relatively new for member libraries. These are research data management, which enters the research process at the grant proposal stage in meeting data management planning requirements, and in various ways during the research process. Second is support for data archiving, at a project's end for preservation and online dissemination to facilitate data sharing, and in providing new data resources for discovery. This survey will use the term "research data management services" (RDMS) to refer collectively to library activities surrounding data management and archiving.

Our survey also addresses contextual requirements for planning, developing, staffing, and managing new research data management services. Finding necessary expertise and funding for new positions is challenging, but creative new models of service provision are emerging. Collaboration across institutional units is one route toward unearthing expertise and knowledge to help researchers at all stages of the research process and to provide data management support. Our survey findings, however, point to the "growing pains" of new service development, with challenges such as initiating and encouraging campus-wide coordination that addresses gaps and overlapping services. Responses throughout the survey show that libraries are still in the early stages of development and implementation of RDMS. In most cases, services are evolving ahead of evidence of which models and strategies will prove most effective or successful. Variables for that development include the structured and unstructured institutional environment for new research services: Are more universities recognizing a community need formally through data policies (See Q1)? Are administration and researchers looking to libraries for solutions or are libraries taking their own initiative (Q6)? At what stage are libraries in conducting needs analysis to guide service development, engaging in active outreach to communicate their provision of new services (Q50), or assessing their early efforts (Q51)?

This SPEC survey of research data management services at ARL libraries joins a growing literature of surveys and case studies covering various dimensions of this emerging domain.[1] To a degree, we are still learning what questions we should ask to assess current practices and provide benchmarks for assessing future developments. We address a breadth of aspects that has not been consolidated before to encourage further research, but perhaps more importantly, to give libraries a timely orientation to the challenges and benefits of offering research data management services (henceforth referred to as RDM services).

Background

This survey builds upon and extends the findings of the 2010 report entitled "E-Science and Data Support Services: A Study of ARL Member Institutions," authored by Catherine Soehner, Catherine Steeves, and Jennifer Ward, and sponsored by the ARL E-Science Working Group.[2] In preparing for our survey, we spoke with the authors of the report to discuss which aspects of the study went well and which areas would have benefited from continued investigation. Among the strengths, the authors specifically noted valuable responses to inquiries about how organizations began their efforts, collaborative approaches, educational efforts, pressure points or barriers to entry, and refocusing of professional interests. As key areas to explore further in our survey, the authors identified semantic and contextual clarity, more detail of academic and work backgrounds, ties between data management and cyberinfrastructure, perceptions of responsibility, areas of intended investment, and actions which defined success. While we did not have enough space to address all new questions, we have attempted to incorporate many of those points into our survey questions. We recognize that the absence of new, comprehensive case studies is a limitation of our study. This may be a fruitful direction for a next round of inquiry.

One of the areas that we recognized as important in building upon the 2010 report was further refinement of terminology, which may have varying interpretations or ambiguity among respondents and more broadly in this field of service. Examples of such terms include "eScience," "cyberinfrastructure," "data archiving," and distinguishing research data management from data services more broadly, the latter sometimes including institutional records management. Perhaps the most ambiguous term still emanating through these conversations is "digital services." As the survey results show, in application this can range from digitization to research support to intellectual property to repository management. While we attempted to specify particular meanings and control response variables for the purpose of a more systematic analysis, this process and some of the responses we received highlight the fact that there is still great variation in maturity and definition of services throughout the ARL community. Based on anecdotes, we believe that there is also even broader variation and understanding of the meaning of these terms and services beyond the ARL community, such as with the consumers of the services. Among other goals, we hope that this type of study may shed some light on the variations of meaning, and provide some opportunity for further maturation and convergence of terminology.

Broad Data Support Services

Seventy-three of the 125 ARL member libraries responded to the survey. All are academic libraries. Respondents were asked whether and how long they have offered 11 broad types of data support services (Q3). All 73 libraries offer at least one of the listed services. Helping researchers locate and use data sources is the most common and long-established service (68 responses, or 93%). Also common and long-offered are support for geospatial analysis (61, or 84%), dataset acquisition (58, or 79%), and copyright and patent advising (53, or 74%). Sixty-four respondents (88%) provide an institutional repository; 48 for more than three years. About half of the respondents plan to add one or more services, particularly data visualization, data mining, and data analysis.

Research Data Management Services

Respondents were also asked whether they offer research data management services, defined as "providing information, consulting, training or active involvement in: data management planning, data management guidance during research (e.g., advice on data storage or file security), research documentation and metadata, research data sharing and curation (selection, preservation, archiving, citation) of completed projects and published data" (Q4). Almost three-quarters do (54, or 74%). Seventeen others plan to (23%). Only two have no plans to offer RDM services. One of those two commented, "Rather, no, we wish we could, but we have no formal plan to do so in the short term because of lack of resources." The other explained that there is little to no demand for such services.

Origin of Research Data Management Services

Several survey questions addressed the current and transitioning climate of support for data services at each library and their institution. The responses suggest that libraries are developing data support services ahead of formal policy requirements of their institutions. Fewer than a quarter of the 73 respondents (16, or 22%) reported that their institutions have some form of policy for research data management or retention (Q1). However, those policies vary from IRB guidelines for sensitive data to institutional records policies, with few specifying that research data be kept and managed. A third of the respondents (24, or 33%) reported that policies are planned in the next one to three years, suggesting a trend of institutions to keep up with expanding funder requirements.

Only four libraries initiated RDM services before 2005 (Q5). The earliest reported was the 1966 library-supported Latin American Data Bank project. Ten others started their RDM services between 2005 and 2009. This correlates to the early initiatives for eScience, which was the hot topic of many papers and task forces.[3] For these early providers the most important reasons for beginning service were researchers' requests for help and a library initiative to expand support of faculty research (Q6). One reports their AUL "at the time was a visionary in terms of DRM and initiated a broad range of services."

Five of the 11 libraries that started their services in 2010 also reported the influence of library initiatives. Four others stated the main reason was the National Science Foundation announcement that they would begin requiring data management plans on January 18, 2011. The NSF requirement was the main reason for 11 of the 16 libraries that started RDM services in 2011, as well.

In 2011 and 2012, ARL and CLIR/Digital Library Federation co-sponsored the E-Science Institute, a workshop series to help libraries develop e-research strategic agendas. The institute is now operated by DuraSpace and is open to non-ARL institutions.[4] Forty-nine survey respondents (67%) have attended one or both of the previous sessions; seven (10%) say they plan to attend a future session, four for the first time (Q2). Of the 49 who have attended the institute, 40 provide some level of RDM service (82%). The four who are planning to attend a future institute for the first time also already provide RDM services.

The core of the survey focused on the RDM services that support the management and curation of research data throughout its life cycle. The following table presents the range of RDM service categories discussed below, with their corresponding survey questions.

RDM Services (N=54)	N	%	Q
Online Data Management Plan (DMP) resources	47	87%	7
DMP training	33	61%	10
DMP consulting	48	89%	11
RDMS besides DMP support	53	98%	18
Data archiving by library	40	74%	19
Data-specific archive (other than institutional repositories)	5	9%	21

Data Management Plans

Many libraries began their RDM service to help researchers create data management plans (DMP), most often for NSF proposals. Two and a half years since the NSF's DMP requirement began, and with several other funders adding similar requirements, the survey asked several questions to gauge the extent to which the libraries with RDM services currently provide online DMP resources, or training and consulting on plan preparation. As seen in the table above, 47 libraries provide online resources related to data management plans. All of those libraries include an explanation of funding agencies' DMP requirements and guidelines for creating DMPs. All but a few have a tool or resource for DMP creation and DMP template examples. In addition to DMP planning information, other online resources include information about digital repository services, long-term data management and preservation, links to related campus services, information on copyright, and workshops offered.

Most respondents are providing links to external resources, but 70% have also created their own, and almost half have customized others' content. Forty-one libraries (75%) have linked to the DMPTool, a prominent online self-service resource for researchers

operated by the California Digital Library and a consortium of contributing institutions, and 29 of those offer training and support for the tool (Q8 & 9).

A rough indicator of the breadth of RDM support at the responding libraries is whether they have gone beyond online resources to add training on DMP preparation and/or consultation on DMPs for grant proposals. Forty-eight libraries offer consultation services. Ten began in 2010, 23 in 2011 (again due to the NSF DMP requirement), 12 in 2012, and 2 more in 2013. Thirty-three of these 48 also provide DMP training. At 25 institutions DMP support is also provided by various other departments, most frequently the office of research and office of sponsored programs (Q10–Q12).

Most direct consultations (other than workshops) are done via e-mail/chat/phone (94%) or meeting with researchers at their office, lab, or other location (92%) (Q13). Of all consultation methods used, many respondents commented that workshops, training, and tutorials have provided the most contacts. One commented on a unique way they get consultations: "In addition to providing our contact information to faculty through departmental grant administrators and our partners in the institution, faculty can also indicate that they would like a consultation through our internal grant tracking system."

While all 48 libraries that offer DMP consultation services interact directly with researchers, it was challenging to gauge the extent of engagement, and how many researchers are actually looking to their library for DMP assistance. When asked about the number and depth of consultations, only 28 libraries reported that they kept track of the number of consultations (Q14). About half of that group reports having more than 10 consultations since their DMP service began. Seven of the ten libraries that started in 2012 and 2013 have consulted on fewer than five plans total. Only two libraries average more than three consultations per month. It is possible that there was confusion of what we meant by "DMP sessions" in our survey question, which had aimed to measure individual one-on-one consultations via any means, e-mail, face-to-face, etc. One response to this question was 300 DMP sessions. This high number may have included number of participants in group sessions; the next highest number was 96 sessions.

When asked which departments/fields of research at their institutions use the DMP services, the majority reported that natural sciences, social sciences, engineering, and humanities use them at least occasionally (Q17). Overall, most of the libraries that are offering more than online DMP resources are receiving modest participation from their researcher communities. If more funders add DMP requirements for proposals, and/or increase compliance and accountability for the quality and follow-through on plans in the next few years, use of these services should be resurveyed, ideally with more detailed usage metrics and qualitative details on models of service provision.

RDM Services Other than DMP Preparation

Helping researchers prepare data management plans for grant proposals is a relatively focused category of service, and for many of the libraries surveyed, an entry point into this relatively new area of support for the research process. The survey next explored whether libraries offer services beyond DMP assistance, and asked about nine categories of additional RDM services that we expected some portion of ARL libraries to offer (Q18). A majority of the 54 responding libraries (36 to 48) offer eight of the nine services, including data management best practices (both online resources and workshops), helping researchers identify (and apply) appropriate metadata standards, research file organization and naming, data citation, data sharing and access, and data storage and backup. The last two services are commonly offered by both the library and elsewhere on campus. Only 14 libraries provide help with securing & anonymizing data. Half said this is only offered elsewhere, such as by the institutional review board, privacy office, or research compliance office. In addition to the listed services, a few libraries report they support data publication, data rights management, and analog to digital image data conversion.

Data Archiving Services

Even with moderately enforced requirements to share datasets from grant projects, publications, and other research by public funders such as NSF, most of the

responding libraries have made initial forays into data archiving services. As funders expand data sharing requirements and efforts at compliance,[5] libraries may see an increasing role and call from researchers to assist in finding data archiving and dissemination solutions.

The survey defined data archiving as providing "longer-term retention of and access to research data by others." A data archive is distinguished from other digital file storage systems by features that include online file access, file integrity checks, and permanent identifiers for locating files (Q19). Forty of the 54 respondents who provide RDM services (74%) report their library offers data archiving. At 13 of these institutions archiving is also offered elsewhere. Five institutions (9%) only offer data archiving elsewhere than the library. The other archiving units are most often central IT and research centers; however, in some cases such services might have been considered centralized storage rather than archiving by our definition.

Whether or not libraries offer their own archiving solution, nearly all the libraries with RDM services offer assistance locating data archiving solutions for research data, for example, at data repositories for a specific domain of research such as neuroscience (Q20). Twenty-six libraries (48%) help researchers deposit data at such repositories, and 21 (39%) have assisted with depositing data at journals, which are increasingly asking for supporting datasets. All but one of the 40 libraries that have their own archives offer direct assistance with depositing data in that archive.

Data Archive Characteristics

From prior literature and informal observations, we expected current archiving solutions among libraries to fall into two categories: repositories designed specifically for research data, and those repositories built for other purposes that can contain datasets. For the latter, we expected institutional repositories (IRs) used primarily for publications to be most common, as well as datasets archived in digital repositories that are used for institutional special collections such as historical photos. We expected data-specific archives built specifically for retaining and accessing research data to be less prevalent because of the higher resource requirements for setting up a system, the lack of ready-made data repository platforms (in contrast to the more mature enterprise of institutional repository software), and the relative lack of demand for libraries to initiate such platforms at this time.

Question 21 attempted to distinguish these two categories, but after comparing Q21 responses to those for other questions, and to links respondents supplied to their archive websites, we found varying interpretations of what respondents reported as a "research data archive dedicated to data deposits and access." All but 5 of the 13 "data archives" appear to be institutional or digital repositories by our definition. The survey questions and responses section lists the original responses, but the table below redistributes the 40 data archiving libraries' primary method into three categories: institutional repositories (IRs) that include datasets in addition to publications; digital repositories used for a variety of institutional collections besides data, such as digital photos, rather than faculty publications; and archiving solutions more clearly dedicated to research datasets.

Archive Type (N=40)	N	%
IR with datasets	30	75%
Digital repository with datasets	5	13%
Data-specific repository	5	13%

Since 88% of the libraries that archive research data use either publication- or digital collection-centered repositories, the analysis of the survey's data archiving questions distinguishes between responses of the five libraries we labeled as having "data archives" from those with institutional repositories or digital collections, which we combined into a single category "IRs with data." This categorization may not be exact in all cases, so we encourage those doing further studies on library data archiving to verify directly with the libraries their current methods.

Funding Data Archives

As is the case for funding RDM services overall (Q49), most of the libraries with data archiving services (31 of 37, or 84%) are absorbing those costs through their internal budgets (Q24). Nine fund archiving through

grants (24%), five charge researchers (14%), and seven have found funds through other means (19%). No library charges users for data access.

There are clear differences between IRs with data and data-specific archives. Thirty of the IRs (94%) absorb any extra costs for research data into the repository budget. Only one of the data archives reported funding from their general budget. Two of the data archives are grant funded, and three charge researchers for archiving. Charging researchers or fees upon grants are much less common for IRs with data. Understanding the costs of archiving in many cases is still under review, and institutions recognize that archiving costs will need to factor in the volume of data and length of hold.

Data Archive Infrastructure

The survey asked respondents to describe the platform and software used for their data archiving solution (Q22). Most of the 38 respondents use open source software for all or part of their solutions; one developed their own software. DSpace is the most commonly used institutional repository and digital collection platform and interface (17, or 43%). Fedora is the platform for eight of the IRs (20%), often along with additional software interfaces such as Hydra or iRODS. The five data archives use Fedora and Data Conservancy software, Chronopolis, a customization of HubZero, a multi-component system that includes Fedora, Archivematica, Dataverse,[6] and iRODS, and a custom-built repository. All are in active development and/or in "beta" phase of implementation.

To assess the use of repositories for data, the survey asked for estimates of the number of researchers currently depositing datasets in the archives (Q23), the typical sources of archived data (Q25), and total deposit size (Q26). Twenty-eight of those with IRs reported that zero to 1000+ researchers have deposited data (a median of 10 and an average of about 91 researchers). Four of those with data archives reported that between two and 100 researchers have deposited data. Twenty-two of the respondents (66%) reported that data deposits are in the gigabyte range; all but three are under 100GB. Eleven others reported total deposits between 1 and 75 terabytes. Follow-up with respondents might yield more precise numbers and distinguish among archives with single large and many small deposits. Clearly, however, these are early days for both data-specific archives and IRs with data, and possibly also for researcher's awareness and adoption of these archiving options.

Data files in both IRs and data archives are coming from a range of sources. Most of the respondents report that datasets are associated with particular publications (88%), full research projects (85%), and graduate theses/dissertations (80%). Twenty-five (63%) report that data was moved from another archive to the library. As data-specific archives expand in use, there may be shifts in data sources that institutional repositories cannot accommodate as well.

The survey also asked about data deposit options. Institutional repositories are generally set up for self-deposit by researchers, and 23 of the IRs with data (65%) do allow data deposits without direct assistance (Q28). However, all but one of these also provides assistance, and 19 say they will deposit data collections for their researchers. Three data archives allow researchers to self-deposit, and they also provide assistance and will deposit data for the researcher. A trend to follow is whether data archive software and support models become more "self-service" for researchers or remain a staff-mediated service.

The final set of data archiving questions addressed details on their architecture for access and preservation. The survey responses show that open access is the policy and intention for all but three of the libraries with archiving solutions, as one would assume based on the literature and public funder requirements. Six of the IRs and data archives also allow controlled access, such as administrative or researcher approval to access data. For data archives in particular, the type of access may be a technical issue, not just policy. Datasets for two institutions are essentially "dark archives" for preservation without a public interface as a direct component of the system, and at least one archive does not currently have the capacity to control access.

Another feature generally considered essential to data archives is support for persistent identifiers so that datasets can be located long-term and reliably cited in publications (including, in some cases, citing particular versions of collections updated with new

data). All but one of the libraries with archives use persistent identifiers. Most common for IRs are the Handle System (21, or 64%). Most of the data archives use DOIs for datasets. Only eight archives use ARKs. Several generate their own identifiers.

Finally, the survey asked about preservation capacities of archives, choosing standards most typical for digital repositories. All five data archives provide file integrity/fixity checking and multiple copy replication, as do 27 IRs (84%). Four data archives practice geographic separation of backups, but this is less common for IRs (22, or 69%). Three data archives and half of the IRs also provide format migration and conversion for data files over time.

Overall responses about archive architecture reflect the different purposes of publication-oriented IR platforms and archives that focus on the specific needs of large and diverse research data collections. If data archiving platforms become less resource intensive for libraries to install and operate, future surveys could gauge corresponding attitudes toward making the library a center for data archiving. Here, the traditional cultural emphasis of libraries and archives on long-term preservation and curation may support libraries' justification for taking on their operation. By contrast, science domains and academic publishers may emphasize shorter-term requirements of data dissemination, yet may hesitate to invest in archiving infrastructure. As interest grows in meeting public funder requirements, and as research practices shift toward data sharing for accelerated discovery and collaboration, academic institutions may recognize libraries as facilitators for research data. Survey responses suggest that infrastructure requirements are significant, and implementation and adoption may be slow. Data archiving by academic libraries, however, is clearly an emerging field that future surveys and case studies should follow.

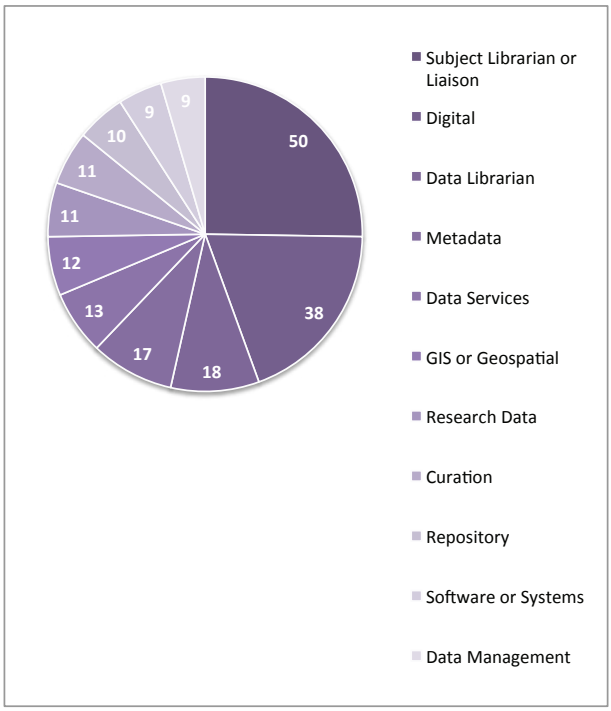
It will be increasingly relevant to follow up on libraries' forays into data archiving, since 30 institutions indicated they plan to offer data archiving within two years (Q56), but nearly a quarter of the respondents included providing data archiving in their top three challenges for RDM services (Q54), most commenting on the difficulty in setting up infrastructure, from software to storage requirements.

RDM Service Staffing

The depth and range of RDM services that libraries offer are, of course, directly proportional to staffing, both in the number of positions and the amount of time given to RDM activities when a position has other responsibilities. Currently, the most prevalent organizational structure for providing RDM services at the 53 responding libraries is a committee of staff from departments within the library (27, or 51%)(Q32). Less common organizational structures include a committee/group comprised of staff from across the university, including the library (9, or 17%), a single position within the library (8, or 15%), and a single department within the library (6, or 11%).

At the libraries where RDM services are provided by staff from different departments, no single department dominates (Q33). About a quarter of the departments provide reference/liaison services, followed by work in collections (19%), digital services (12%), research/instruction (12%), and systems/IT (11%).

The range in RDMS position titles reported shows that staff expertise is diverse and that no one type of position dominates either (Q38).



Out of 231 positions, the most frequently reported title is subject liaison/librarian (50 positions). One might expect RDMS roles to be more prevalent for liaisons to

the hard sciences, but subject specialists for all academic disciplines were represented nearly equally. The next most common position titles include the words "digital" (38 positions), "data librarian" (18 positions), or "metadata" (17 positions). Given that these positions come from across the library, it is not surprising that the majority (146, or 63%) have RDM activities added to their existing job duties, though 49 are new RDMS positions (22%). Only 34 positions were substantially redesigned (15%) to focus on RDM services (Q39).

The survey asked about these positions' responsibility for three broad RDM service roles: data management plan support, RDM guidance other than DMPs, and data archiving assistance. All but a few of the 53 libraries provide some level of service for all three categories (Q42). RDM guidance is the most common activity among both libraries (51, or 96%) and positions (183, or 82%). Data archiving and DMP support for grants follow closely (49 libraries, 154 and 150 positions, respectively). The majority of positions perform all three roles. In addition to providing RDM service, these positions devote a significant portion of their time to subject reference services (42%), cataloging/collection development services (27%), administrative tasks (25%), and other data services (24%), among many other duties (Q44).

How much staffing is required to provide RDM service so comprehensively? The number of positions and time spent on RDM activities varies widely across the responding libraries. Groups made up of staff from both the library and other departments in the institution have 10 members on average (Q34). Library committees/groups average about 8 members (Q33). Single library departments that provide RDM services average about 6 staff (Q36). Eight libraries have a single position that provides RDM services.

Although 90% of the positions are full-time, most of these individuals spend only a portion of their time on RDM activities. Respondents were asked to estimate the percentage of time spent on RDM for up to six positions within their institutions (Q43). Only 27 of the 213 reported positions (13%) spend 100% of their time on RDM services. Twelve spend between 60% and 90%, and 17 spend half their time on RDM activities. The remaining 147 positions (69%) spend less than half their time on RDM activities; nearly a quarter (47) spends less than 10%.

Education and Skills

In addition to assessing staffing models, the survey explored the training and educational backgrounds of current RDM service providers. We expected this emerging service area to require new and diverse skills, not always found within library service environments, and which often draw upon experiences from other professions or disciplines. The majority of staff with RDMS roles have MLS/MLIS degrees (172 individuals, or 75%) (Q45). Thirty-eight of these have masters degrees in other disciplines (22%), and six have PhDs in other disciplines (3%). Forty-three individuals (19%) hold only graduate degrees in a discipline other than library science; 20 hold a masters, 17 hold a PhD, and six hold both. Fourteen individuals have a degree with a data curation emphasis. Eight others have degrees with an archives emphasis. The range of other disciplines is tremendous and includes very minimal concentration in any single area. Only ten individuals have only an undergraduate degree.

Recognizing that formal academic backgrounds may not be the only influential factor in building a qualified RDM service team, the survey also asked respondents to select the three most important skill areas for the staff now in these positions (Q46). The top three areas are subject domain expertise (38 responses, or 75%), digital/data curation training (31, or 60%), and IT experience (30, or 59%). Respondents also explicitly noted the importance of training and experience in research methods, data analysis, research data management best practices, and scholarly communication.

Training Needs

Although many respondents stated that the function of libraries (access and preservation of knowledge) and current skill set of librarians lends themselves to RDM, many acknowledged that new skills and training are needed to optimally perform RDM. For future planning purposes, the survey asked which skills RDM staff most need further advanced training on. The most frequent response was identification and

application of appropriate metadata standards (25, or 49%). A second cluster includes digital preservation (17, or 33%), data ownership policies (16, or 31%), ethical and legal issues (15, or 29%), and subject domain expertise (15, or 29%). Some respondents also noted the need for deeper technical skills in related areas such as data acquisition, wrangling, analysis, interpretation, visualization, and deeper knowledge of research administration practices and forces. A few also acknowledged that their services were not yet developed enough to know what they would need next.

The survey next looked at the training methods that libraries have used to develop their RDM staff. Perhaps predictably, workshop attendance (48 responses, or 92%), conference attendance (44, or 85%), independent study (35, or 67%), and training provided by professional organizations (32, or 62%) rose to the top.[7] With more iSchools and MLIS programs offering data curation and digital collection emphases, and other fields of science and information technology emphasizing training in big data and digital data support, it will be interesting to follow trends in educational background of those entering the emerging library specialty of RDMS.

Funding RDM Services

Many survey respondents identified RDM service funding as a key challenge. This is not surprising since all but one library covers the costs through the regular library budget (Q49). Only a few have received external grants or a portion of research project funds. Only three have tapped endowment funds. Expectations for additional funding to support RDM services in the future don't indicate much change. Slightly more than half of the respondents to Q61 indicated that additional sources of funding have not yet been determined. The most frequently anticipated future funding sources are the regular library budget (21, or 36%) and external grant funding (15, or 26%). Some libraries expect to tap a temporary or special project budget, or receive funding from the parent institution. At the same time, 66% of survey respondents expect the allocation of funds for RDM services will increase in the next three years (Q62).

Partnerships

Building RDM services involves collaboration within the library, across a campus, and sometimes across institutions. Respondents' institutional models and levels of service development have varied widely for all the components of RDM services discussed thus far. Similarly, responses about which departments library RDMS staff refer researchers to reveal wide diversity in the degree to which these units interact. Survey respondents most frequently refer researchers to central IT and research administration, units with whom they also frequently collaborate on projects (Q52). Referrals are also directed to IRB and general counsel, but only a few libraries collaborate with these departments. A smaller number of respondents both make referrals to and collaborate with institutional administration, institutional archives, and other units ranging from statistical consulting groups to department or school IT units.

Although inter-institutional partnerships for service provision are happening throughout ARL libraries, the number of formal collaborations is still relatively small. Only 13 respondents (26%) have participated in an external partnership (Q53). Several of these were joint developers of the DMPTool. Other partnerships center on software and tool development, creation of training materials, and research for service provision.

Conclusion

Given the current technical and political environment, we (and most other followers of the research data curation field) predict that the need for universities to manage their research data for both access and preservation will grow, due primarily to two factors: the reliance in many fields of science upon technical ability to create large and complex digital data, and the increasing requirements and enforcement of data sharing policies by research funders. As the results from this survey show, at least 54 ARL libraries are responding to this need by providing RDM service in some capacity, and another 17 have plans to do so in the next few years. The majority of these libraries provide service across all three RDM activities: data

management plans for grants, guidance on data management, and research data archiving. Half draw on staff from different departments within the library to deliver these services.

A common theme throughout the survey is the recognition that, in order to provide comprehensive RDM services and to support scientists throughout the data lifecycle, libraries need to collaborate, either formally or informally, with other units at the institution. This is true even where the library has a dedicated department for RDM. At the very least, these units will then be aware of the services the library offers and can refer researchers to them (e.g., research administrators can send proposal writers to the library for data management plans). Forming these partnerships is listed as the biggest challenge by respondents, and in some cases has led to uncertain roles at the institution-level over which units have primacy over RDM. With so many aspects of RDM services overlapping domains and defining new territories of collaboration among multiple units within an institution, it will be interesting to follow how libraries continue to position their roles within the school as such services expand, and how unified or diverse an approach an institution may choose to take overall in supporting research data management.

It costs money to provide quality services. RDMS requires a diverse range of skills, many outside the typical expertise of library staff and not all libraries can afford to hire new and/or retrain staff. Creating archiving infrastructure and curating research data are also expensive endeavors. Right now, the amount of archived data is relatively small for the majority of institutions; however, as funders become more stringent in data retention and sharing requirements, libraries will need to employ an alternative funding model, such as through fees to researchers or their grant projects.

Finally, the third biggest challenge reported is faculty (non)engagement due to a lack of awareness of services that the library provides, low perceived value of services, and resistance to data sharing. Respondents stated that the most effective marketing techniques were through workshops and presentation to researchers, referrals from research project (grants) administration, and direct emails to researchers (Q50). The low perceived value of services is a more difficult issue to overcome. Some researchers do not view the library as a resource for data management, and as one respondent commented, "preservation and data sharing are a hard sell when the researcher only need[s] to write a plan." Also, some respondents noted that funders are not taking mandates seriously, giving the researcher a low incentive to care about data management and sharing. As noted, the push for open data access for publicly funded research and compliance for data sharing policies, both in North America and around the world, may change the environment significantly. Our survey results suggest that many ARL libraries have at least a start toward growing services to meet an increasing demand.

This survey provides a snapshot of what RDM activities ARL libraries are currently involved in, what human resources are being used to provide these services, and projected service provision. Although providing RDM services is not easy and requires a heavy investment in hiring/retraining staff, building technical infrastructure, and continually reaching out to and collaborating with other data management players on campus, many respondents felt that library could and should support RDM activities to some degree (Q66). Although RMD services are relatively new, institutions are taking diverse approaches to providing them, and will likely evolve over the next few years. The exact nature of how service will be provided will likely depend on institutional and funder policy, technical skills of library staff, and the financial position of a library.

Limitations of Survey

The survey responses clearly show that RDM services do not happen just in the library. They require pan-institutional collaboration. One major limitation of this survey is the absence of responses from those outside the library. If RDM services go beyond the library but are unknown by the library, the answers in this survey may not be a complete picture of the institutions' RDM services. One particular library commented: "We don't know the answers to any of these and don't want to speak for other units."

The survey underwent review and testing before it was sent out. We adjusted many definitions and

questions, but that did not eliminate problems with respondents misinterpreting concepts and definitions. Several of the metrics of service, such as the extent of data management plan assistance and archive use, were particularly difficult to define and ask in ways that yielded precise responses. We recommend that further studies involve case studies, and focus on particular topics such as archiving or staffing, since we presented our respondents with a particularly long and complicated survey. We greatly appreciate their efforts to complete our survey, and hope these results will be a useful benchmark and basis for inspiration in this new and expanding field of research library service.

Acknowledgements

We would like to sincerely thank the following individuals and groups for reviewing our survey instrument: Andrea Denton, Mike Furlough, Brian Gunia, Patricia Hwse, David Lowe, Karl Nilsen, Susan Payne, Lizzy Rolando, Jennifer Ward, Jonathan Wheeler, Lynda White, Stephanie Wright, the Canada Institute for Scientific and Technical Information (CISTI), and the E-Research Working Group. We would also like to thank the authors of the e-science report for providing helpful background and context for the report and for advising us what we ought to consider including in the RDM services survey. Tim Dilauro, Johns Hopkins University, provided his perspective and expertise in developing the survey instrument. Finally, we'd like to thank Wendy Mann of George Mason University for arranging a space at the Fenwick Library for the University of Virginia and Johns Hopkins University authors to hold in-person meetings.

Endnotes

1 For a sample see the references in the Selected Resources section of this SPEC Kit.

2 http://www.arl.org/storage/documents/publications/escience-report-2010.pdf

3 See heading "Key Papers in the Development of RDMS" in the Selected Resources section of this SPEC Kit.

4 http://www.arl.org/focus-areas/e-research/e-science-institute

5 See White House Office of Science and Technology Policy (OSTP) Feb. 22, 2013 memorandum on open access to funded research data and publications.

6 Dataverse is in relatively widespread use as a repository for specific disciplines and research centers, but only one library reported being directly involved with a Dataverse implementation.

7 Attendance at an E-Science Institute workshop noted earlier in responses to Question 3 was another influential training resources for many respondents.

SURVEY QUESTIONS AND RESPONSES

The SPEC Survey on Research Data Management Services was designed by **Barbara E. Pralle**, Head, Entrepreneurial Library Program and Interim Manager JHU Data Management Services, **David Fearon** and **Betsy Gunia**, Data Management Consultants, at the Johns Hopkins University Sheridan Libraries; and **Andrew L. Sallans**, Head of Strategic Data Initiatives, and **Sherry Lake**, Senior Scientific Data Consultant, at the University of Virginia Library. These results are based on data submitted by 73 of the 125 ARL member libraries (58%) by the deadline of May 6, 2013. The survey's introductory text and questions are reproduced below, followed by the response data and selected comments from the respondents.

This study surveys ARL member libraries on their activities related to access, management, and archiving of research data at their institutions. This introduction will help identify who should respond to the survey questions, and we encourage involving others at your institution to assist in filling out this survey.

Over the last decade, most research libraries have provided some degree of support services for research data access and use. Over the last few years, many found they needed to extend and unify services around more aspects of data acquisition, management, dissemination, and preservation. The steady increase in e-Science—digitally mediated research with large datasets and networked collaborative use—is one reason researchers look to their libraries for help with organizing, sharing, and archiving data.

The survey's purpose is to assess the current landscape of how libraries, in relation to their parent institutions, are providing research data management services to their community. It will not only provide benchmarks for trends, but will also help libraries gauge their level of service for further development, and discover and share new models of service. The survey explores the organization of research data management services (including a few questions on broader data support services), how they are staffed and funded, and what services they offer and to whom, among other questions.

This survey expands the 2009 ARL E-Science Task Force survey on E-Science and Data Support Services, updating recent developments and adding scope and detail for services addressed in the 2009 survey. You may have received similar surveys. This one is intended to go more deeply into the details of RDM services, as a benchmark survey of ARL member institutions.

Research Data Management Services support the management and curation of research data throughout its life cycle. RDM includes services such as: data management plan consulting, data documentation/metadata, data organization, data security and backup, data citation, funder requirements, ethical and legal issues, preserving digital data, sharing data and archiving data. For this survey, services are for research data, not institutional data such as departmental records keeping or government archives. It can include research data in the sciences, social sciences, and humanities.

NOTE: Respondents whose libraries are providing data management consulting, and/or operating data archives or institutional archives that host data, will answer the majority of the questions. In some cases, this survey may take more than one hour to complete. We thank you in advance for taking the extra time required for these questions, which will be an important benchmark for mapping the development of research data services in research libraries.

BACKGROUND QUESTIONS

1. Does your parent institution have a data retention policy or formal research data management policy in place? (It may be associated with legal/regulatory compliance, intellectual property, technology transfer, or research administration policies. It may also be associated with a document retention policy or research practice policy. Please share links to the policies, if available, in the "Call for Documents" section at the end of the survey). N=73

Yes, there is a policy now	16	22%
No, but one is planned in the next 1–3 years	24	33%
No	32	44%
Don't know	1	1%

Comments N=21

Yes, there is a policy now

Current policy only covers IP & technology transfer.

The parent institution's policies dictate records management and retention guidelines for institutional data, whereas individual investigators have ultimate responsibility for their own research data. Research data that is also institutional data is subject to institutional data handling requirements. With respect to research data per se, the institutional policy is simply that "researchers are encouraged to retain research data and records for a period of at least five years following publication to provide verification of the validity of the reported results."

There is a policy, but it is not formally resourced or "centralized." It is currently under revision and goes before the Board of Governors in June 2013.

The university has many policies regarding research data management and retention. They are not currently sufficiently comprehensive and the supports for the actual practice of data management are not yet sufficiently in place for the full support for all involved for the full data curation lifecycle. These full supports are in development for implementation.

University Records Retention Schedule.

Yes, but only a general data retention policy that dates to 2004 and is oriented towards laboratory notebooks and general recordkeeping. We lack a comprehensive, formal research data management policy.

No, but one is planned in the next 1–3 years

A plan is currently in draft form but there is no clear implementation timetable. The draft was completed in December 2012.

A research records policy has been approved by the Board of Governors. We are in the midst of consultations with stakeholders on campus around further procedures and guidelines.

At least I hope it's in place by then. We've had two provost's task forces investigating what a policy should include.

The library plans to have a data retention policy within the next three years, although we are at the early stages of engaging campus leadership on these issues.

We are still formulating our policy but we have laid a foundation for data services already, which we want to include in this survey.

No

But, the university's retention policies for grant-funded research data are covered as a category in the retention schedule.

IRB policies, but few other guidelines for non-sensitive data.

No formal institutional policy on data management. The "Responsible Conduct of Scholarship and Research" policy addresses retention, but not ownership, of research data.

Nothing beyond regular records retention policy.

The Office of Sponsored Programs references the Council on Government Relations publication, "Access to, Sharing and Retention of Research Data: Rights and Responsibilities," published March 2012.

The university provides data retention guidelines to the campus research community.

There are guidelines on responsibilities for research ethics that talk about data, but that's it. We don't have a policy or guideline specifically for research data management.

We have a data access policy, which is currently being updated, but nothing else.

We have a records retention policy, but not a research data policy that the library is aware of.

We have related policies, but none that actually require research data be retained or managed. Records retention applies to a broad range of topics but is not interpreted as related directly to data.

2. Please indicate if your library has participated in the ARL/DLF E-Science Institute(s) or plans to participate in the ARL/DLF/DuraSpace E-Science Institute when it is next offered. N=53

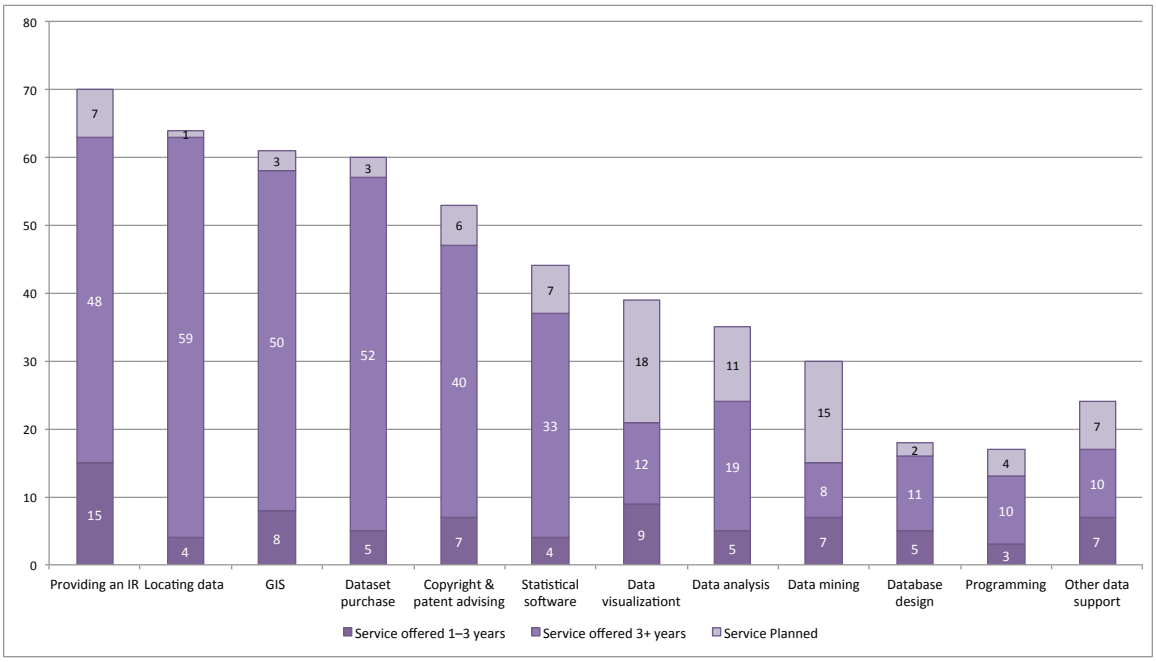
2011–12 ARL/DLF E-Science Institute	43	81%
2012–13 ARL/DLF E-Science Institute	7	13%
Future ARL/DLF/DuraSpace E-Science Institute	7	13%

BROAD DATA SUPPORT SERVICES

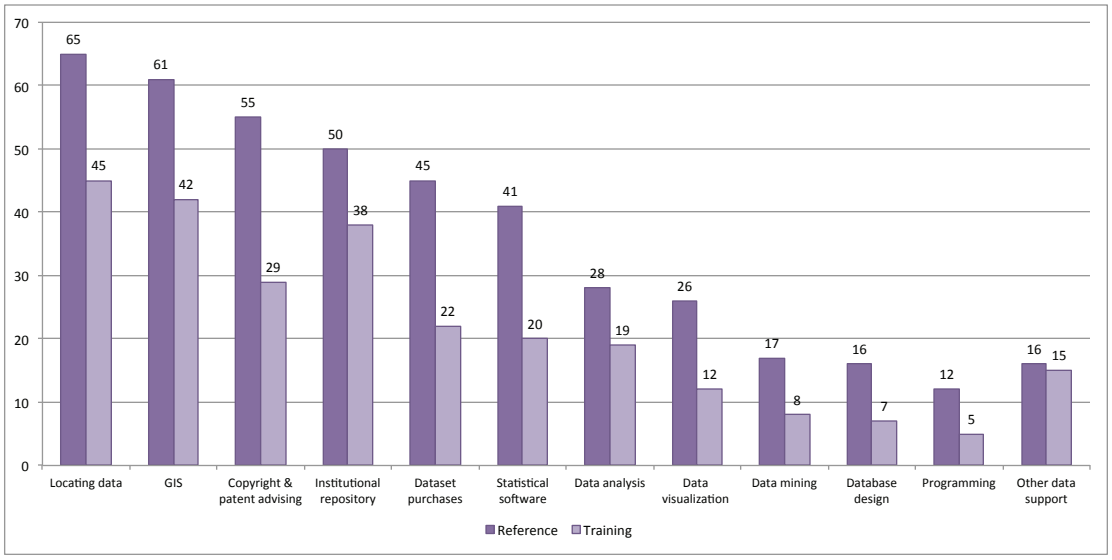
Most of this survey will focus on research data management (RDM) services as described in the introduction. Our first question, however, asks about the broader range of services provided at your institution that support research and data access.

3. Please indicate whether your library offers information resources, assistance, and/or provides training on each topic below. Also indicate how long your library has provided the particular service (1–3 years, or more than 3 years) or if there are plans to provide the service. Check all that apply. N=73

Topic	Offers Service	Service offered 1–3 years	Service offered 3+ years	Service Planned	N
Providing an institutional repository	64	15	48	7	71
Locating & using existing data sources (for sciences, humanities, government, medical, etc.)	68	4	59	1	69
GIS and geospatial analysis, support	61	8	50	3	64
Dataset purchase, acquisition, subscriptions	58	5	52	3	61
Copyright & patent advising	54	7	40	6	60
General statistical software support	42	4	33	7	49
Data visualization support	26	9	12	18	44
Data analysis support	28	5	19	11	39
Data mining	20	7	8	15	35
Database design & management	20	5	11	2	22
Programming/software development	17	3	10	4	21
Other data support service(s)	21	7	10	7	26
Total Responses	72	42	66	37	72



Topic	Reference	Training	N
Locating & using existing data sources (for sciences, humanities, government, medical, etc.)	65	45	65
GIS and geospatial analysis, support	61	42	61
Copyright & patent advising	55	29	55
Institutional repository	50	38	51
Dataset purchase, acquisition, subscriptions	45	22	45
General statistical software support	41	20	44
Data analysis support	28	19	29
Data visualization support	26	12	27
Data mining	17	8	20
Database design & management	16	7	18
Programming/software development	12	5	14
Other data support service(s)	16	15	18
Total Responses	72	63	72



If you selected Other data support service (besides research data management services), please briefly describe the service(s) your library offers. N=23

Data preparation support: geocoding, georeferencing, distance/area/density calculations, converting formats, merging/joining, extracting/subsetting, and digitizing.
Data access: maintain a data extractor and data fileserver.

Informatics support 1–3 years, personal archiving training 1–3 years, subject repository 3+ years, media repository 3+ years, digitalization (data creation for mining) 3+ years, data transformation/normalization (campus-level support, i.e., faculty metrics) 1–3 years.

Locally host purchased numerical and GIS data as necessary. On-demand digitization of paper maps.

Measuring research impact—guides and workshops on research impact metrics (e.g., h-index), measuring journal impact, altmetrics, pricing information (scholarly materials), citation tracking.

Metadata consulting, Persistent Identifiers/data citation.

Programming/software development is only available to faculty who have funding to contribute. We act as an intermediary for acquiring data for free, including providing confirmation of academic need on behalf of students, brokering gifts of desired data from governments (local, state or federal) or non-governmental organizations and, if needed, filing FOIA requests for specific data sets. We also provide file format conversion help. We also offer consulting on data management solutions, referring to appropriate services across campus (including to our own IR) as well as training generally on data management issues. While these services have mainly been in place since January 2011, the library committee behind them has been in place since summer 2008. Several organizations on campus provide advanced, customized and often automated data management solutions, some of which are mediated and some of which are self-serve.

Through our digital humanities campus partnership, the Libraries provide some level of support for textual analysis and visualization (which you might consider a form of data mining and data visualization). We also consult on issues related to long-term data retention and/or preservation for a wide variety of data types, including "image" data (photographs, illustrations, etc.)

We provide an API for accessing, using, and querying the digital collections/libraries hosted at the university. Current needs are being reviewed for the addition of other services, trainings, and supports.

We are exploring a range of metadata support and curation services.

We have been supporting GIS visualizations and are exploring new options such as Tableau. On the Chinese Canadian Stories project we worked with the Stanford Spatial History Project to create some visualizations.

Web scraping/harvesting, metadata management.

While the Libraries' liaison to computer science has provided reference help with data mining, data visualization, database design and management, and programming/software development questions over the years, these support services are not officially part of our research data management services framework of services.

Working with campus IT to develop a university data management plan.

RDM Services

A LibGuide of resources for data management is currently in progress.

Advice for data management plans.

Data curation and legacy data conversion; active research data storage; data management planning.

Data Library has been in existence since 1992 offering a number of data support services. We are engaged in a wide spectrum of data curation, RDM and preservation activities and partnerships.

Data management plan training and support, DMP tool

Data Management Planning (DMPTool)

Data management plans

DMP Tool in place for the university community. We provide training on use of the Tool.

EZID, WAS, Merritt support via LibGuide

Research Data Management Services

RESEARCH DATA MANAGEMENT SERVICES

For the rest of this survey, you will be asked about your library's provision of research data management (RDM) services, which we define as providing information, consulting, training or active involvement in: data management planning, data management guidance during research (e.g., advice on data storage or file security), research documentation and metadata, research data sharing and curation (selection, preservation, archiving, citation) of completed projects and published data.

4. Does your library offer any research data management (RDM) services as described in the introduction and above? Answer "yes" even if the extent of your services are reference resources for data management plans on your library's website. You will indicate the range of services and any additional planned services in follow-up questions. N=73

Yes	54	74%
No, but we plan to	17	23%
No, and we have no plans to	2	3%

Comments N=11

Yes

Minimal. Information only through a LibGuide.

Some campus solutions even offer management solutions for projects in progress (not just completed ones) and unpublished data, i.e., embargoed data that may be shared in the future but is not publicly available now.

Some of the services are provided at the system-wide level.

We are starting up these services now.

We have resources and have assisted faculty in the development of data management plans and the preservation and access of data. The service is limited but we hope to extend if staffing can be acquired.

We have strengths in several of these areas (e.g. metadata design) but are continuing to build and develop a suite of services to better support all of these roles.

We provide basic level support of research data management planning.

No, but we plan to

Planned LibGuide on website.

The university libraries in partnership with the Department of High Performance Computing are working on a RDM partnership for fall 2013.

We have done ad hoc data management planning, metadata, and curation support when faculty have come to us, but we do not yet advertise or fully staff a standing service.

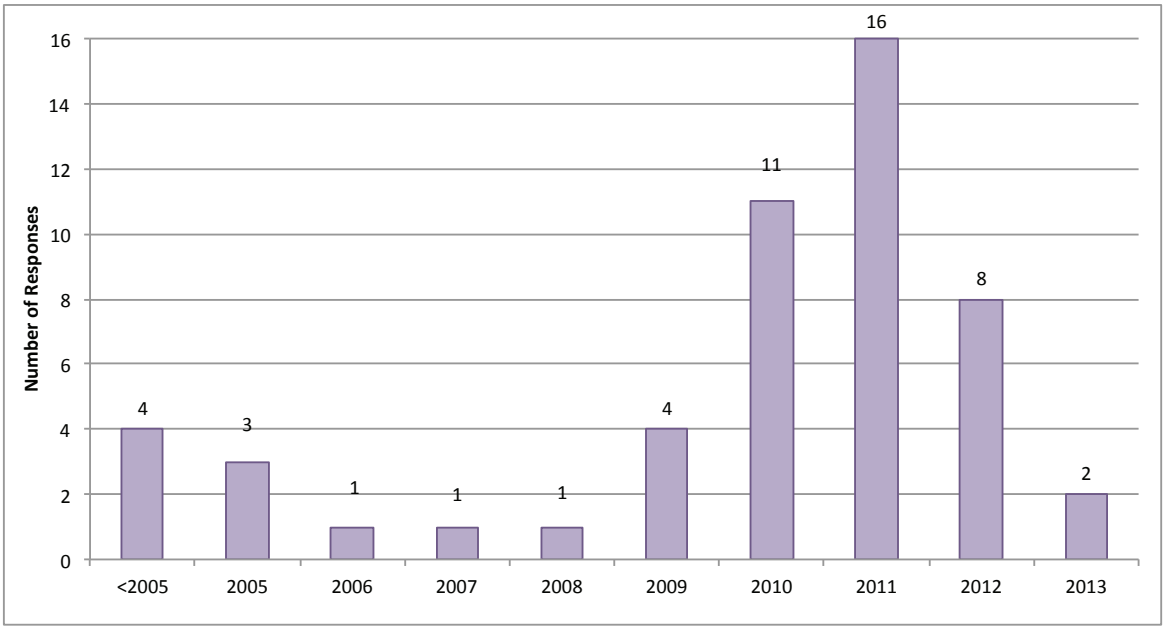
No, and we have no plans to

No, we wish we could, but we have no formal plan to do so in the short term because of lack of resources.

If you answered Yes, skip to the section on **Origin of RDM Services**. If you answered No but plan to, skip to the section on **Other Units Offering RDM Services.** If you answered No and no plans to, skip to the section on **No Library RDM Services**.

ORIGIN OF RDM SERVICES

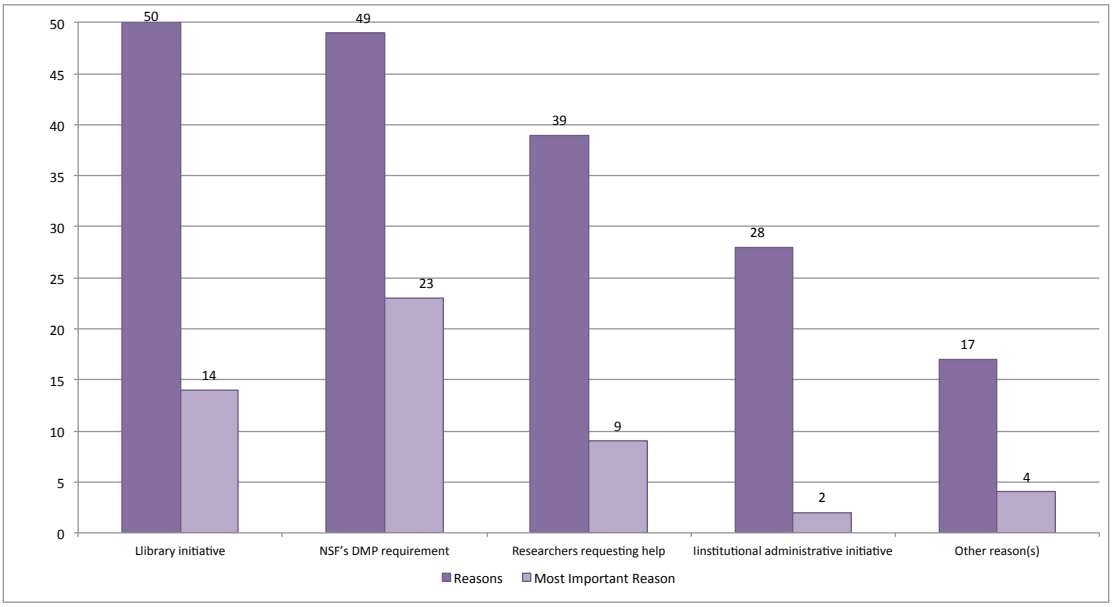
5. Please enter the year when your library initiated RDM services. (This can include simply adding references to your webpage if this is the extent of your RDM services). N=51



Range: 1966 to 2013

6. Please indicate your library's reasons for initiating RDM services, either as a new service or an expansion of existing services. Check all that apply. Then select one reason that is the most important to the library. N=54

Reasons	N	Most Important One
A library initiative to expand support for faculty research	50	14
NSF's Data Management Plan requirement of January 2011	49	23
Researchers requesting help with data management, data sharing, or preservation	39	9
An institutional administrative initiative to support research data services	28	2
Other reason(s) for initiating RDM services	17	4
Total Responses	54	48



Please briefly describe the other reason(s) for initiating RDM services. N=17

Area of research interest on the part of our data librarian.

Broadly built collections that have data as well as other content. Institutional initiative for supporting research data.

Developing trend among peer institutions.

Enable the success of graduates by providing data management resources and training.

In addition to NSF, other funding agencies requiring data management plans. Also, these data services are a natural extension to the support that the library already provides to the research community.

Institutional history for archiving research data.

Invitation to collaborate with our Sponsored Programs and research administration office on campus

Librarians established a data management working group in the summer of 2008 to begin educating themselves about data management in response to a couple of requests from faculty for help in this area. We were well positioned then to step up our efforts when the NSF announced the gist of its DMP requirement in the fall of 2010. Our most recent five-year plan for the library has been being formulated over the last year and data issues have made a major appearance therein, partly as a result of our participation in the e-science institute and two university task forces on formulating institution-wide policies for data management. These three initiatives together have highlighted that the campus has a LOT of offices offering data management services and solutions, which have grown up gradually without an institutional administrative initiative to support research data services, but the institutional administrative initiative now is to better coordinate services, make researchers aware of their existence and protect the university from liability issues that might arise from not adequately managing data.

Library strategic direction.

Our associate university librarian at the time was a visionary in terms of DRM and initiated a broad range of services, including hiring new non-librarian technical staff.

Our first research data librarian was hired in 2006, thus initiating our RDM services. Additional services were implemented as a result of recommendations from a library working group, set up to review current research data landscape on campus and provide recommendations and suggest opportunities for the library to pursue related to data curation and sharing. When the NSF's DMP requirement was announced in 2010, the university, with support from the university librarian and the vice provost for research, formed a cross-institutional group to offer additional RDM services to our researchers.

Our university librarian co-authored 2007 ARL report on "Agenda for E-Science" and this work led to the creation of an E-science and Data Services Collaborative in 2007.

Partnership with campus IT to provide updated infrastructure and support for research data.

Recognition of the growth of e-Science.

Service initiated in part as an extension and implementation of the NSF funded Data Conservancy project and data archive led by the university.

The university has provided research data management services since at least 1966 with the Latin American Data Bank project, which was supported through the libraries. Because the libraries have been actively engaged with faculty for managing, sharing, and preserving data, new supports and services have changed with technologies and needs.

Three library faculty appointed to the "Research Data Management Task Force" created by the ice president for research.

Please enter any additional details you wish to share about how and why RDM services were initiated at your library. N=17

1992 start date came out of an action item from the library strategic plan. The plan was based on community needs as well as the library's anticipation of new service demands.

Addressing a need on campus.

Although we list 2011 as the date for library initiated RDM services, some levels of support for research data have existed in our libraries and in partnership with other entities on campus since at least 2000.

Filling gap identified by faculty.

General sense of urgency out of ARL and other information organizations around RDM, data, and e-Science.

In 2005, an ad hoc data group was established. In 2010, research data management services began to be offered. In 2011, a framework of defined research data management services was developed with further coordination and defined staffing. In 2012 to the present time, a research data services working group was established to further assess data needs and gaps in data services.

In 2008, the working group recommended educating researchers on emerging requirements and helping researchers meet those requirements. Other opportunities for supporting researchers included providing information on best practices in managing data and referring researchers to appropriate sources of information and expertise across campus. With the advent of the NSF DMP requirement in 2011, the Research Data Management Service Group (RDMSG) was formed as a cross-campus, collaborative organization to provide RDM services to faculty, staff and students; the library took a leading role in the formation of this data support group.

In 2011, the university created a top-level IT group for research computing, building on existing excellence in high performance computing. The new group oversaw dramatic improvements in the already excellent research computing infrastructure including seeing a new data center come online in 2013 and the network access increased to 100Gbps. This recent work paralleled ongoing work by both research computing and the libraries to support research data needs, and to partnered work on the shared need for more research data management support. That work has led to the current data management/curation task force that is developing more plans and supports.

It seemed that this was an important new area for academic libraries to become involved with.

Our library director also made this effort a priority here.

Our support position for this service is currently vacant, but we are recruiting.

Restructuring of prior service unit combination of library "digital services" and IT research computing services with new focus exclusively on issues of research data management. Gained administrative buy-in from offices of VP/CIO and VPR.

The library is well positioned in all this to act as an information and referral point for all campus RDM services. It's clear that we cannot address issues like technological infrastructure but we're building important partnerships to share information among these many campus stakeholders.

The Social Science Library began collaborating with the StatLab early on in the StatLab's existence. The StatLab and Social Science Library provided research data management services to the university community, focusing on social science data and research. Now as a combined unit the Center for Science and Social Science Information provides research data management services to the community.

We deploy a diverse, multidisciplinary support team.

We identified prominent research groups and faculty on campus and began systematic outreach to them. After assessing their needs, we began building the technical and human architecture and staffing to support their research efforts.

We offer a mix of services that are driven by library initiatives and broader campus ones.

ONLINE DATA MANAGEMENT PLAN (DMP) SERVICES

Note: The following questions are for currently implemented services. If your library is *planning to* offer any of the particular RDM services listed below, you will be able to indicate planned services later in the survey.

7. Does your library have an online resource related to Data Management Plans (DMPs) for NSF proposals or other funding agencies? Choose "Only offered elsewhere" if the library does not offer the service, but you are aware of this service being provided by a department or group outside the library. Choose "not offered" if this service is not provided at your institution. N=54

Yes	28	52%
Offered by the library AND elsewhere	19	35%
Only offered elsewhere	2	4%
Not offered	5	9%

If yes, please indicate whether the components below are included in the online resource and how they were developed. Check all that apply. N=47

Component	Linked to	Developed our own	Customized another's for own use	Not included	N
Explanation of DMP requirements by different funding agencies and/or NSF directorates	34	24	11	0	47
Guidelines for creating DMPs	32	24	14	0	47
Template examples of DMPs	29	18	12	5	46
A tool or resource for DMP creation	36	9	11	1	43
Other component(s)	2	9	1	0	11
Total Responses	42	33	21	5	47

If you selected Other component(s), please briefly describe the component(s). N=11

A data planning checklist.

Boilerplate text to include to use our data repository (not a full template).

Copyright considerations, data citation guidelines, metadata examples, examples data sets on campus, subject repositories for data (list), data archiving guidelines, file naming best practices, data storage options on campus, related campus services, recorded workshops,

Data registry for collecting information on data developed at the university.

FAQs for data management, depositing data in digital repository, list of data repository options, copyright information, educational materials.

Information about digital repository services.

Information on long-term management and preservation of research data. Tools for managing metadata for research data.

List of services and resources available on campus for supporting research data management; slidedecks of workshop on basics of data management planning.

Local listserv for data management questions and data management plan reviews (moderated by the library), contact information, DMP examples and example language.

Our section on development of a data management plan also links to university resources that can assist the researcher for each main section included in a plan.

Workshops offered in person and streamed online, and recorded for future use on issues of concern such as what the NSF requirement was and would mean for researchers; a variety of solutions for managing data and how to write plans that would include those solutions; dealing with sensitive data (including GIS data); dealing with the many regulations for sensitive data; and more. Brief, online, stand-alone tutorials to help any researcher better understand what data management, DMPs and some foundation basics of data management like file naming conventions.

If online resource related to Data Management Plans are offered elsewhere, please specify which departments/units (e.g., central IT, research administration/sponsored projects office, IRB, health sciences data management, etc.) have online DMP resources (besides links to your library's web resources). N=20

Offered by the Library AND Elsewhere

Central IT, IRB

Central IT, Texas Advanced Computing Center (TACC)

Heath Sciences, Office of Sponsored Programs

High performance & research computing center (called Center for Computing and Visualization)

Office of Research

Office of Sponsored Research Projects

Office of sponsored research, University of California system-wide resources

Our Center for Computational Science has also done some work in this area.

Research center

Research computing (central IT)

Research Computing and the libraries offer DMP resources. Also, they partnered for the addition of the university to the DMPTool, which is another resource linked to, customized, and supported in partnership.

Research Data Management Services, Sponsored Research, Graduate School, Various Colleges, etc.

Research office/sponsored projects, Office of Responsible Conduct of Research, Technology Transfer, central IT

Responsible Conduct of Research, under the Office of Research Integrity Assurance

Some research institutes offer limited DMP services to researchers affiliated with the Institute.

The Odum Institute for Research in Social Science. (Our library data management committee has representatives from our health sciences library as well as the main academic library. And while our IR is part of the library, they have separate DMP materials from our DMC ones more specifically related to deposit in the IR.)

The university's online resource for DMPs is provided by our Research Data Management Service Group (RDMSG), of which the library is a part. We specifically chose to not duplicate data management information on the library website, but the library has had an active role in creation and maintenance of content on the RDMSG site since inception, and many of the services referred to in the DMP planning portion of the site refer people to the library.

We work closely with Sponsored Projects and their web pages link to Libraries' web pages, and vice versa.

Only Offered Elsewhere

Central IT, Research Administration, Office of Sponsored Programs

Research Computing/Office of Research and Technology Management

8. Does your library online resource provide links to the DMPTool (http://dmptool.org or https://dmp.cdlib.org)? N=54

Yes	41	76%
No	13	24%

9. Does your library provide direct, in-person guidance or training for use of the DMPTool, beyond their online help guides? N=54

Yes	29	54%
No	25	46%

DMP TRAINING & CONSULTATION SERVICES

10. Has your library offered training classes, sessions, or workshops on data management plan (DMP) preparation? N=54

Yes	33	61%
No	21	39%

11. Does your library offer direct assistance or consultation, either in-person or by email/chat, on data management plan (DMP) preparation for grant proposals, and/or data management planning support. (This could range from answering inquiries about plan preparation to more formalized consultation meetings). Choose "Only offered elsewhere" if the library does not offer the service, but you are aware of this service being provided by a department or group outside the library. Choose "not offered" if this service is not provided at your institution. N=54

Yes	25	46%
Offered by the library AND elsewhere	23	43%
Only offered elsewhere	2	4%
Not offered	4	7%

If this service is offered elsewhere, please identify the department or group that provides this service (e.g., central IT, research administration/sponsored projects office, IRB, health sciences data management, etc.) N=23

Offered by the Library AND Elsewhere

All data management planning services, including training and workshops for DMP preparation are offered through our Research Data Management Service Group. The library is a part of this cross-campus, collaborative organization.

Center for Computational Science and possibly the Office of Research

Central IT, TACC

Clinical and Translational Science Institute, Data Coordinating Center

College of Arts & Sciences: Office of Research Funding and Support. Research and Health Sciences IT.

Colleges, research & sponsored projects.

Heath Sciences

High performance computing and research center

Institute for Policy and Social Research, Center for Research Methods & Data Analysis

IRB, Sponsored programs, Research Administration, but this is ad hoc and not a centralized or organized effort.

Office of Research

Office of Research, Office of Sponsored Programs

Office of Research and Engagement

Research administration

Research center, Research administration/sponsored projects

Research Computing (central IT) Research Computing links and directs these to the Libraries, but provides support specific to High Performance Computing as well.

Some research institutes offer limited DMP services to researchers affiliated with the Institute.

Sponsored Program Services/Pre-Awards (mostly done in conjunction with the Libraries).

Sponsored programs proposal development office

The Odum Institute also offers such consultations. Other campus groups that provide DM solutions (for which they would help a researcher write a DMP) include a division of campus IT called Research Computing; a semi-independent computing institute based on campus called RENCI; and the Data Intensive Cyber-Environment (DICE) group based in the library school.

Only Offered Elsewhere

Office of Research Services and University Information Technology departments consult on an ad hoc basis only.

Office of Sponsored Research Projects

If you answered Yes or Library and elsewhere, skip to the section DMP Consultation Services. If you answered Only offered elsewhere or Not offered, skip to the section on RDM Services other than DMP Support.

DMP CONSULTATION SERVICES

12. Please enter the year when your library initiated direct DMP consulting or guidance services. N=47

2010	2011	2012	2013
10	23	12	2

13. Which consultation method does your library use? Check all that apply. N=48

By email/chat/phone	45	94%
Meetings with researcher at researcher's office, lab, or other location	44	92%
Faculty/researcher visits library (reference desk-style support)	29	60%
Other consultation method	12	25%

Please briefly comment on the consultation method your library uses. N=21

An email form for questions is provided on the data management webpage and personal consultations are mentioned.

Basically, we'll meet with researchers in whatever way/location they prefer.

Faculty survey

In addition to providing our contact information to faculty through departmental grant administrators and our partners in the institution, faculty can also indicate that they would like a consultation through our internal grant tracking system.

In regards to "faculty/researcher visits library" option: we don't design our services for consumption at the reference desk, but we have received some referrals for service via our reference desk staff.

No faculty has contacted the library and asked for help.

Online training and videos

PI generally contacts the library and their needs dictate how the consultation proceeds.

Researchers contact me directly via email most of the time.

The vast majority of the consultations have taken place via email.

Training classes on DMPs, with shared Q&A. This is rather recent, and is still developing. This support is not currently/yet being tracked separately.

Tutorials, training, collaborations with Office of Contracts and grants

We can meet with faculty either in the library as an appointment or at the researcher's office.

We extended our Ask-A-Librarian digital reference service to include data management, including DMP consultation, with special training and routing for data reference questions.

We offer all modes of consultation, but most contacts so far have been in the form of workshops.

We offer group information sessions on DMP preparation and have a ticketing system where researchers can request help in whatever format they prefer: email, phone, web conference, or in-person. We also offer walk-in style assistance several hours per month at libraries across campus where people can get any data management guidance, including DMP preparations assistance.

We receive requests for consultation via our dedicated listserv or through direct email to one of the Data Working Group members. We ask for one-week to turnaround reviews, but frequently have to meet deadlines in one or two days. All members of the working group can provide feedback; feedback is collated, prepared and emailed back to the faculty requesting consultation. Follow up phone calls or emails are offered and welcomed.

We respond to questions sent to our email address, data-management@mit.edu, also, our "outreach" efforts sometimes lead to consultations.

Whatever the researcher needs.

With Sponsored Projects, we offer one-hour workshops for faculty and graduate students on data management planning and other data management topics, roughly once per semester.

Workshops

14. Does your library track the number of consultation sessions held? N=48

Yes	28	58%
No	20	42%

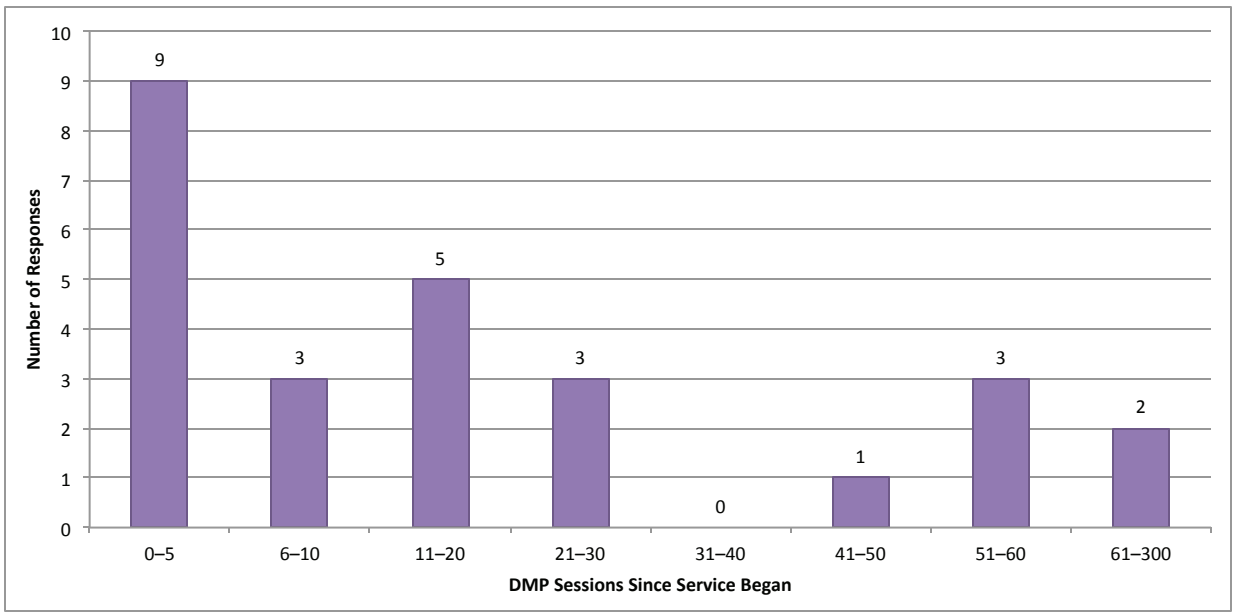
If yes, please enter the number of data management plans consulted on per month, on average. If possible, please also enter the number of DMP sessions since this service began. N=26

Average number of plans per month

Average	Responses
0	4
>0 but <1	7
1	7
1.58	1
2	2
2.5	1
2.7	1
3	1
8	1
14	1

DMP sessions since service began

Minimum	Maximum	Mean	Median	Std Dev
0	300	31.62	11.50	59.72



15. Does your library offer iterative feedback on researchers' data management plans? N=46

Yes	37	80%
No	9	20%

16. To the extent possible, please estimate roughly how much overall time is spent per researcher consulting on DMPs for each grant proposal. (Include meetings, feedback on drafts, and other associated work on a proposal.) N=31

Transactional Only (in Minutes) N=7

Minimum	Maximum	Mean	Median	Std Dev
30	60	45.0	45.0	15.00

Iterative Only (in Hours) N=12

Minimum	Maximum	Mean	Median	Std Dev
1	10	4.6	3.0	3.26

Both Transactional & Iterative N=5

Duration	Minimum	Maximum	Mean	Median	Std Dev
Minutes	5	240	70.0	30.0	96.11
Hours	1	4	1.5	1.5	1.24

Transactional, Iterative, & Project N=6

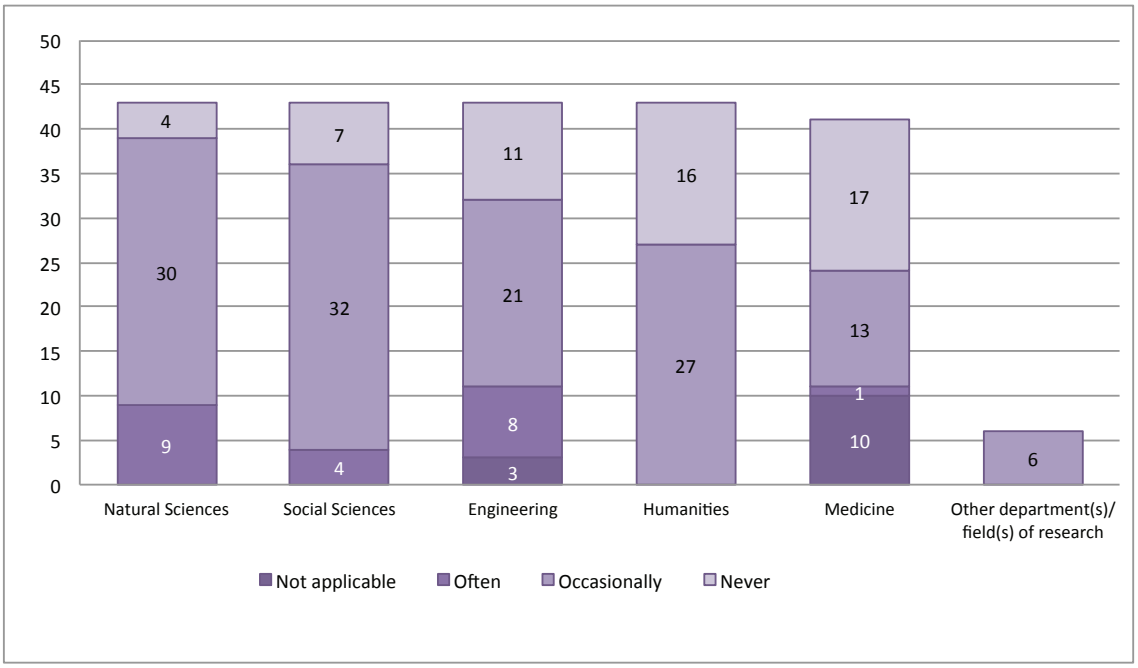
Duration	Minimum	Maximum	Mean	Median	Std Dev
Minutes	30	90	60.0	60.0	18.97
Hours	1	3	1.8	2.0	0.75
Days	.8	10	3.6	2.5	3.48

Additional Comment

The overall time varies greatly from transactional, iterative, and projects.

17. Please estimate how often each department/field of research at your institution uses your DMP services. N=43

Department/Field	Never	Occasionally	Often	Not applicable	N
Natural Sciences	4	30	9	0	43
Social Sciences	7	32	4	0	43
Humanities	16	27	0	0	43
Engineering	11	21	8	3	43
Medicine	17	13	1	10	41
Other departments/fields of research	0	6	0	0	6
Total Responses	28	42	13	12	43



Please specify the other departments/fields of research. N=6

Agriculture

Business

Libraries

Other areas include the arts and agriculture.

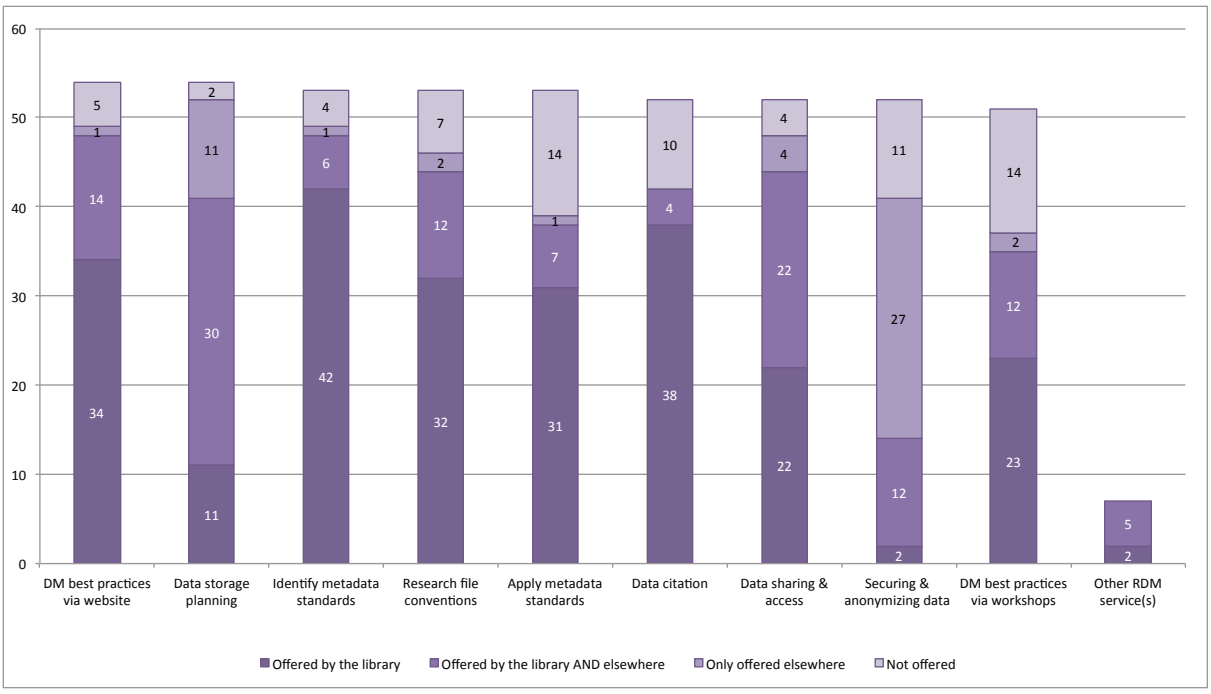
Our contacts have come from all different departments but we've had so few, the data aren't reliable for describing trends. We're initiating in the next month an analysis of DMPs submitted to NSF since their requirement was implemented (both funded and not), partly to check whether researchers are claiming they'll use our IR for data management without having first secured a deposit agreement.

Physical sciences and other various unspecified.

RDM SERVICES OTHER THAN DMP SUPPORT

18. Please indicate which of the following RDM resources, consulting, or training services your library provides. These services support ongoing research beyond grant proposal preparation. (Data archiving will be addressed later.) Please select one response in each row. Choose "Only offered elsewhere" if the library does not offer the service, but you are aware of this service being provided by a department or group outside the library. Choose "Not offered" if this service is not provided at your institution. N=54

Services	Offered by the library	Offered by the library AND elsewhere	Only offered elsewhere	Not offered	N
Data management best practices via website resources & links to relevant literature	34	14	1	5	54
Data storage and backup planning	11	30	11	2	54
Helping researchers identify appropriate metadata standards	42	6	1	4	53
Research file organization and file naming conventions	32	12	2	7	53
Helping researchers apply metadata standards	31	7	1	14	53
Data citation	38	4	0	10	52
Data sharing & access	22	22	4	4	52
Help with securing & anonymizing data per research conduct policies & Institutional Review Board Advising on institutional data polices (e.g., retention, IP ownership)	2	12	27	11	52
Data management best practices via workshops/ direct training	23	12	2	14	51
Other RDM service(s)	2	5	0	0	7
Total Responses	51	37	33	25	54



If you selected Other RDM service(s), please briefly describe the service. N=7

Offered by the library

Conversion of analog image data to digital data.

Helping researchers document their data (not confined to the use of metadata standards.) Also, we do outreach to research groups to initiate conversations about their data management challenges as part of our assessment of needs.

Offered by the library AND elsewhere

Beyond the Libraries, there is at least one data management initiative. The Research Data Repository operates outside of the Libraries but is attempting to address some data management and curation needs in the health sciences. The repository collaborates with the Libraries on design, metadata standards, and other points of interest.

Data publication, data rights management.

The libraries are working with campus partners, especially IT, on data management tools, such as electronic lab notebooks.

We did a survey of researchers about their data management habits and needs; set up a listserv for data management questions for researchers on campus; beginning this month will offer a set of user group meetings exploring various software tools; with our central IT expanding storage available for research data we are beginning a pilot to support long term managed arching of data through our planned digital repository.

We in the library offer information to help researchers assess their RDM needs and refer them to appropriate services on campus that can address those needs. To my knowledge, no other campus organization offers researchers such referral information on its web site, though they may provide it informally/in person.

If the service is offered elsewhere, please identify the department or group that provides the service (e.g., central IT, research administration/sponsored projects office, IRB, health sciences data management, etc.) N=38

Campus cyber-infrastructure organization

Center for Computational Science, Central IT at the School of Marine and Atmospheric Science

Central IT (2 responses)

Central IT, Research and Health Sciences IT

Central IT also offers services for data storage and backups. The research office leads services and support related to IRB and other sensitive or restricted data management.

Central IT, college-level

Central IT, departmental IT, High Performance Computing Center, IRB, Office of Research Integrity Assurance, Information Security

Central IT, human subjects protection, IRB

Central IT, IRB, Office of Research

Central IT, research centers, IRB

Central IT, TACC

Data storage and backup planning also offered by central IT. Help with securing/anonymizing data per policies also offered by central IT information security, policy, and records office.

Data storage and backup planning: Central IT

Data storage is offered by campus IT services as well as the OCUL-Scholars Portal library consortium that is hosting an instance of Dataverse, a data storage repository for Ontario researchers.

Faculty of Nursing (HRDR as described above), Biological Sciences (GIS), Research Ethics Office (compliance, policy)

File organization and file naming: central IT and Libraries (guidance). Data storage and backup planning; central IT (offers storage and backup services); Libraries offers guidance, but not storage and backup services. Secure data and IP issues: Libraries offers guidance, especially with IP issues; SPARCS (Sponsored Programs) offers guidance on security, IRB, compliance issues.

Health Sciences, Office of Sponsored Programs, University Archives

High performance computing and research center

In most cases, services listed above are offered through Research Data Management Service Group, a cross-campus, collaborative organization that the library is a part of. In the case of Data Sharing and Access, the library offers some services directly (such as deposit and sharing in our institutional repository), but others are offered across campus through central IT, the Center for Advanced Computing and the Institute for Social and Economic Research. In the case of application of metadata standards, service is provided by the group the most closely aligns with the research

area requesting help. For example, social science metadata creation is supported by both the Institute for Social and Economic Research and the library, but environmental science metadata creation would be primarily supported by metadata librarians.

Information Services & Technology, Clinical and Transitional Science Institute, Data Coordinating Center

Information Technology Services

Institute for Policy and Social Research, Center for Research Methods and Data Analysis, Central IT for the campus

IRB

Odum; ITS Research Computing; RENCI; DICE. The Carolina Population Center offers anonymization and securing services only to their own fellows; the Sheps Center and other institutes on campus likely have similar services for their own affiliates.

Office of Grants and Contracts Administration, Research Compliance, Institute for Social Science Research

Office of Research Services

Office of Research Services: Help with securing & anonymizing data per research conduct policies & Institutional Review Board, advising on institutional data polices (e.g., retention, IP ownership). UIT: Data storage and backup planning.

Office of Sponsored Research, School of Medicine, College of Engineering

Other academic departments and organizational units

Our research office (office for sponsored programs or OSP) works with researchers also in data sharing requirements. Central and departmental IT also work on storage, file organization and anonymizing of data, and security concerns. They are also partners in our data management best practices workshops.

Research Computing (central IT), IRB

Research computing and office of research

Research Computing includes some materials on their website in collaboration with the Libraries and links to the Libraries. For help with securing and anonymizing data, the different control support units like IRB, the Privacy Office, and others offer online and other training on data security as a matter of conformance and compliance with policies and law, and the Libraries refer people to those groups.

Social Sciences Research Institute (SSRI): Data Services Core; Central IT academic computing services; Health Sciences, data management services

Sponsored Programs

Storage and backup planning is offered to a limited extent by college IT groups and by central computing's RCC/HPC service.

Unit and campus IT groups provide help with data storage and backup, and with securing confidential data. The office of responsible conduct of research addresses IRB issues and also consults on confidential/sensitive data. Technology transfer also consults with researchers on intellectual property and data sharing issues.

DATA ARCHIVING SERVICES

Data archiving provides longer-term retention of and access to research data by others. Archiving involves additional services beyond temporary storage for preserving digital data including assigning of unique, durable identifiers and checking for file corruption.

19. Does your library archive research data? Choose "Only offered elsewhere" if the library does not archive research data, but you are aware of this service being provided by a department or group outside the library. Choose "Not offered" if data archiving is not provided at your institution. N=54

Yes	27	50%
Offered by the library AND elsewhere	13	24%
Only offered elsewhere	5	9%
Not offered	9	17%

If data archiving is offered elsewhere, please identify the department or group that provides the service (e.g., central IT, research administration/sponsored projects office, IRB, health sciences data management, etc.) and explain whether the library is involved with the archive in any way, from recommending it to researchers to assisting them with deposits. N=18

Offered by the library AND elsewhere

Central IT (2 responses)

Central IT and TACC

Central IT provides storage space. A faculty member is piloting an NSF project for a Qualitative Data Repository. The library's repository provides a "front-end" for data set access.

Central IT, Health science IT, Institute for Digital Research and Education.

Colleges and research centers provide domain specific repositories. Library maintains a comprehensive cross-campus service catalog and provides a robust referral service.

Consortially through Scholars Portal

Institution for Social and Policy Studies for their user population and affiliated researchers. The library helps advise and will help reorganize their data archive, but it's run independently.

Many

Research center

The DRYAD repository is based out of the university and we refer researchers to ICPSR and other disciplinary repositories as appropriate. The Odum Institute is a partner in the Dataverse repository network. The library's IR is appropriate for completed projects only but does have secure/embargoed storage; deposit in the IR is largely mediated, with limited self-serve deposit enabled for ETDs and scholarly posters.

Varies

Only offered elsewhere

Center for Computational Science

Central IT offers a storage environment; the Libraries and central IT are currently piloting a dedicated data curation platform (Dataverse network) to experiment with other options.

College of Literature, Sciences and Arts

Colleges and departments

The museum of natural and cultural history maintains its own systems for fossils and other collections, and the library has occasionally consulted with them about some of these data sets.

Various research centres run their own archiving systems. The library is working to offer a central spot for this data, but are waiting on central IT to up capacity.

20. How does your library provide assistance to researchers for archiving research data? Check all that apply. N=54

Assistance locating solutions at existing data repositories for particular research domains (for example, directing sociology students to the ICPSR repository)	52	96%
Direct assistance with depositing data at the library's or institution's data archive	39	72%
Direct assistance with depositing data at a domain repository	26	48%
Assistance depositing data to a journal publisher	21	39%
Other data archiving assistance	4	7%

Please briefly describe the other type of data archiving assistance. N=5

Dataverse

Deposit into Chronopolis, a campus-managed service.

We are building a solution for data only.

We are currently developing infrastructure and more nuanced services to support these needs. We are also developing several cross-institutional partnerships to provide data preservation infrastructure.

If you answered Yes or Library and elsewhere above, skip to the section Data Archive Characteristics. If you answered Only offered elsewhere or Not offered, skip to the section on Staffing: Organizational Structure.

DATA ARCHIVE CHARACTERISTICS

We would appreciate more details about the characteristics of your current research data archiving solution.

21. Please specify your library's archiving solution. Check all that apply. N=40

Our institutional repository is used to archive research data (and used for text documents)	32	80%
The library hosts or is directly involved with archiving research data at an external archive or repository service, such as DuraSpace or Dataverse	17	43%
The library operates a research data archive dedicated to data deposits and access	13	33%
The library is involved with data archiving that is primarily operated by another department/group at our institution	6	15%

[Editor's note: After a further analysis of respondents' answers and a review of archive websites, the survey authors assigned these 40 responses to two categories: IRs with Data (35 respondents) and Data Archives (5 respondents). Those categories are represented in questions 22 through 31 below.]

22. Which platform(s) and/or software are you using for your archiving solution (e.g., DSpace)? N=38

BePress Digital Commons (2 responses)

BePress Digital Commons, DuraCloud; campus is investigating DSpace and Fedora Commons.

BePress Digital Commons. We are developing and will soon transition to a Fedora managed data repository that will contain research data but many other kinds of data and objects as well.

Chronopolis

Custom repository built with open source tools including Python, Django, Soir, Lucene, Jquery, ubuntu, Celery, PostgreSQL and WordPress.

DAITSS

Dataverse

Drupal

DSpace (11 responses)

DSpace (with a planned conversion to Hydra/Fedora)

DSpace (for the institutional repository), Dataverse (for an archive external to library)

DSpace, Dark Archive

DSpace, Dataverse

Fedora Commons (3 responses)

Fedora, Hydra (2 responses)

Fedora, iRODS to manage storage and preservation functions

Fedora, Data Conservancy software

Fedora, Merritt

Fedora, Archivematica, SDA (Survey Documentation and Analysis), file servers

Fedora, Archivematica, Dataverse, iRODS, and other software are currently in use or being evaluated for use.

HUBZero

Meta Archive Cooperative

Self developed software

SobekCM

Archive Platform Type	IRs with Data (N=35)	Data Archives (N=5)	Total (N=40)
DSpace	17	0	17
Fedora	8	2	10
BePress Digital Commons	4	0	4
Dataverse	2	1	3
Archivematica	1	1	2
Hydra	2	0	2
iRODS	1	1	2
Custom repository	1	1	2
Drupal	1	0	1
Meta Archive Cooperative	1	0	1
DuraCloud	1	0	1
SobekCM	1	0	1
DAITSS	1	0	1
Merrit	1	0	1
HUBZero	1	0	1
Chronopolis	0	1	1
Data Conservancy	0	1	1

23. How many researchers currently have data deposits in your archive? An estimate is acceptable. N=33

Archive Type	N	Minimum	Maximum	Mean	Median	Std Dev
IRs with Data	28	0	1000+	90.79	10	258.93
Data Archives	4	2	100	31.00	11	46.78

IR with Data Comment

Hard to say. We have hundreds of datasets that have been submitted along with articles for publication. Additionally we

have the research data sets from a handful of humanities institutes on campus, which include the work of many faculty and student researchers.

24. How does the library pay for data archiving? N=37

Funding Method	IRs with Data (N=32)	Data Archives (N=5)	Total (N=37)	% of Total
Absorb the cost	30	1	31	84%
Grant funded	7	2	9	24%
Charge to researchers	2	3	5	14%
Charge to users	0	0	0	0%
Other method	4	3	7	19%

Please briefly describe the other method.

IRs with Data

Currently the library absorbs the cost. In the pilot phase of the managed repository we will also absorb the costs. When we operationalize the repository we will have a service charge for users.

Depending on the material, with most data sets being small and typical of any item in the Institutional Repository and Digital Collections, the cost is part of the standard cost of operations. When there are grant funds available, the libraries are included for the work to support stronger overall growth and the inclusion of library materials to the larger research project as born-digital and digitized materials. This is for the curation of materials from the libraries' holdings and/or partner holdings with those materials part of the larger research project. For other cases, if funds are needed, the libraries work with researchers to find funds from appropriate sources.

For DSpace, the Libraries absorb the cost up to 2 GB; for Dataverse, the two-institution partnership is funded by academic departments.

Paid for consortially.

Data Archives

Partnerships, in-kind contributions to development. We are looking at cost models for our different categories of projects, services and partnership initiatives.

The first 18 months for developing the repository were provided by the university after submitting a proposal and budget.

To be determined.

Please briefly describe any details of the cost model that you'd be willing to share (e.g., amount and how it is applied, is the charge cost recovery?) N=9

IRs with Data

Five-year pilot project of Scholars Portal. No charges for members of Ontario Council of University Libraries.

One-time charge model based on the data set size, per file.

We are developing a business model that will grant researchers a fixed amount of storage space (e.g., 5 GB) for free. After that, projects/researchers will be charged on a per meg basis. We do not yet have a figure for the per meg cost.

We have a limit on the size of files we can accept and we're grappling with how long we can keep deposits. Data archiving is so much more complex than other kinds of digital curation that we are just now testing what can be archived, but we planned from the start that our IR would include data sets. We would like to begin charging for long-term preservation of data at some point, but our library is not designated as a chargeback center on campus, so we are currently unable to charge at this time.

We plan cost recovery by establishing ourselves as a cost center, per OMB A-21.

Data Archives

Services are based on cost recovery. The funding mechanisms vary depending on who the data owners are.

The cost is 2% of total direct cost of grant, though if someone has data from a completed project that they would like to archive, we would explore funding models with them to pay for archiving that data.

Too early to share costs. We are still working towards understanding our costs for both our in-house services as well as several of the partnership initiatives that we are involved with.

Under review.

25. What are the current or anticipated sources for the data in your data archive? Check all that apply. N=40

Source of Data	IRs with Data (N=35)	Data Archives (N=5)	Total (N=40)	% of Total
Data directly associated with a particular publication	30	5	35	88%
Data associated with a full research project (e.g., all processed data for an NSF grant)	29	5	34	85%
Data from graduate research, dissertations, or theses	30	2	32	80%
Data moved from another archive to the library data archive or IR	22	3	25	63%
Other data source	5	1	6	15%

Please briefly describe any other data source(s). N=6

IRs with Data

Data from long-term faculty research archives. This is especially the case with humanities researchers who travel around the world, sharing often the only other copies of primary materials. These archives were once on paper and microfilm and often eventually found homes in library special collections. We are working with the same for digital materials and

making sure that transfer process can happen before researchers retire, to have the data access and archiving as part of their workflow so that they, their field, and the world benefit from their research.

Directly from faculty, Archives and Special Collections.

For Dataverse, the data include both primary data and data compiled from secondary sources.

Raw data, active storage.

We are doing outreach to near retirement age faculty to get them to review their research history and identify data sets, grey lit, and other files that should be ingested in a repository (or digitized and then ingested).

Data Archives

Many other sources, e.g., campus planners have indicated interest in depositing their data with us. Interest also from archives and other agencies, e.g., government.

26. What is the total overall size, in gigabytes, of research data sets currently deposited to the data archive? An estimate is acceptable. N=33

Size	Minimum	Maximum	Mean	Median	Std Dev	N
Gigabytes	.009	500	38.37	5	107.12	22
Terabytes	1	75	15.61	2	29.42	11

27. Are there limits to the amount or size of data deposited for a given project or researcher? N=39

Limits	IRs with Data (N=34)	Data Archives (N=5)	Total (N=39)	% of Total
Yes	15	3	18	46%
No	19	2	21	54%

If yes, please briefly describe the limits. N=17

IRs with Data

<200 MB (per file) as determined by vendor.

1 TB

2 GB per project.

2 GB. The http protocol limit of 2 GB often serves as a ceiling to what is feasible to deposit in either DSpace or Dataverse.

50 MB and then must ask permission.

If researcher expects to have project files (data and other files) totally more than 500 GB, we ask them to talk to us to devise a custom solution.

Individual files deposited in eCommons should normally be less than 1 GB in size. The total size of a collection of files associated with a research project and deposited into eCommons should not exceed 10 GB per year. In many ways this is merely a functional limit as large files are very hard to handle/download with our current infrastructure. In the future these limitations may be revised if usage or technology dictates a change.

No total limit, but individual files can't be larger than 700–800MB because of problems with uploading and downloading through DSpace.

There is no hard limit but if there is a large data set, we check with Scholars Portal.

Upload limits of 1GB per file (repository limitation).

We accept up to 2 TB of data per project, but we will consider exceptions to this rule.

We are still working on establishing appropriate limits.

We try to keep videos to no larger than 4GB. We have negotiated coverage of 1 TB of collections for a fee.

We're not sure at this point of the limits. When uploading if there are issues, we can have the vendor load them for us.

Data Archives

2 TB per project, but a researcher could pay for additional storage.

Currently 50MB for data associated with a project is not part of a grant. 100G for data associated with projects that are supported by a grant.

We currently mediate deposits of data so the limits—if there were any—would be based on our capacity limits for ingest as they relate to storage, personnel time, etc.

IRs with Data Answered No

Have ability to override limits.

There is a limit of 100 MB per file per upload, but no limit per project or researcher.

28. Which of the following options apply for how researchers deposit data into your data archiving solutions? Check all that apply. N=40

Data Deposit Options	IRs with Data (N=35)	Data Archives (N=5)	Total (N=40)	% of Total
The library deposits data collections for the researcher	30	5	35	88%
The library provides assistance when researchers deposit data	27	5	32	80%
Researchers can self-deposit their data collections without direct assistance	23	3	26	65%

Comments N=7

IRs with Data

For DSpace, some staff involvement is required to set up an account. For Dataverse, no staff involvement is required.

If we create the metadata and do the ingest or create a batch ingest solution, we charge a fee.

Most of our self-deposit transactions are mediated in some way.

Self-deposits are always mediated.

We prefer that they self-deposit, but we sometimes end up either providing them with assistance or depositing the files for them.

Data Archives

All data submissions are reviewed by the librarian before being accepted.

We are working towards a suite of services that provides very basic, self-directed RDM and deposit (Dataverse) and more robust services for doing more comprehensive data management and curation.

29. Does your data archiving solution provide a persistent identifier for deposited datasets? N=39

Yes	38	97%
No	1	3%

If yes, which identifier type is used? Check all that apply. N=38

Identifier	IRs with Data (N=33)	Data Archives (N=5)	Total (N=38)	% of Total
Handle System (handle.net)	21	1	22	58%
DOI	7	4	11	29%
ARKs	5	3	8	21%
Other identifier	10	1	11	29%

Please specify the other identifier(s). N=11

IRs with Data

EZID

Fedora PID URI

Locally created UUID [Universally Unique Identifier]

Locally generated, but will be implementing DOI and possible ARKs in coming year.

Locally generated Persistent Identifier (PID)

Persistent URL (2 responses)

Resource Object ID

This will happen in our planned managed digital repository.

UUID

Data Archives

Self-generated

30. Does your data archiving solution have any of the following digital preservation capabilities? Check all that apply. N=37

Preservation Capability	IRs with Data (N=32)	Data Archives (N=5)	Total (N=37)	% of Total
Integrity checking	28	5	33	89%
Replication for more than one copy	27	5	32	87%
Replication with geographic separation to protect copies	22	4	26	70%
Format migration/conversion	16	3	19	51%
Other preservation capability	9	1	10	27%

Please briefly describe the other preservation capabilities. N=11

IRs with Data

Academic Preservation Trust

Audit trail

Checksum (fixity)

Ditto [integrity checking and format migration], planned for managed digital repository.

DSpace: Integrity checking, replication for more than one copy, and replication with geographic separation. Dataverse: Replication for more than one copy, replication with geographic separation, and format migration/conversion.

Format migration is manual on upload and only if required.

Normalization of formats, characterization of the files

Periodic virus scanning

Preservation actions are limited to a) quarterly dark archive and b) participation in LOCKSS PLN (forthcoming).

We do [all but format migration] more or less depending on the project.

Data Archives

We have several systems that we use for digital preservation. Format identification, fixity checks, and other common DP actions are applied.

31. Is the primary purpose of the data archive for open access (e.g., unrestricted access for online users) or controlled access (e.g., admin or researcher approval required before accessing data) or another purpose? N=39

Archive Purpose	IRs with Data (N=34)	Data Archives (N=5)	Total (N=39)	% of Total
Open access	30	3	33	85%
Controlled access	7	2	9	23%
Another purpose	4	2	6	15%

If another purpose, please briefly describe the conditions for data access and use. N=6

IRs with Data

Archive service for faculty.

Both, as well as programmatic access to data for digital projects, feeds, etc.

The archive is for access and preservation. The two are problematic to decouple within the data curation lifecycle.

We allow both open access and controlled access depending on the needs of the researcher and the nature of the data.

Data Archives

The emphasis would be on sensible access, open or controlled. Proper data management for long-term access and reuse would be of primary importance.

We accept data according to the data owners' mandates. The archive itself is dark and not an access tool.

STAFFING: ORGANIZATIONAL STRUCTURE

32. For the personnel in your library who currently provide RDM services as all or part of their responsibilities, please indicate which of the following best describes their organizational structure. N=53

A committee/group of staff from two or more departments within the library	27	51%
A committee/group of staff from the library and other departments in the institution	9	17%
A single position within the library	8	15%
A single department within the library	6	11%
Other organizational structure	3	6%

LIBRARY COMMITTEE/GROUP

33. Please enter the name of the library committee/group of staff that provides RDM services, the names of the departments, and the number of committee/group members. N=27

Committee Name	Department Names	Members
Ad hoc RDM Working Group		8
Center for Digital Scholarship	Digital Technology, Research and Outreach Services	9
Data Archive and Curation Services Team	Center for Southwest Research, Outreach, Maps	6
Data Curation Committee	Archives, Technical Services, public services, IT, Institutional Repository	7
Data Initiatives	Center for Digital Scholarship; Collection Development; Integrated Technology	5
Data Management Advisory Group	Digital Scholarship Center; Science Library; Document Center	4
Data Management Committee	Research and Instructional Services, Digital Repository Services, Resource Description and Management, University Archives	10
Data Working Group	Systems, Science and Engineering Library, Reference & Liaison Services, Scholarly Communication, Special Collections and University Archives	8
Data Working Group	Learning & Research Support, Digital Publishing, Administration	4
Digital Case Committee		6
Digital Services Operations Team	Collection Development, IT, Cataloging,	6
E-Sciences Working Group	Education & Outreach, RSMAS Library, Cataloging & Metadata Services	8
Library Data Curation Service Team	IT, Digital Library, Science and Engineering Library, Cataloging and Metadata	10
No formal name: Research Data Services	Technology, Scholarly Communications, metadata service	12
No name; team approach	Publishing & Curation Services, Reference Collections & Research, Digital Library Technologies, Metadata and Cataloging Services	3
Research Data Services	Program Management Center, Research, Collections and Scholarly Communication, Cataloging	7
Research Data Committee	Collection Management, Digital Library Initiatives, Research and Information Service, Collection Development and Special Collections, Centennial Campus Reference Services, Administration	7
Research Data Management Committee	Bibliographic Services, Map Library, Reference, Science Library, University Librarian's Office.	7
Research Data Management Working Group	Services Division, Electronic Data Center	11

Committee Name	Department Names	Members
Research Data Management Guidance	Digital Curation, Data Services, Systems, Digital and Multimedia Center	7
Research Data Services	Information Technology, Public Services	4
Research Data Services working group	Specialized Content & Services, and Curation & Preservation Services	5 (plus 2 advisers)
Research Lifecycle	Publishing; Science, Engineering, and Data; Collections; Health Sciences; and Learning and Teaching	15–20
Science Librarians	Reference, technical services, health sciences	4
	Digital Initiatives	15
	Desktop Network Services, Science & Technology Department	2
	Reference, branch libraries	8

LIBRARY AND INSTITUTION COMMITTEE/GROUP

34. Please enter the name of the committee/group of staff from the library and other departments in the institution that provides RDM services, the names of the departments, and the number of committee members. N=9

Committee Name	Department Names	Members
Data Curation Workgroup	Library Technologies, Digital Scholarship & Technical Services, Campus IT	6
Data Management Planning Group	Libraries, Information Services & Technology	5
e-science working group	Research & Information Services, Branch & Off Campus Services, IT, College of Arts and Sciences IT, Bibliographic services, Center for Scholarly Communication & Digital Curation, Library IT	8
eScience Planning Group	Center for Science and Social Science Information, Information Technology Services, University Library	6
eScience Team	Library, Office for Sponsored Programs, University IT, Health Center Library	8
Research Data Curation Program	Research Data Curation Program	5
Research Data Management Service Group	Center for Advanced Computing, IT, Institute for Social and Economic Research, Medical College, Astronomy Department	13 consultants, 7 member of a management council, 10 additional member of a faculty advisory board
Research Data Services	Libraries, Research Computing (central IT), Office of the Vice Chancellor for Research	5
Research Data Services	Libraries, Central IT/CIO	20–25 individuals

SINGLE LIBRARY POSITION

35. Please enter the name of the department that the position that provides RDM services belongs to. N=8

Center for Digital Research and Scholarship

Digital Library Services

Information Technology

Physical Sciences and Engineering Library

Scholarly Communication

Scholarly Communications and Digital Curation

Scholarly Publishing and Data Management

Systems Office

SINGLE LIBRARY DEPARTMENT

36. Please enter the name of the library department that provides RDM services, the number of individuals in the department, and the FTE (e.g., Individuals: 3, FTE: 3 or Individuals: 3, FTE: 2.5). N=6

Department	Individuals	Total FTE
Center for Digital Research and Scholarship	8	3
Data & GIS Services	5	4
Data Management Consulting Group	5	3.6
Digital Access Services	2	2
Data Management Services	6	5
Research and Data Services	9	7.45

OTHER ORGANIZATIONAL STRUCTURE

37. Please briefly describe the other organizational structure. N=3

All of the liaison librarians, curators, and archivists provide some level of support for research data management. There is a current task force working to extend the support for all involved, and for all of our researchers as well, but the organizational structure is that this is a core part of work and we all support this work.

Currently provided by representatives from two different departments. We are in the process of hiring a data curation librarian who will coordinate these efforts and develop a more robust suite of services.

There is no committee, three people responsible in varying ways. I work at the policy level, another librarian works with GIS data, and a library assistant trains students on software packages.

STAFFING: POSITION DETAILS

Please list the position(s) of the person(s) who are responsible for RDM services as part or all of their job. These may include library professional staff, interns, students, or non-library staff positions that are part of your library's core RDM services, or are established partners in your core services (i.e., not a research institute data manager unaffiliated with the library's services).

If you are reporting on multiple positions that have varying levels of RDM responsibility, please start with the position that has the most responsibility and work down. We acknowledge that some of these roles may be in flux or not fully implemented, so for the following questions, please answer with what you anticipate will be the most accurate configuration of personnel in 2013.

38. Please enter the position title for up to six individuals who provide RDM services. Use official job titles when possible, or "intern," "volunteer," etc. N=54 respondents, 231 positions

Single Library Position N=8

Data Curation Librarian

Digital Repository Coordinator

Digital Services Librarian

Director of Research Systems Development

Research Data Librarian

Research Data Manager

Research Services Librarian

Science Data Management Librarian

Two Positions N=4

Digital Repository Librarian

Metadata Librarian, Head of Digital Access Services

Research Data & Metadata Librarian

Research Data Specialist (Research Computing)

Science Data Management Librarian

CLIR Data Curation Fellow

Web Support Librarian, Desktop Network Services

Science & Technology Librarian

Three Positions N=2

Digital Curation Librarian

Data Services Librarian

Metadata Librarian

Faculty, Digital Repositories Librarian

Faculty, Digital Technologies Development Librarian

Faculty, Data Services and Informatics Librarian

Four Positions N=9

(e-)Science Librarian II

Science Librarian II

Digital Publishing Librarian I

Library Director

e-science librarian

Head, Digital Collections and Scholarly Communication Services

Biosciences Librarian

Associate Director for Academic and Research Technologies, Research Computing Services

Engineering Librarian

Health Sciences Librarian

Assistant Head of Reference

Knowledge Management Specialist

eScholarship, ePublishing & Digitization Coordinator

Data Management Librarian (to be hired in 2013)

Liaison librarian

Library assistant

Head of the RSMAS Library

E&O Librarian for Sciences

Metadata Librarian

RSMAS Librarian

Head, ScholarSphere User Services

Metadata Librarian

Data Management Specialist

Subject/liaison librarian (multiple librarians, depends upon subject area)

Post Doctorate for Biology

Science and Health Librarians

Metadata Specialist

Associate University Librarian

Research Data Librarian

Assistant Dean for Digital Humanities Research

Engineering/Research Data Librarian

Data Curation Assistant (Student Position)

Research Liaison Librarians

Data Librarian

GIS Librarian

Scholarly Communications Librarian

Five Positions N=7

CLIR Data Curation Intern

CLIR Data Curation Intern

Head of Science, Engineering, and Data Team

Head of Institutional Repository

Physics and Astronomy Librarian

DACS Committee Chair

Data Curation Librarian

Data Librarian for Economics and Business

Research Librarian for Engineering

Fine Arts Librarian

Head, Data and GIS Services

GIS Librarian

Data Visualization Coordinator

Systems Librarian

Data Services Intern

Librarian

Librarian

Associate Director, Scientific Computing and Visualization Group

Institutional Repository Librarian

Assistant Director of Library Computing & Systems

Lorry I. Lokey Science Data Services Librarian

Social Science Data & Government Documents Librarian

Head, Digital Scholarship Center

Coordinator, Institutional Repository

Map/GIS Librarian

Research Data Curation Program Director

RDC Technical Manager

Faculty Liaison Project Manager

Faculty Liaison

Metadata analyst/consultant

Scientific Data Curator

Social Sciences Data Librarian

Digital Humanities Librarian

Digital Repository Manager

Head, Digital Production Services

Six Positions N=23

Associate Dean

Data Curation Librarian (search in progress)

Graduate research assistant / intern

Development Programmer

Digital Production and Publishing Manager

Head, Digital Initiatives

Associate Dean for Library Technologies

Senior Associate Dean for Digital Scholarship and Technical Services

Director of Digital Scholarship

Digital Library Architect

Head of Metadata and Cataloging

Information Technology Enterprise Architect

Associate Head, Collection Management, and Director of Research Data Services

Head, Digital Library Initiatives and Digital Projects

Data Services Librarian

Director, Copyright and Digital Scholarship

Digital Collections and Preservation Librarian

Associate Head, Research and Graduate Studies

Data Library Coordinator

GIS Librarian

Digital Repository Services Librarian

Digital Initiatives Coordinator

Digital Preservation Officer

Metadata Librarian(s)

Data Management Specialist

e-Science Librarian/Postdoctoral Fellow

Coordinator of Geospatial Data Services

Coordinator of Numeric Data Services

Copyright & Scholarly Communications Librarian

Multiple other members

Digital Assets Librarian

Map Librarian

Data Librarian

Business Librarian

Science Librarian

Social Sciences Librarian

Digital Humanities Librarian

Reference & Liaison Librarian in the Health Science Center

Agricultural Sciences Librarian

Political Science Librarian

GIS Librarian

University Archivist and Head of Records Management

Digital Scholarship Librarian

GIS Specialist

Digital Librarian

Research Services Librarian(s)

Team Leader: Digital Learning & Scholarship

Team Leader: Acquisitions & Metadata Services

Digital Strategies Coordinator

Reference & Liaison Librarian(s -- 2 science, including Head of Science & Engineering Library, and 1 social science)

Head, Systems and Web Management

Digital Repository Resident

Science Librarian for the Center for Hierarchical Manufacturing

University Archivist

Head of Maps & Data Initiatives

GIS Specialist

Technology Specialist

Statistical Computing Specialist

eSciences Librarian

Social Sciences Librarian

Head of reference

Director of collection development

Life sciences librarian

Physics/astronomy/chemistry/mathematics librarian

Engineering librarian

Geology librarian

Head of Specialized Content & Services

Social Science Data Services Librarian

Bioinformatics and Biosciences Librarian

Electrical Engineering and Computer Science Librarian

Head of Curation & Preservation Services

Digital Curation and Preservation Services Library Fellow

Head of Strategic Data Initiatives

Data Consultant

Health Sciences Data Consultant

Senior Data Consultant

Data Consultant

Intern

Interdisciplinary Research Librarian

Data Services Specialist (2)

Digital Data Repository Specialist

Software Developer (2)

Metadata Specialist

Digital Archivist

Librarian

Librarian

Academic Technologist

Academic Technologist

Information Architect

Information Architect

Librarian for Numeric Data Services and Data Management

(Health Sciences Library) Bioinformatics and Translational Science Librarian

(Health Sciences Library) Cataloging and Metadata Coordinator

Head, Kenan Science Library

GIS Librarian

(IR) Information Infrastructure Architect

Manager

Data Management Consultant

Data Management Consultant

Systems Administrator

Software Developer

Technical Consultant

Numeric Data Services Librarian

Scholarly Communication Librarian

GIS Librarian

Metadata/Cataloging Librarian

Research & Assessment Analyst

Director, Program Management Center

Research Data Librarian

Digital Projects Specialist

Metadata and Digital Resources Librarian

Acting Head, Collection Development

Research Librarian

Director, Information Technology

Research Data Management Service Coordinator, Science Data and Metadata Librarian

Consultant, Research Data and Life sciences Librarian

Consultant, Life sciences Librarian

Consultant, Physics and Astronomy Librarian

Consultant, Research Associate at Cornell Institute for Social and Economic Research

Consultant, Research Associate in Astronomy

Science and Social Science Data Librarian

Science and Social Science Data Librarian

Statistical Consultant

StatLab Manager

Director of Academic IT Solutions

Director of Science and Social Science Libraries

Science Librarian Team Leader and Librarian for Biology and Agriculture

Preservation and Data Management Services Librarian

Cataloging and Metadata Librarian

Librarian for Psychology, Sociology and Communications

Engineering Librarian

Collection Development Librarian (UC Health Center)

Software Engineer Lead

Software Engineer

Digital Projects Manager

Geospatial and Research Data Manager

Collections & Scholarly Communications Librarian

Metadata Librarian

Other

All liaison librarians are responsible for data management advising.

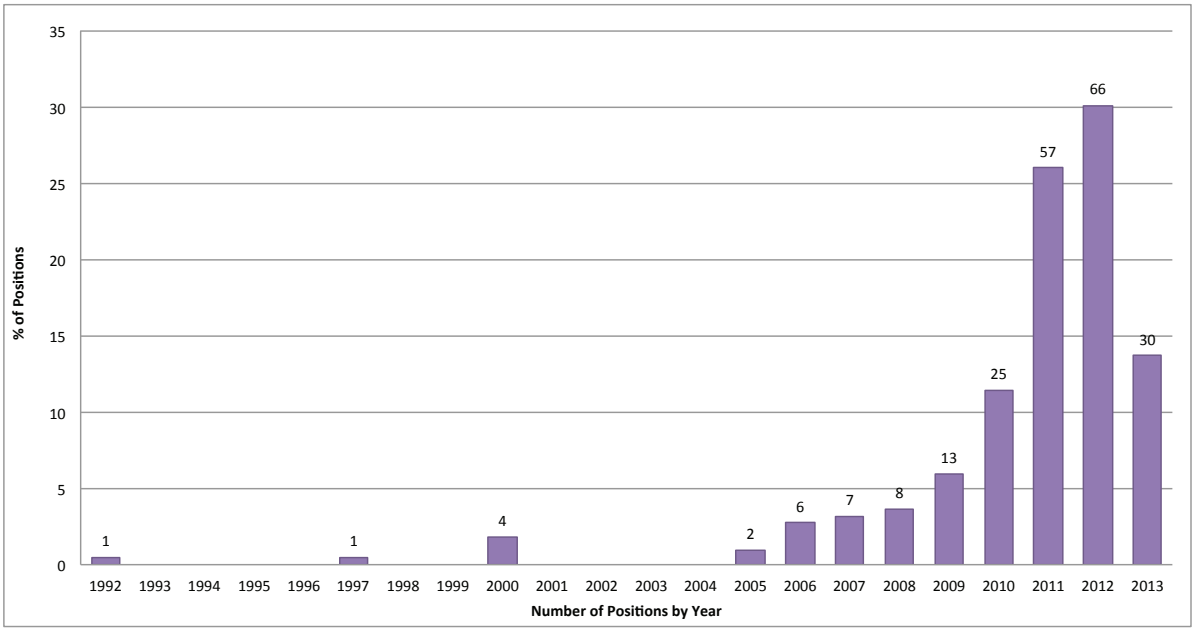
39. For each position, please indicate if it was created specifically for RDM services, substantially redesigned to focus on RDM, or had RDM activities added to existing responsibilities. N=53 respondents, 230 positions

RDM activities added	146	63%
New RDM services position	49	22%
Substantially redesigned	34	15%

Comment

Position also has co-director responsibilities of the institutional repository.

40. Please enter the year that the person in each position began providing RDM services. N=52 respondents, 220 positions



41. Please indicate whether each position is full-time permanent, part-time permanent, or temporary (e.g., grant-funded, internship, etc.) N=52 respondents, 224 positions

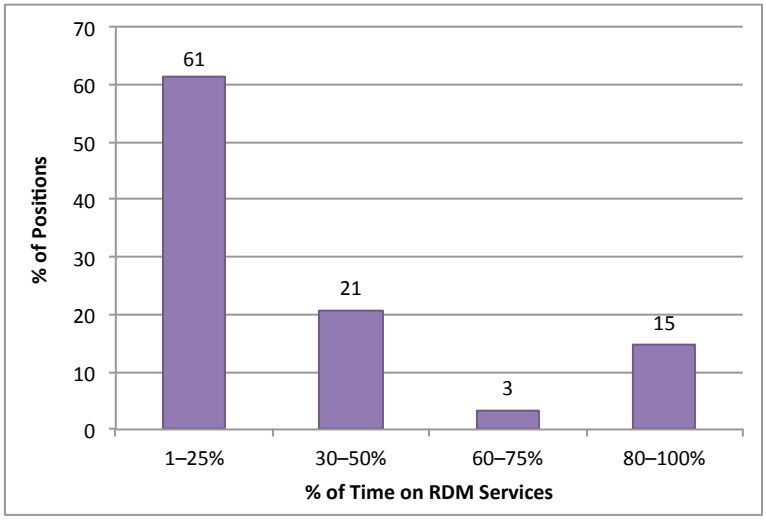
Permanent or Temporary	N	Positions	% of Positions
Full-time (permanent)	51	203	90%
Part-time (permanent)	4	8	4%
Temporary (full-time or part-time)	11	13	6%

42. Which RDM service roles apply to each position? Check all that apply. N=53 respondents, 222 positions

RDM Service Roles	N	Positions	% of Positions
RDM guidance (other than DMPs)	51	183	82%
Data archiving assistance	49	154	69%
Grant proposal DMP support	49	150	68%

43. What is the approximate percentage of time each person spends on RDM services? Enter a whole number without a % sign. N=51 respondents, 213 positions

Minimum	Maximum	Mean	Median	Std Dev
1	100	31.99	20.00	31.73



% of Time	1–25	30–50	60–75	80–100
# of Positions	130	44	7	32

44. Which non-RDMS tasks, if any, are a significant portion (more than 50%) of each position's job description/time allocation? Check all that apply. N=49 respondents, 194 positions

Non-RDMS Tasks	N	Positions	% of Positions
Subject reference services	34	81	42%
Cataloging/ collection development	28	52	27%
Administrative	30	48	25%
Other data services (e.g., GIS training)	25	47	24%
Library IT, systems admin	21	30	15%
Metadata librarian	19	28	14%
Other tasks	28	75	39%

Please briefly describe any other non-RDM tasks. N=27

Task Category	N	%
Repository-Related Work	10	37%
Teaching	7	26%
Copyright and Scholarly Communications	6	22%
Management	4	15%
Digital Collections or Project	3	11%

All librarian members of the group perform reference, liaison, and collection development tasks, as well as other projects.

Analog Media preservation.

Content management, work on research, instructional activities, and professional service.

Copyright and digital scholarship guidance.

Copyright and Scholarly Communication

Copyright checking and consultation; training and workshop presentations on RDM; outreach and marketing; participation on grant funded research projects.

Digital projects, collection development

Digital Repository Coordination

Head of Specialized Content & Services and Head of Curation and Preservation Services are department heads who oversee a portfolio of services in their areas and also have various administrative responsibilities. Specialized Content & Services includes GIS, Social Science Data Services, Statistical Software Support Services, the Aga Khan Documentation Center, Digital Image & Visual Collection Services, Distance Education Video and Streaming Services, and Community Video Sharing Services, in addition to Research Data Management Services. The librarians have liaison roles that include collection selection, soliciting open-access articles, reference and instruction responsibilities. The Library Fellow is involved with digital curation and preservation activities.

Impossible to do briefly. Serials and other collection review projects; weeding; special IT projects such as developing more robust infrastructure for the IR; professional development; special cataloging projects; administrative projects like formulating the library's strategic plan and developing a patron-driven acquisitions project; classroom instruction; etc.

Liaison activities overall, which include many of these elements but also extend into other work, and there is much overlaps with RDM tasks.

Leadership and coordination for research services overall (publishing support, impact metrics, etc.) and IR services (as co-director).

Library instruction and faculty liaison work.

Manager of digital repository.

Other digital project related work.

Other tasks could include too many activities to detail here.

Position 1 and 2: Repository Development. Position 3: Data Literacy.

Position 1: supervision, management, budgeting, planning, etc.

Position 2 is a joint appointment between the Libraries and the university's digital humanities center. Position 3 also provides information literacy instruction.

Position 5: Copyright and intellectual property related tasks, primarily related to scholarly publications (not datasets) for institutional repository.

Position 6 works with archive system development.

Position 4 is primarily responsible for scholarly communications and the institutional repository.

Project management, supervision, outreach, committee work, etc.

Repository Administration.

Research-oriented work towards building repository tools to manage data.

Statistical analysis.

Teaching, scholarly communication; digital repository development.

45. For each position, indicate the degree(s) that the individual holds. Check all that apply. N=53 respondents, 229 positions

Degree	N	Positions	% of Positions
MLS/MLIS	52	172	75%
Masters in a discipline other than Library/IS	33	64	28%
PhD in a discipline other than Library/IS	21	29	13%
Data curation emphasis (w/MLS or other degree)	12	14	6%
Archives emphasis (with MLS or other degree)	7	8	4%
Non-graduate degree only	7	10	4%
Other degree	5	5	2%

Please specify the other degree. N=4

For position 4, this could refer to any of our subject librarians, so it's impossible to say what other degrees might be included.

JD (2 responses)

PhD in Information Science

Non-graduate Degrees

Bachelor in computer and information technology

BA, economics

BS, Biochemistry

Geography, English, Curriculum Design

If you selected Masters or PhD in a discipline other than Library/IS, please indicate the disciplines for the graduate degree. N=29 respondents, 73 positions

Fields of Study

Social and Behavioral Sciences	17
Life Sciences	16
Arts and Humanities	12
Physical Sciences & Mathematics	8
Engineering	7
Other	13

1) Biology, 2) Geography, 3) Natural Resources, 4) Law, 5) German Literature and Humanities Computing, 6) Entomology

At least a Masters in a science or quantitative social science discipline will be required (this position has not been hired yet).

Biology

Biology (MS), Ecology and Evolutionary Biology (MS), Botany (MS), Demography (MA), Theoretical Physics (MSc and PhD), Development Sociology (MS and PhD)

Biology, Bioinformatics, Computer Engineering

Computer and Information Sciences

Computer Science; Religion & Literature

Computing Science, Political Science

Digital Humanities

Earth sciences

Engineering

English/Digital Humanities; Political Science; GIS

Environmental Science

Geography

History

Masters in Instructional Technology

MS Industrial and Systems Engineering, MS Geology, PhD Archives Studies, Certificate of Advanced Study in Library and Archives Conservation

Neuroscience, Geography, Political Science, Anthropology, Genetics, Public Health

PhD anthropology/archaeology; MS Geographic Information Systems

PhD in Mass Communication, MA in Social Sciences, MS in Engineering (2)

Philosophy, Geosciences, Plant and Soil Science, Higher Education

Political Science and Biochemistry

Position 1 has MA in English Literature. Position 3 has PhD in Computer Science and Engineering

Position 4 has MA in Italian Studies, PhD in Humanities Computing

Position 1: MBA, Position 2: PhD in sociology, Position 3: MS in environmental science, Position 4: Germanic Languages

Position 1: MS in Management of Information Technology, Position 5: MA in Astronomy, PhD in Astronomy

Position 1: Educational Technology, Position 2: Botany and Plant Pathology

Psychology, Neurobiology, English

Russian Language & Literature

Russian; Engineering

STAFFING: SKILLS & TRAINING NEEDS

Providing the RDM services addressed in this survey requires a broad range of skills, some of which may not be well developed in research library staff yet. These questions are intended to help us identify the most important training needs.

46. Please pick up to three of the training/experience categories listed below that are the most important for the staff at your library who provide RDM services. N=51

Subject domain expertise	38	75%
Digital/data curation training	31	61%
IT technology or services experience	30	59%
Library MLS/MLIS training	24	47%
Traditional archives training	3	6%
Other training or experience	12	24%

Please briefly describe any other training or experience that is particularly important. N=12

Research Methods and Data Analysis

A degree of proficiency in discipline-specific aspects of data acquisition, wrangling, analysis, interpretation, and visualization. Understanding the role of data in scholarly communication. (also Scholarly Communication)

Data mining; database creation; data analysis

Research methods and data analysis

Research Data Management

Research data management practices

Self-education in data management and big data

Data Curation

Data curation

Data curation, data management, data analysis (also research methods and data analysis)

Metadata

Scholarly Communication

Copyright, how faculty actually conduct research (and how a research University actually runs- priorities) data formats and overall RDM best practices. (also Research Methods and Data Analysis and Research Data Management)

Policy/copyright/IP management

Scholarly communications training and experience

Other

The librarians who are liaisons to departments, labs and centers have knowledge of the research activities of the areas to which they are liaisons and experience supporting researchers with data management.

47. Please indicate up to three skills that your RDM staff most need additional training on. N=51

Identifying and applying appropriate metadata standards	25	49%
Digital preservation	17	33%
Data ownership policies	16	31%
Ethical and legal issues	15	29%
Subject domain expertise	15	29%
Data security	12	24%
Data sharing & access	11	22%
Data storage and backup planning	10	20%
Understanding funder requirements for data management plans	8	16%
Data retention policy	7	14%
File organization and naming conventions	6	12%
Data citation	4	8%
Other skill	8	16%

Please briefly describe any other skill that your RDM staff most need additional training on. N=8

A degree of proficiency in applications used for data acquisition, wrangling, analysis, interpretation, and visualization.

Because we are in the process of assessing data service needs, we are not at the point of defining more services and/or identifying additional skills for library staff. These activities will follow later this year and next.

Deeper knowledge of research administration practices and forces (i.e., sponsored research offices, research administration offices, higher education research business model, etc.)

Funder requirements; and available services and tools for research data management (on-campus and elsewhere)

Research methods and skills

Selected top three, but all skills mentioned generate interest among our staff, and seem pertinent at different stages in the development of our RDM services.

Systems Administration

The university has a robust repository in place with SobekCM, so the technical aspects and the related work for digital preservation are already supported. Liaison librarians already have expertise in providing data reference, liaising with their departments, and more. More training in various areas will be needed for all of the librarians based on their individual skills and specific liaison group needs. What is expected to be needed are instances of team-teaching with faculty across campus to help create an overall culture of data management. Other training is more about connecting and translating existing expertise to specific terminology for data management.

48. Please indicate the methods your library has used to facilitate additional training for the RDM services staff. Check all that apply. N=52

Workshop attendance	48	92%
Conference attendance	44	85%
Independent study	35	67%
Training provided by professional organizations	32	62%
Local courses in computer or digital technology	13	25%
Training provided by vendors	9	17%
Hire consultants	4	8%
Other training methods	10	10%

Please specify other training methods. N=10

Cross-training among librarians

Dorothea Salo's Introduction to Research Data Management course

ICPRS Training

Independent research on best practices

Online course on data management course for librarians from U of Wisconsin, Dorothea Salo

Participation in data curation pilot projects with related organizations.

Scientific Data Management course at a local MLS Program.

We have a dedicated committee, the Data Education Working Group that meets monthly to design and arrange training for librarians to support data services.

We haven't done formal training; learned on the job. Sending specific staff to training workshops (like the Purdue workshops, or the ARL E-Science Institute) was very helpful.

Webinars

FUNDING

49. How are your library's current RDM services funded? Check all that apply. N=53

Internal library regular budget	52	98%
Direct administrative funding (separate from library funds)	6	11%
External grant funding	6	11%
Internal library temporary or special project budget	5	9%
Department or research institute/project group funds (e.g., a research project funds specific RDM assistance)	3	6%
Endowment fund	3	6%
Fee to researcher or researcher's grant	2	4%
Facilities and Administrative (F&A) funding	1	2%
Other source of funding	5	9%

Please specify any other source of funding. N=5

Contracts for services

Funding from the state library, which is a project partner for the managed digital repository being developed.

IT internal budget

RDM services are part of the core work of the libraries, so the funds are the same as for other core liaison, reference, technical, and other standard operations.

The technology support funding comes from Information Technology Services (ITS), the parent organization of our Digital Library Technologies unit. Except for this one question we have considered DLT to be a part of the Library when answering all other questions because they largely operate that way.

If your library received external grant funding, please identify the funding agency(ies). N=6

Alfred P. Sloan Foundation (DMPTool2), National Science Foundation (Engineering Genome Project)

Government of Canada

Grant funding is received for collaborative projects with the libraries and researchers proposed by UF to various funding agencies including: NSF, NEH, NHPRC, IMLS, and LSTA.

National Library of Medicine

NSF-EPSCOR

Sloan/CLIR data curation fellowship

RDM SERVICE OUTREACH & ASSESSMENT

50. Please indicate which outreach methods are used to encourage faculty/researchers to use your RDM services, then select up to three of those that have been the most effective. N=53

Outreach Method	Used	Most Effective
Library staff referrals/promotion	44	7
Website links cross-posted to other library site pages	43	5
Workshops or presentations to faculty or student groups	41	19
Direct emails to faculty/researchers	35	12
Website links cross-posted to administrative sites	30	9
Research Projects (Grants) Administration referrals/promotion	30	14
Academic department administrator referrals/promotion	25	4
Website links cross-posted to academic department sites	14	2
Other outreach method	14	4
Total Responses	53	41

Please briefly describe any other outreach methods. N=13

Collaboration with and promotion through other campus offices like the Odum Institute and the Center for Faculty Excellence.

Director of SPARCS (Sponsored Programs) sent email to all researchers about library services and support for RDM.

Faculty outreach through a survey.

Held Campus Data Summit and other promotion/outreach events.

LCD and other social media (Twitter, blogs, etc.) are also used.

Posting flyers in campus buildings.

Print promotional materials

Representation at the university's New Faculty Orientation and similar events. Articles in the Libraries' newsletter for faculty.

Social media

Targeted promotion techniques, e.g., open data during open access week

Webpages and LibGuides, communication with liaison librarians

Word of mouth

Working with SPS has resulted in the inclusion of the repository in the data management plans of over 700 proposals to date. Awards are tracked and PIs directly contacted. Postcards mailed to all faculty and staff. Posters. Extensive

outreach campaign. Subject librarians working directly with their liaison departments has been very effective, also. The Distributed Data Curation Center (D2C2) that was founded at the end of 2006, as well as librarians collaborating on interdisciplinary research grants, helped establish a reputation for the Libraries as a place to go for help with research data.

51. Has your library assessed the impact or success of your RDM services? N=53

Yes	6	11%
No, but we plan to	28	53%
No	19	36%

If yes or you plan to, what measures are (will be) used to assess these services? N=23

Answered Yes

An assessment of the repository was conducted in March 2013. There has not been an assessment of data service writ large.

Evaluation of last year's data management plans; assessment of web stats through Google Analytics; review of anecdotal comments from users.

Surveys, feedback forms/evaluations, tracking consultations.

Tracking statistics for consultations provided to and interviews conducted with individual researchers, and workshop/training attendance.

No, but we plan to

As yet undetermined.

Don't yet know how we will do this, but have planned to.

Downloads/article accesses and other analytics, consultations, uptake across the university.

Faculty survey to be distributed in the very near future.

Follow up surveys to users; focus groups.

More systematic tracking of requests for services; pilot project to assess workflow for digitization of data from multiple formats (audio, handwritten field notes, video).

Not sure yet. (2 responses)

Number of consultations. Metrics are still under development.

Number of service requests, statistics on data downloads, analysis of DMP's submitted with grant applications.

Our assessment method and criteria have not be defined or determined at this stage.

Review of all data management plans submitted with grant proposals, using a rubric/template created from the basic components of data management plans as expressed in funding agency guidelines and local best practices.

Statistics around reference transactions have been kept for some time but more meaningful metrics analysis and

assessment are being worked on.

Survey.

Survey and data curation profile interviews.

The impact is still developing as more faculty and researchers realize how much the libraries have to offer, so assessment will be determined after the scope/scale of impact becomes clearer for use in selecting the assessment measures and methods.

Uncertain at this time.

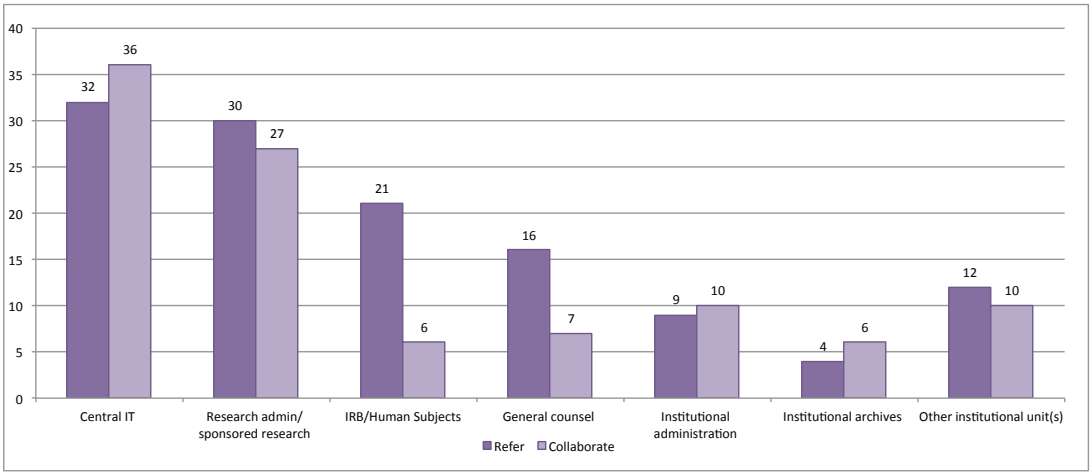
Usage of university sponsored DMP forms. Increased use of the institutional repository.

We don't know yet; probably a survey. We also were a principle developer of the DMPOnline tool. As a founding member of the DMPOnline we have access to institutional statistics.

PARTNERSHIPS

52. Please indicate whether your library RDM services staff has referred researchers to the groups listed below for particular RDM services, or has collaborated with these groups on RDM projects. Check all that apply. N=50

Group	Refer Researchers	Collaborate on Projects	N
Central IT	32	36	46
Research administration/sponsored research	30	27	42
Institutional Review Board/Human Subjects	21	6	23
General counsel, legal department	16	7	19
Institutional administration (deans/presidents, academic councils, etc.)	9	10	15
Institutional archives (if not operated by the library)	4	6	7
Other institutional unit(s)	12	10	15
Total Responses	42	42	50



Please identify the other institutional unit(s). N=15

Academic Health Center IT related departments, Minnesota Supercomputing center, Office for the commercialization of technology

Belfer Audio Archive (part of library's Special Collections)

Center for Computational Science: this is a separate department from central IT

College and departmental IT groups

Data Coordinating Center, Clinical and Translational Science Institute

Departmental or school IT

Departments/institutes on campus

Institute for Social and Economic Research, Center for Advanced Computing, University Statistical Consulting Unit, Computational Biology Service Unit, Center for Technology Enterprise and Commercialization, Copyright Information Center, Office of the University Counsel

Institutes, laboratories, other data producers on campus

Institutional repository

Professional school libraries (i.e., school of medicine, law school)

School of Information Science

Statistical Consulting, Survey Research Consulting, High Performance Computing (housed in a research center)

TACC

The Odum Institute, RENCI, DICE

53. Has your library partnered with external institutions or groups to provide RDM services? (By "partnering" we refer to a more formalized collaboration of people and projects outside of your institution, in contrast to referring patrons to external institutions such as to science data repositories.) N=51

Yes	13	26%
No	38	74%

If yes, please identify the partner(s) and briefly describe the nature of the partnership. Address what has made the partnership successful and/or challenging. N=13

A number of services/tools such as the Data Curation Profiles (http://datacurationprofiles.org), Databib (http://databib.org), and DMPTool (http://dmptool.org) are some examples of collaboration outside of our institution. The software we have been developing for PURR is in collaboration with the HUBzero Foundation.

California Digital Library and several members of the Orbis-Cascade Alliance, for the DataCite/EZID service for our repositories. Successful coordination to achieve a lower cost for the service by signing up as a group.

Canadian Polar Data Network, CPDN. http://polardatanetwork.ca/. Successful: highly collaborative, wide pool of expertise, coordinated access services, distributed preservation infrastructure, shared ingest/processing services. Challenging: business model, sustainability, governance.

Data curation pilot project with ICPSR's Director of Curation Services and data librarians at multiple universities. Applied data curation practices, processes, and tools developed by ICPSR to prepare, process, and ingest datasets into their archive. Challenging to allocate personnel time to the project. Successful due to the collegial people involved.

DMPTool.

DMPTool (joint development and planning). DataONE (users group, best practices development). ASERL/SURA (data policies development). Virginia Tech (graduate student bootcamp).

MIT has partnered with Harvard over many years to provide the Harvard-MIT Data Center that includes access to high performance computers, special application software, training for the software applications, and the development of the Dataverse Network. Many researchers who use these services have affiliations at both institutions. The two institutions are physically close, and relationships with colleagues are very positive. Ensuring that MIT needs and/or requirements are considered for Dataverse developments can be a challenge as the development and system support for Dataverse is at Harvard.

Participate in consortial committees through Ontario Council of University Libraries.

Research computing and IT each have representative members on the e-science working group. Willingness to cooperate has made this collaboration successful.

Virginia Tech and University of Virginia jointly sponsored a data management boot camp for graduate students at each institution participating on their local campuses and virtually.

We have partnered with CDL in use and development of some of their RDM tools.

We want to.

We were principle developers of the DMPOnline tool. We collaborated in the development tool with eight other institutions.

RDM SERVICE CHALLENGES

54. Please briefly describe up to three challenges your library has encountered in providing RDM services and the measures that have been tried to overcome those challenges. N=48

Qualitative answers to this question were organized and sorted, resulting in the following overall themes.

Collaboration campus-wide	18	38%
Funding	17	35%
Faculty Engagement	15	31%
Technology Infrastructure	13	27%
Limited Staffing	12	25%
Marketing Services	12	25%
Staff Training	11	23%
Scoping services	9	19%

Staff Roles	8	17%
Institutional commitment	7	15%
Faculty Education	5	10%
Evaluating demand	4	8%
Policy	3	6%
Scaling services	3	6%
Agency ambiguity	2	4%
Other	3	6%

Collaboration Campus-wide

Buy-in from various constituencies (i.e., research administration workflows, researchers, graduate students). Measures: lots of information sessions, one-on-one meetings to make pitches of value, training workshops, boot camps.

Collaborating with the office of research. Successfully worked with OR on joint web pages on DMPs and on presenting a panel at a Research Development network meeting

Collaboration with campus units.

Communicating with non-science researchers about how their work is a part of RDM.

Decentralized campus IT and data-related services. Science Data Services Librarian has made many connections between various players, initiating coordinated activities, attending meetings, and bringing various units together to work on solutions. New CIO with a renewed focus on research IT support is also bringing about positive changes.

Decentralized nature of the university. Thinking strategically and cultivating relationships with stakeholders across campus to make mindful recommendations about services to develop or improve.

Interest and support from potential collaborators. We've received some assistance from the Center for Computational Science but haven't been able to develop anything sustainable.

Potential collaborators (e.g., division of research) have not yet recognized the value of RDM. (*also Institutional commitment)*

Some issues related to primary responsibility/leadership for certain RDM service roles and responsibilities between library and other units (campus IT, technology transfer). New CIO and reorganized campus IT has addressed most of the IT-related issues, but there is a need for ongoing coordination at administrative and operational levels to keep everyone on the same page.

The largest challenge we have has been silos within the institution. The office of research, campus IT, and high performance computing have all had stakes in this and were for a while working in a vacuum from one another. We have now established some cross-divisional working groups to help alleviate the problems. It is not totally fixed, but is better.

Trying to partner with campus IT department (some meetings have taken place).

We have moved too slowly and the rest of campus is launching disjointed services.

Funding

Addressing costs of research data management: value of services, costs for preservation, incentive structure for researchers. Measures: carrot (make services available and easy) and stick (information sharing about requirements and

admin buy in).

Allocating resources and facilitating professional development within the Libraries to support the wide range of data management situations, problems, and issues that arise at a large research university. This is an ongoing challenge and will be resolved, in the short term, by restricting our services to a selection of data management situations and, in the long term, by sufficiently demonstrating the value of RDM services to the Libraries and the university in general that more resources are allocated to RDM services. *(also Staff training)*

Build case for more secure funding and/or expansion of related services.

Cost of infrastructure and personnel for RDM services. Measures tried: Collaborating with other units on campus; seeking grant funding. *(also Collaboration campus-wide)*

Describing the need and justifying a budget.

Developing a sustainable financial model, in particular with the inclusion of fees to the researcher for data archiving.

Financial support for service

Funding and staffing. Demand (by faculty) has so far been light but with increased funder expectations our staffing is too low and funding for the expected amount of research data to store, tools (visualization, analysis) to develop will soon be a major problem. *(also Limited staffing)*

Funding and support for program and services.

Funding issues have been a challenge. This exhibits itself in inadequate staffing, underestimating the needed resources, and sustainability concerns. We are establishing a governance group that can hopefully address financial and policy issues. *(also Institutional commitment)*

Implementing sustainable services.

Insufficient resources to significantly move this forward as a service.

Money: hard to do very much with little to no funding; considering grant funding but currently the library is not well set up to manage the regulations surrounding grant funding.

Research is paid for and supported by the funders out of direct grant charges, not the institution, so gaining knowledge about what research is happening is a challenge. The Libraries are not involved in the normal workflow of research administration on campus. Measures: The Libraries continue to form partnerships with other administrative groups such as the Office of Sponsored Programs, Information Technology Services, and Office of General Council in order to better understand the research landscape and identify gaps in data services.

Faculty Engagement

Attracting participants in training sessions.

Connecting with faculty: easier to connect with research support staff but hard to identify researchers' needs.

Convincing researchers to speak to the library about data management.

Faculty buy-in is always tricky. Proving to them that the library is in fact the place for their data needs.

Lack of core faculty interest, particularly in sharing data in open environments.

Many faculty are not taking these mandates seriously or have determined a process that has allowed them to submit, and they are sticking with it.

Our data archiving services as a recognized value added to research on campus. Preservation and data sharing are a hard sell when the researcher only needs to write a plan...not implement one.

Reaching grant writers well before their deadlines.

Researchers have few tangible incentives to actively support data management and curation. This challenge is not limited to our institution, so measures to overcome it will come from a variety of individuals, groups, organizations, and institutions. For our part, we aim to minimize the administrative burden associated with our data management and curation services and, generally, advocate for data as a first-class, credit-worthy research product.

The independent nature of the researchers and the highly decentralized culture is a challenge. Measures: We try to partner with our library liaisons who have existing relationships with researchers in the departments, labs, and centers to know what research data issues are being encountered and to promote our RDMS.

Technology Infrastructure

Developing an institutional technical infrastructure for research data storage and management. Measures to overcome this challenge: proposed a major storage infrastructure acquisition to the university's provost office (declined); developed and submitted an NSF Major Research Instrumentation grant proposal to acquire a petascale storage array (currently under review by NSF).

Developing the technology infrastructure.

Digital preservation for research data.

Highly decentralized nature of the campus IT infrastructure. *(also Collaboration campus-wide)*

Improving the technical infrastructure to archive and curate research data: four pilot projects currently underway along with major DAMS upgrade.

Inadequate long-term curated storage [advocacy for central funding for local and cloud solutions] *(Also Funding)*

Providing archiving and storage options.

Providing support while lacking a data repository.

Technical infrastructure: we are limited by both the Libraries' infrastructure as well as the campus' infrastructure options. We are in the process of proposing a campus-wide task force to address this issue (specifically, research data management and storage infrastructure, as opposed to computation).

Technology limits the size of files we can accept into our repository.

The absence of appropriate technical infrastructure and support for researchers. The library has begun looking towards OCUL/Scholars Portal's Dataverse instance as a potential data repository for researchers.

We lack an appropriately configured repository for optimal data storage.

What repository to recommend to faculty for sharing their data? Some disciplines don't have a repository. We can recommend that faculty use the campus repository, but there's a limit to file size that can be loaded. If they don't have a disciplinary repository, what should we recommend? Looking at this institution-wide with central IT and the Office of the VP for Research involved; also being looked at from a university perspective**.**

Limited Staffing

Gauging resource capacity to potential service demand is a challenge. There is a limit on the time that staff can dedicate to RDM services. We've tended to be conservative with our service offerings until we have a better understanding of

data service needs and funding for those services. Measures: We are proceeding with needs assessment and a digital content review of research data to help inform our service offerings.

Lack of staff time: We all have full time work, to which this has been added; measures include prioritizing, planning, doing what we can with what we have; at this point there is little chance of adding personnel in the near term.

Lack of staff time to devote to providing RDM services. At this time, there is no plan to hire additional staff, so this will continue to be a challenge.

Lack of staffing devoted to RDM.

Limited staffing.

No dedicated staff to data management.

Staffing: the Data Working Group is a voluntary group; we are hoping to present a case for a staffing scenario that recognizes the need for sustained work and increasing visibility of RDMS.

Staffing: We have been identifying the staffing needed and strategies to recruit within an overall library reorganization.

Stretching existing staff capacity.

Time: hard to keep content fresh and plan workshops when this is an added-on task for all of us in addition to our regular jobs.

We do not have current staffing to support RDM services.

Marketing Services

Awareness of services.

Building awareness among library colleagues and campus administrators.

Campus awareness of services continues to be low.

Communication of services offered.

Create awareness of service to campus constituency. Develop new marketing techniques, identify new collaborators, etc.

Maintaining awareness of services offered by the Research Data Management Service Group (RDMSG) across campus over time is an ongoing effort for us. In 2011–2012, we offered regular (each semester) information sessions about the NSF DMP requirements, which made the RDMSG name visible, but with reduced attendance at those sessions, we've begun additional outreach efforts such as "walk-in" office hours. We continue "advertising" by way of limited email announcements, twitter and LCD screens across campus, and by maintaining communications with the full group of library liaisons.

Marketing the existence of our service. We've identified many of the units at the university that interact with researchers at some point in the research lifecycle. We've systematically reached out to these units, perhaps on more than one occasion, to update them on our service and look for points of collaboration and information sharing. This helps us understand what all the different units that support data management do so that we can refer researchers to them, and in turn, it helps these units refer to us if the occasion arises.

Marketing to faculty, what's the best way to let faculty know about these services? There is an established Campus Data Management Committee with representatives from the library, Office of the VP for Research, central IT, and faculty from several departments. Response: Sending out memos to faculty and researchers on campus. Added information about

services to message that sent out by Sponsored Projects when researchers receive grant award (going forward in 2013). Directly contact researchers with new grant awards (received in 2012). Work with librarians who interact regularly with faculty to let them know about our services.

Marketing value of our service to researchers. Because proposals with poor data management plans are funded and data sharing via a publicly accessible repository is not mandated, selling the value of our service has been a challenge. We try to find ways to personalize how our service impacts them. For example, we discuss how good data management will allow them to find and understand their own data in five years time. Also, we discuss the how deposited data in the data archive receives DOIs with their data for citation by others and as a research product to list on future NSF grants. *(also Faculty Engagement)*

Promoting awareness of services to faculty and university administration, particularly getting faculty to embed the services into their research workflow. *(also Faculty engagement)*

Providing consistent and effective outreach to inform the research community about the library's RDM services. Measures to overcome this challenge: provide a monthly standing research-data working group that includes academic department IT and research liaisons to use as referral mechanisms for researchers.

Research awareness/receptivity. Measures tried: A variety of outreach/promotional efforts, including websites, presentations, print materials, emails, news announcements, etc. Also, providing internal training to subject librarians to increase outreach skills around RDM. *(also Faculty engagement)*

Staff Training

Infrastructure: capacity building; can be expensive; toolkit development and/or adoption of existing tools (growing but still relatively few).

Lack of domain knowledge of librarians (i.e., STEM)

Lack of experience / knowledge in IT areas; also lack of experience working with research data. Measures include partnering with campus IT and OSP, and doing pilot projects to learn about research data, just about to embark on pilots. But this effort is a first in collaboration of this kind with OSP and central IT so we are learning as we go. *(also Collaboration campus-wide)*

Interest and expertise among already existing staff. Some have agreed to take it on but have many other responsibilities that limit their ability to focus on RDM services.

Educating librarians on research methods.

Ensure staff who support RDM services have appropriate knowledge and skills. We have begun to support staff to attend RDM workshops organized by professional associations. Staff who have attended these workshops have begun to share their knowledge with colleagues upon their return.

Education and training of liaison librarians.

Lack of specialized training in data management issues—specifically the use and sharing of data as it relates to individual subject areas. Those involved in RDM services have tried to attend training sessions to increase their knowledge in this area.

Lack of awareness around actual research processes (i.e., subject liaisons uncomfortable, RDM staff need to gain deeper knowledge of processes, etc.) Measures: data interviews, embedding ourselves in research processes when possible, exposure to as much research practice as possible, workshops with Q&A as new data points.

Scoping Services

Attitudes and needs of campus researchers for RDM are largely unknown. *(also Evaluating demand)*

Balancing specialized/project and generalized/services among a great diversity of needs, data formats, etc.

Challenge: defining needs. Measures: faculty survey.

Creating a clear services statement.

Current focus is almost entirely on the active phase of the research, not on long-term preservation and access of data. No action has been taken on this challenge yet.

Defining and articulating what "data services" we offer, determining roles and responsibilities and situating them, not only with personnel in the Libraries, but with our colleagues in IT, SPS, and the research office in our collaboration.

Defining and scoping services.

Getting the service off the ground, planning, etc.

Locating, measuring, and evaluating the demand for RDM services. This is an ongoing challenge that will be resolved through comprehensive engagement and dialogue with researchers, administrators, and technologists. To date, we have interviewed researchers, discussed RDM issues with researchers, analyzed funding proposal and award data, analyzed DMPs written by researchers, shared insights with campus IT, and pursued a number of ad hoc information gathering activities to understand demand. *(also Evaluating demand)*

Managing expectations. Some faculty, when they hear that the Libraries will help with DMS, then just expect us to do everything.

Primary challenge pertains to meeting the researcher needs brought about by the changing data management landscape. Continue to monitor trends, assess service, etc.

Time it takes to develop services. Conducting a comprehensive assessment of researchers' needs (survey completed, interviews ongoing, partnerships with other institutional departments/offices under consideration).

Understanding faculty needs as we define the scope of the program, and matching training of staff with local needs.

Staff Roles

Balancing traditional library services with RDM.

Challenges on understanding and accepting roles for data management.

Defining library RDM service roles in relation to other units on campus (e.g., research computing). Measures tried: Keeping communication between units open and frequent; examining how other institutions have defined roles.

Exploring and defining team roles in data management projects that can involve multiple units, such as solutions for better data management in the research environment (IT as developers, library as curation/preservation resource, researchers as the client). This is a new venture and will require open communication and collaboration from all. *(also Collaborations campus-wide)*

Fully recognize and benefit from existing experience and knowledge of data management that exists within the library. The library's RDM Working Group includes members with data management knowledge & experience and members for whom data management are a relatively new domain. *(also Staff training)*

Getting outreach librarian supervisors to make RDM a priority.

Motivating outreach librarians to have research conversations with faculty.

Subject librarians do not see this as part of their job and do not participate. We are trying to slowly engage them with LibGuides and workshops.

Institutional Commitment

Institution-wide leadership on RDM issues is lacking.

Institutional support: We have reached out to the Office of Research Development and are building relationships there. This is an ongoing effort that is tied into both the staffing and infrastructure challenges that we face.

Lack of institutional and administrative support

Lack of institutional priority

Faculty Education

Developing workshops for faculty and researchers, could develop general workshops for faculty, but may be more effective to contact colleges. Contact colleges asking to attend a department head's meeting where we present data management services and offer to provide workshops for faculty in that department. Planning on developing videos that faculty can view at any time.

Development of a culture of data management, curation and preservation. Raising faculty awareness of issues, opportunities, challenging, IP, access (open vs closed), ethical use, rewards, attribution, citation.

Educating/knowledge of faculty in understanding data management

Finding time for the boot camp was during spring break.

Mainstreaming data: understanding the implications of RDM as core library service. (*also Faculty engagement)*

Evaluating Demand

Development of a service in the absence of demonstrated need on the part of researchers. The e-science working group was formed to help address this.

Finding out about DMPS already submitted. Working on a library policy on DMP deposits by successful awardees and have worked with Office of Research on acquiring DMPs provided by PIs *(also Collaboration campus-wide)*

Gathering data on data management to assess the need for assistance by researchers. The e-science working group has drafted a survey and enlisted the support of the vice president for research. *(also Scoping service)*

Getting a university-wide data management policy that is clear and has teeth.

Participating on assessment of faculty computing needs. Able to insert wording in a final report on need to go beyond addressing faculty data storage needs to address data management, curation, preservation/archiving, and discovery. *(also Scoping services)*

Policy

Silo nature of policies and budgets at university level *(also Funding)*

Scaling Services

Identifying the limitations of the envisioned services *(also Scoping services)*

Challenge: scalability. Measures: investigating cross training, researcher empowerment, online training options.

Scaling for consultation services [build campus service catalog]

Agency Ambiguity

Ambiguous requirements by funders for DMP and data sharing; measures include consulting with OSP, staying abreast of developments.

No incentives for faculty to value of data management *(also Faculty engagement)*

Other

Creating local infrastructure.

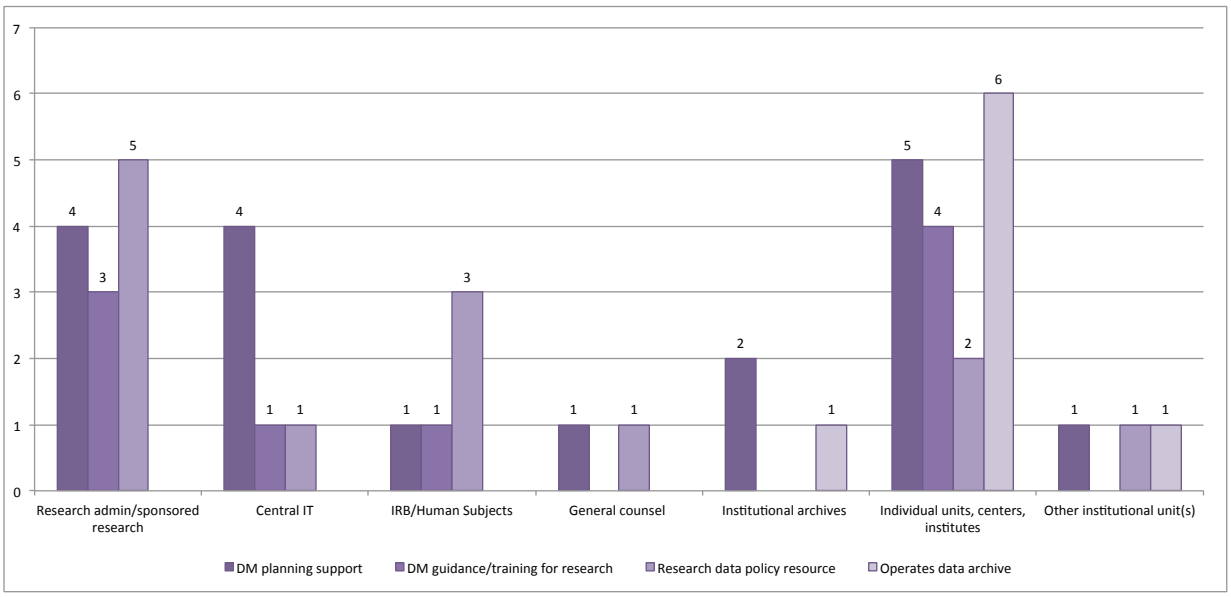
Finding participants for pilot projects.

The library did not launch a focused effort to leverage resources effectively and efficiently for entire campus.

OTHER UNITS OFFERING RDM SERVICES

55. The following is a list of departments and units in an institution that typically may be involved in providing RDM services. Please indicate which, if any, RDM services (such as data management planning for grants, training on data backups and security, or research data archiving) these units provide at your institution. Check all that apply. N=13

Department/Unit	DM planning support	DM guidance/ training for research	Research data policy resources	Operates a data archive	N
Research administration/ sponsored research	4	3	5	0	9
Individual academic units, research centers, or institutes	5	4	2	6	8
Central IT	4	1	1	0	5
Institutional Review Board/Human Subjects	1	1	3	0	4
Institutional archives	2	0	0	1	3
General counsel, legal department	1	0	1	0	2
Other institutional unit(s)	1	0	1	1	3
Total Responses	8	9	9	6	13



If you selected other institutional unit(s), please identify the unit(s) and specify the RDM service. N=3

DM Planning Support

Research Computing Center: Data Management Planning/Guidance/Training

Research Data Policy Resources

University Libraries and the Department of High Performance Computing

Operates a Data Archive

Special Collections & Archives, within the Library

Additional Comments

Office of Research Ethics focus is on research ethics. The protocol submissions include statements about the length of time that data will be retained, how data will be stored, where it will be stored, who will have access, and how the researchers will protect the privacy, confidentiality, and anonymity of participants in the research, and in some cases questions arise whether the data will be archived.

Our approach will require cross-functional and coordinated efforts, knowledge, and experience of many constituents (internal and external).

We don't know the answers to any of these and don't want to speak for other units.

FUTURE PLANS FOR RDM SERVICES

If your library is currently offering RDM services, please answer the following questions about any plans your library has to offer **additional** services, or make staffing or funding changes in the future.

If your library is NOT currently offering RDM services, but plans to, please answer as many of the following questions as possible at this time.

56. Please indicate which, if any, of the research data management (RDM) services below your library plans to offer but has not yet implemented, and when they are anticipated to begin. Check all that apply. Do not check boxes for services your library currently offers (which would have been indicated in prior questions). N=56

RDM Services	Within 1 year	Within 2 years	Within 3–5 years	No current plans	N
Services to support ongoing research (e.g., back-up planning)	11	16	2	14	43
Other online resources for research data management besides DMP preparation	20	15	0	6	41
Research data archiving	17	13	6	5	41
Training sessions on DMP preparation and/or other data management topics	15	11	3	4	33
Online resources related to Data Management Plans for NSF proposals or other funding agencies	20	5	0	7	32
Direct assistance/consulting with researcher for DMP preparation	13	10	3	5	31
Other RDM service(s)	8	3	0	8	19
Total Responses	40	29	7	23	56

Please briefly describe the Other RDM service(s). N=11

Within 1 year

Although we are able to support data archiving in our institutional repository, we plan to develop a data repository in the future.

Dataviz tools and guidance (tentative)

Expand archiving in the local data repository

Metadata creation and related quality control

More robust data curation services

Provide metadata for computational data that can be harvested from the repository. Systematically contact faculty with current grants to offer metadata and archiving services for research data.

Raise awareness of data governance, data citation training, use of identifiers for researchers (ORCID) and data (DataCite).

Training opportunities

Within 2 years

Expansion of identifier assignment, assistance with sensitive/restricted data

Statistical reference services; database creation

The development and offering of other RDM services will be informed by our current needs assessments.

Additional Comments

As an institution in Canada, we don't have a lot of researchers applying for NSF grants, hence our answer to the first question. The Canadian funding agencies don't yet required DMPs.

Providing access to research data for use in further discovery, research, and scholarship.

If you plan to phase in services over time, please enter any additional details about your phase-in plans and timeline. N=11

Staffing

At this time there are plans to hire a Research Data Coordinator. Currently there are no plans for any of the above until the coordinator is hired and decisions are made on priorities.

Data services will start on July 1, 2013, with the arrival of our newly hired Data Librarian. The first three activities she will work on are implementation of a website delineating our services, customization of the DMPTool, and a survey of the data landscape at the university.

We are hoping to begin with a Business Analyst position who will lead the assessment of needs and the development of an implementation plan for research data acquisition, management and curation at the Centre for Hip Health and Mobility (CHHM), aligning business initiatives with information technology solutions. The Business Analyst will provide expertise in the areas of requirements definition, business process analysis and design, functional design, configuration, implementation, testing, training and documentation to deliver an enterprise level solution that will provide the framework for implementing a data management and curation strategy for the larger university research community. The majority of work will be performed working closely with researchers and staff at CHHM, with support from formal project teams within Research IT Support and the library. This project has a high degree of complexity, influence, and impact as the solution must provide the basis for a research data management and curation strategy that meets the needs of the majority of constituents in the university research community.

We are hoping to hire a data management specialist this year to organize our services.

Assessment

Currently, the university libraries is planning on developing a data management day with the Department of High Performance Computing in fall 2013. Once we have completed the data management day event, HPC and the libraries (under the guidance of the E-Science Librarian) will submit a campus-wide data management survey, much like the survey developed by GA Tech's Susan Wells Parham. Once the survey data has been collected, it will be presented to the OVPR in order to develop an approach to campus-wide data management services. This is where the campus is on data

management planning at this point.

Planning

Timelines for DMP are dependent on Canadian funding agencies introducing data management plans as part of the grant application requirements.

We are doing some work in each area now but plan to expand our services going forward. For data management we want to populate the DMP Tool with institution specific information; for DMP training we want to do more subject specific workshops with graduate students; for archiving we have 3-4 pilot projects coming soon.

We have a proposal in to support hardware/software/staffing for a data archive and consultation services.

We plan in the next year to increase effort around training sessions, operational RDM, archiving through our IR, and cost models/budgeting for RDM.

We would like to do that, but it will depend on staffing and expertise as well as resources to move it forward.

We've just started thinking about it, we don't know about any of this yet.

57. If you indicated above that your library plans to offer research data archiving, please indicate the repository model. Check all that apply. N=47

We plan to assist researchers with locating external data repositories	34	72%
We plan to directly assist researchers with depositing data at a repository (whether library-operated or external)	33	70%
Our institutional repository will be used to archive research data (and used for publications)	23	49%
The library will host or be directly involved with archiving research data at an external archive or repository service, such as DuraSpace or DataVerse	14	30%
The library will operate a research data archive dedicated to data deposits and access	13	28%
We will be involved with data archiving that is primarily operated by another department/group at our institution	12	26%
Other repository model	9	19%

Please briefly describe the other repository model. N=9

Chronopolis

Dark replicated archives, library-hosted

Possibly look toward other repository models as we mature our data archiving services.

The libraries are investigating other supports, which could include external hosting options to be added with the existing excellent, robust internal supports through the university.

The library-run Fedora repository being developed will include research data but also many other kinds of data and objects.

This process is still under discussion.

We are evaluating the Cornell University Library Archival Repository (CULAR) as a possible future archival service for large data sets. Current development trajectory does not include support for data archival, but CULAR remains one possibility for such a service and may evolve over time as dictated by library priorities and researcher needs.

We are still discussing the models with the university IT department and the vice-president research office. Several of the models listed above are being considered.

We need to do some analysis to decide what would be the most appropriate technology solution(s).

FUTURE PLANS FOR RDM STAFFING

58. Please indicate if your library plans to add staff positions for RDM services or add RDM roles to existing staff positions to support services in the future. Check all that apply. N=68

We will be adding one or more staff positions primarily for RDM services	31	46%
We will be adding RDM roles to existing staff positions	29	43%
No staff changes planned at this time	23	34%

Please comment on your staffing plans, including how many staff positions will be created and/or reallocated. N=26

New Positions and/or Postdocs Being Added

3–4 new positions will have expectations in this area.

CLER fellow for Digital Arts and humanities. Possible other roles in research support.

Currently working on a Copyright Services Librarian position, which will provide some support for RDM services, intending to fill by fall 2013. Also drafting research data services consultant position, which would hopefully be filled by EOY 2013. GIS librarian search will commence later year and expect that position to have increased responsibility for GIS data management support.

Just added two in past year and do not have current plans to add more. Will adapt as needed with changing environmental factors.

Research Data Coordinator

The Business Analyst position will inform the sustainability and staffing planning.

The most recent budget plan includes a Data Curation Librarian, to be hired if the budget allows. Additionally, the objectives for the coming year include building a full system of data curation, which includes integrating RDM roles into existing positions.

We are adding a Research Data Librarian (faculty) and a Research Environments Librarian (faculty).

We are currently hiring a third data management consultant. We hope to have this person hired in the next few months.

We will adding a Computational Sciences Information Specialist as part of a campus hiring cluster focused on Big Data, and also a Molecular Biology Information Specialist as part of a campus hiring cluster focused on Bioinformatics.

We will be creating 1–3 new positions and reassessing staff from other work into this area.

We will be hiring a data librarian in 2013 and hopefully re-allocate some staff to help with data repository work in 2014.

We will be hiring a metadata librarian who will have some RDM role in the next 6 months. Also we have a project team in the library looking at grant proposal support.

Will be hosting a Postdoctoral Fellow in data curation 2013–2015. We may continue to fund a librarian internship position in data curation. Budgets make this difficult to forecast.

New Positions Proposed/Hoped For

Possibility of hiring Post Docs.

We have a proposal in which, if funded, would add five new positions (from technical to user support).

We hope to add a post-doc position to focus on data management if resources can be found in the next year or two.

We hope to add a repository manager, an e-research manager, a digitization manager, and another repository software engineer (programmer).

We hope to add one librarian position.

When the budget allows, we will add a data librarian position. We do not know when that will be.

No New Positions/Repurposing Existing

Data and data management are acknowledged issues of importance in the next few years, and a variety of options have been discussed including adding a data visualization or text mining position, or a metadata position that could help a variety of library units. However, due to the current budget situation, likely nothing will happen unless current employees leave/retire and we can re-purpose those positions. New positions are very unlikely.

Many subject librarians are taking a greater role in data management consultation.

Position Planning/Assessment

Generally speaking, more staff resources are needed to support RDM services.

I am unaware of any RDM staffing plans, but additional staffing would be necessary to meet patron needs.

Our library is current doing some strategic planning around digital content.

Our staffing plan will eventually be informed by the outcome of our current needs assessments.

We haven't given this any thought, yet.

59. If your library has plans to add RDM positions or allocate responsibilities to existing staff, please indicate the primary responsibilities for the position(s). Check all that apply. N=45

Data archiving assistance (includes IT & deposit assistance)	30	67%
Data management guidance for researchers (other than DMPs)	29	64%
Grant proposal Data Management Plan support	26	58%
Subject reference services	22	49%
Metadata librarian	19	42%

Other data services (e.g., GIS training)	18	40%
Library IT, systems admin	13	29%
Cataloging/collection development	9	20%
Administrative	7	16%
Other responsibility(s)	6	13%

Please briefly describe the other responsibility(s). N=6

A Business Analyst position who will lead the assessment of needs and the development of an implementation plan for research data acquisition, management and curation.

Data Curation Librarian to identify, store, describe (curate), retrieve, and re-use data, particularly data not available in public or government repositories. This is comparable to the program at Johns Hopkins University Libraries Digital Research and Curation Center (http://ldp.library.jhu.edu/dkc), providing support for researchers in meeting federal mandates and enhancing curation and accessibility of their research product. (1 faculty)

Data visualization support

Digital arts and humanities

Digital Collections creation and management

Research using the data

60. What educational background/experience will be most crucial for the position(s)? Check all that apply. N=46

Data curation emphasis (with MLS or other degree)	34	74%
Subject domain expertise	26	57%
Digital preservation training	25	54%
MLS/MLIS - Library or information science emphasis	20	44%
IT technology or services experience	18	39%
Graduate degree in a discipline other than LIS	10	22%
PhD in a discipline other than LIS	7	15%
Archives emphasis (with MLS or other degree)	5	11%
Other training or experience	10	22%

Please briefly describe any other educational background/experience that will be particularly important. N=10

Education/training in statistics

Faculty positions in the libraries normally stipulate "ALA-accredited MLS or equivalent advanced degree in a relevant field" and required and preferred qualifications. Qualifications for this position will include data curation experience/expertise and appropriate combination of expertise, experience, and education.

Informatics

Legal training for the Copyright Services Librarian role

Metadata

Project management

Some experience working with research data.

Statistics, data wrangling, data visualization, computational data retrieval, scholarly communication programming, and other functional education/experience

Statistics; R

The Business Analyst will provide expertise in the areas of requirements definition, business process analysis and design, functional design, configuration, implementation, testing, training and documentation to deliver an enterprise level solution that will provide the framework for implementing a data management and curation strategy for the larger research community.

FUTURE PLANS FOR RDM FUNDING

61. If your library is currently offering RDM services, please indicate any new funding sources that are anticipated for these services in the future. If your library is not currently offering RDM services, but plans to, please indicate the anticipated funding sources. Check all that apply. N=58

Not yet determined	30	52%
Internal library regular budget	21	36%
External grant funding	15	26%
Internal library temporary or special project budget	9	16%
Direct administrative funding (separate from library funds)	7	12%
Department or research institute/project group funds (e.g., a research project funds specific RDM assistance)	7	12%
Fee to researcher or researcher's grant	5	9%
Facilities and Administrative (F&A) funding	2	3%
Endowment fund	2	3%
Other source of funding	4	7%

Please specify any other planned source of funding. N=4

Any of these are possible in the future, depending on needs and for specific projects budgets and for programmatic budgets and operations.

Contract for services

ITS internal budget, special projects budget

The potential funding sources will depend on the model that is selected. Potentially, funds will be sought from other

university units, central administration, or perhaps the government.

62. In the next three years, do you expect your library's overall allocation of funds for RDM services to increase, decrease, or stay about the same? N=64

Increase	42	66%
Decrease	1	2%
Stay about the same	21	33%

Please enter any comments you may have on anticipated funding of your RDM services. N=10

Because our archiving service is paid through charges back to the grant, our funding should grow as more and more researchers archive with us.

It is difficult to talk about the source of funding for future at this time because of uncertainty about senior leadership positions, e.g., VPR, provost, president.

This is currently unknown.

Very hard to predict though our strategic plan is committed to RDM services.

We are actively involved in securing both federal agency grant-based and endowment funding for increased research data services in our library. Even if our current efforts are less than successful, we will continue to pursue these routes as ways of supplementing our library's limited budget for research data services.

We are currently doing some basic things with resources and tools, but if resources allow will begin to develop real sustainable services in this area.

We have no specific goals with regard to funding but anticipate the need growing and our resource allocation changing as well.

We haven't given this any thought, yet.

We really don't know what the funding outlook is because it will be determined by the new UL who is starting in August.

While RDM work is core and continuous with other library curatorial activities, the work can open into a higher ongoing level of engagement with researchers. This could lead to stronger and broader impacts from all research. This would also increase overall demands and require more funding overall.

NO LIBRARY RDM SERVICES

63. Please briefly explain why your library is not providing RDM services. N=2

Little to no demand.

We do not have the resources to introduce new services at this point.

64. Please indicate which, if any, of the circumstances below would influence a decision to begin providing RDM services at your library. Check all that apply. N=2

A Data Retention/Research Ethics policy at your institution (e.g., allowing potential data audit)	2	100%
Increased recognition of the need for better data sharing	2	100%
More requests from faculty/researchers for assistance with data management planning	2	100%
New grant funder requirements	1	50%
Growth of data intensive research	1	50%
Increasing academic credit for sharing data (e.g., as a data collection author or for tenure/promotion)	1	50%

Comment

All of these factors could influence us but the bottom line is we would still need more resources (expertise, staff time, technologies...) to offer those services (short of being able at this point to drop any existing services).

65. The following is a list of departments and units in an institution that typically may be involved in providing RDM services. Please indicate which, if any, RDM services (such as data management planning for grants, training on data backups and security, or research data archiving) these units provide at your institution. Check all that apply. N=1

Department/Unit	DM planning support	DM guidance/ training for research	Research data policy resources
Research administration/sponsored research	1	0	1
Individual academic units, research centers, or institutes	1	1	1

ROLE OF RESEARCH LIBRARIES

66. Please briefly describe the role you see research libraries playing in supporting research data management, now and/or in the future. N=30

Academic libraries have always had a role in educating researchers—extending that education to cover RDM is no different. I also wonder whether it wouldn't be useful to offer metadata consulting in the library, another area where libraries have long had expertise. We don't offer reference desk services worried about how we will serve every single person on campus—why not metadata consulting?

Actively archiving research data. This is a natural extension of our role as the institutional memory.

An educational and supportive role, helping researches learn of things like the DMPTool and existing data repositories and practices.

At the most fundamental level, research libraries will be expected to assist their researchers with planning for the management of data throughout the data lifecycle. Beyond providing a library-operated data management writing/

planning service, I cannot say exactly where research libraries will find themselves when it comes to data management planning. So much of where libraries will go depends on the university in which the library is operating. I see universities as being unique, each university having its focus and individual culture. To say that every university library will offer RDM services in the future is a stretch. Unlike book loaning or providing catalog searching services, RDM services do not have to be offered by a library. In fact, there's no universally accepted expectation that libraries will be offering RDM services. RDM planning, for example, could be offered by an office of research just as easily as a library. The question comes down to partnerships because RDM services touch so many stakeholders on campus, and no one campus department, unless it has been designed specifically for the purpose, will provide every RDM service a researcher could possibly hope for. Thus, partnerships must be formed and agreements made on who will offer these services on a research campus. However, I am of the opinion that libraries will be the future service providers of RDM resources and RDM plan writing, due to the fact that it is the one place where they can establish a foothold without too much investment in the changing research paradigm.

Based on survey responses and interactions with researchers I've had during our campus-wide study thus far, I see opportunities for libraries to support RDM through DMP consulting (with certain disciplines), data archiving (in institutional or disciplinary repositories), and perhaps licensing of datasets. I believe instruction and liaison librarians can play a role by promoting resources and incorporating data literacy in their instruction. Many researchers, however, would not be best served by the libraries or they would prefer to find assistance elsewhere. Any success in RDM services will come from active partnerships with existing research services on campus, such as IRB, central IT, research administration, and designated research support services within colleges or departments.

Continue to provide service, support and training for data management.

Faculty rarely physically visit the library now; online resources provide easy access from anywhere. If the library is to stay relevant to faculty research, this is an area we must expand into in a significant way.

For our university, the library is one branch in a tree of research data management support that spreads across our entire campus. By working with other campus providers such as IT, the Center for Advanced Computing, the Institute for Social and Economic Research, and the faculty themselves, the library is part of a collaborative system that strives to offer services and expertise to fill research data management needs throughout the full data lifecycle.

I believe libraries have a significant role to play in the development of research data management, but finding the right skill set will be difficult. We will need to look for expertise outside of the profession. Collaborations and partnerships will be essential to create a useful plan for the campus.

I see libraries assisting researchers with metadata issues, planning for ongoing access, archiving of data, and referral to additional services offered by the university or external entities.

I see research data management training belonging within the purview of libraries. I see data archiving, at least for some data, also within the purview of libraries; however, domain repositories and perhaps publishers also playing a part in archiving of research data. Regardless of where data are archived, a library should at least have a catalog of data sets generated under that university/college, which points to the archived location of that data set. Public funders may be pushing changes in the near future and are looking to libraries to help.

I think it will only increase.

Independent of infrastructure for data management: education and training (students), policies (interpretation and development), data management planning support, and partnering with archives.

It is already abundantly clear that research libraries can play a major and important role in supporting research data management, to the benefit of researchers and the library alike. Research libraries, in fact, *must* enhance their

services in this area if they wish to remain relevant to the needs of researchers at their institution, particularly in the sciences (who have increasingly turned away from libraries' traditional services).

Just as librarians have organized and made literature available, so we should do for research data so it will not be lost. We should continue to develop subject, IT, and preservation technologies and expertise.

Libraries are well positioned to provide support and education to faculty and students on data management planning, metadata creation, partners on research projects to provide the data management support. Development of a repository will require collaboration with other campus partners.

Libraries have always been stewards of information in various formats and this is another realm that we could play a role in organizing, preserving, and providing access to information.

Libraries will and should increasingly play a support role in that area because of their expertise in organizing and providing access to information. Universities and funding agencies must realize though, that not unlike long-term preservation of books and other publications, this will require major ongoing investments, whoever will be doing it.

Libraries will continue to become more engaged earlier in the research process and apply library science and archival principles to research data and the full research data lifecycle. In the future, we will acquire, describe, preserve, and facilitate access to data like we've done for centuries for paper. Datasets will be a part of our collections and nothing special.

Research data management is an integral part of the role libraries play in supporting their research community. The skills supplied by librarians are very similar to those required by RDM, and so libraries and librarians have the potential to advance the cause of RDM at their universities in the future. Librarians also occupy a unique niche within universities that enable them to serve a broad clientele without undue regard to the interests of finance or compliance.

Research libraries have assisted faculty with data in print, microfilm, and digital. This is part of reference, curation, collection development, and other core library work. New work will continue to extend from this work and should leverage that work for greater benefits and broader impacts made possible in the digital age. RDM activities and leveraging that work for additional benefits also offers new opportunities for collaboration with librarians and other researchers.

The library, as a neutral unit on campus, can provide service via the subject specialists that can help researchers archive, manage, and share data.

This will vary greatly from one institution to another, but in general, all research libraries will continue to gain knowledge of research data solutions (tools, repositories, etc.) that they and others offer, and connect researchers with these solutions.

Two main roles, which leverage our natural competencies. One is metadata creation and quality assurance. The other is our ability to preserve objects over long time periods, whether books or digital files. Everything else flows from these two roles, which are not duplicated elsewhere on campus.

Until now, we have only bought data sets and sent people to the appropriate departments for help. In the future, we will consult with our researchers about how best to shepherd their data through the data lifecycle.

We expect RDM and data curation to be seen as increasingly important. Collaboration around shared data curation services will be key to driving this forward. Libraries are one of many stakeholder groups that are critical to the success of data curation efforts.

We see it is a critical to growing the stewardship role that the library has for scholarship and research.

We see the key library role as collaborating with researchers earlier in the research process, to provide guidance in data collection and documentation that will best facilitate future accessibility of datasets, including, but not limited to, deposit at domain-specific or institutional repositories.

We see the library as playing a central role in supporting research data management, particularly in the organizational aspect.

We've always supported access and preservation. The fact that it's now bytes instead of books shouldn't change the service levels we offer.

ADDITIONAL COMMENTS

67. Please enter any additional information that may assist the authors' understanding of your library's research data management services. N=21

Although we have ten members on our Data Management Committee, only two of us have data management responsibilities formally in our job descriptions. It is difficult to make progress this way.

Critical for us is to avoid recreating the mistake we collectively made with institutional repositories, i.e., every institution setting up its own little silo. Data management needs to happen at much higher levels, at least regional if not national/international. Our intent is not to build an empty castle in our own yard.

In the last two and half years, the university has hired a new president, provost, vice president of research, and dean of arts and sciences. As a result of the changes in senior level management of the university, it has been difficult to gain traction for the support of research data management services. Moreover, the libraries believe the only way to move forward with RDM services is in partnership with the OVPR, which has a strong focus on a number of other issues due to administrative level changes. Consequently, it has been difficult to push forward with a strong RDM agenda by the libraries.

Library is relatively new to research data services; most services currently under development.

Our data management services are new and will evolve once the Data Librarian starts in July.

Our library is currently doing some strategic planning around digital content. We hope to leverage our work to demonstrate that the Data Working Group needs additional resources to maintain and expand the level of services that we currently provide.

Ours is a fairly decentralized RDM and will need to be a collaborative venture between the library, Office of Research, individual faculties, provost office and (possibly) central IT.

The library collaborates with the Office of Research Services to aid researchers.

The Libraries' Head of Curation and Preservation Services, and Director of Research; Head/Scientist, Program for Information Science are both very involved in national and international research projects related to digital curation and preservation, and research data management services.

The library's RDM services are in the planning stages and the development of RDM services and broader e-science initiatives are included in the library's five-year strategic plan. As you can see in our survey responses, some elements of RDM support are currently being offered through existing data expertise in the library.

We are in the process of establishing our data management services. We have hired the Digital Assets Librarian to provide leadership in this area, and formed a Data Management committee in late 2012.

We are just at the beginning stages of looking at offering RDM services. We've completed the ARL/Durspace e-Science Institute. We are working with our central IT on choosing and implementing a collaboration tool (e.g., Sharepoint, or something similar) in order to provide researchers with a place to work with and store their data securely. Canadian funding agencies do not yet mandate DMPs, but we are anticipating that they will in the near future. We hope to get a head of the game by training our librarians on RDM so as to assist researchers.

We are just at the stage of gathering resources, learning about things like the DMPTool, and starting to talk to subject librarians and systems librarians about how RDM support has to become part of their support portfolio.

We are planning our research data management services in collaboration with the campus Research Computing Center and IT Services (campus IT). Budget request will come jointly from the 3 units and new staff will be distributed between them as appropriate.

We are very active in the domain of research data curation, both in developing practices and operational services as well as exploring it as a research area itself. We have a data services unit in our library, a D2C2 research center, a Data Education Working Group, and most importantly, all of our librarians have responsibilities for data (per policy). For the repository, we are collaborating outside of the Libraries with a steering committee representing the interests of libraries, IT, the research office, sponsored program services, and our faculty. Our organization is complex and not easily mapable to some of the questions in this survey.

We are undergoing an overall library staff reorganization at this time that makes some things uncertain.

We do not have a formal data management services program or dedicated department, but we do offer some services on an ad hoc basis, such as data management plan and data archiving consulting. We are currently conducting a campus-wide study (survey and interviews) of research data needs on campus as well as exploring possible roles for the Libraries through an e-Research task force, and we believe we will have a better sense of future plans after both of these groups have reported out.

We have been collaborating with our OSP and central IT units since 2010 and will continue to do so for RDM services. We are beginning to work with departmental/school IT, as well.

We have just begun offering services and so most of them have not yet been taken advantage of.

We took part in the e-science institute and have a strategic agenda, which is in the early stages. We also do not have an institutional repository, but hope to have one soon.

We're adding data responsibilities into subject librarian job descriptions as they get reviewed, and we'll continue to do that.

RESPONDING INSTITUTIONS

University of Alberta
University of Arizona
Arizona State University
Auburn University
Boston University
Brigham Young University
University of British Columbia
Brown University
University of Calgary
University of California, Irvine
University of California, Los Angeles
University of California, San Diego
Case Western Reserve University
University of Chicago
University of Colorado at Boulder
Columbia University
University of Connecticut
Cornell University
Duke University
Emory University
University of Florida
Florida State University
Georgia Institute of Technology
University of Hawaii at Manoa
Indiana University Bloomington
University of Iowa
Iowa State University
Johns Hopkins University
University of Kansas
Kent State University
University of Kentucky
University of Louisville
McGill University
McMaster University
University of Manitoba
University of Maryland
University of Massachusetts, Amherst
Massachusetts Institute of Technology
University of Miami
University of Michigan
Michigan State University
University of Minnesota
Université de Montréal
University of Nebraska–Lincoln
University of New Mexico
University of North Carolina at Chapel Hill
North Carolina State University
Northwestern University
Ohio University
University of Oklahoma
University of Oregon
University of Ottawa
University of Pennsylvania
Pennsylvania State University
Purdue University
University of Rochester
Southern Illinois University Carbondale
University at Albany, SUNY
University at Buffalo, SUNY
Syracuse University
Temple University
University of Tennessee
University of Texas at Austin
Texas A&M University
Texas Tech University
University of Virginia
Virginia Tech
University of Waterloo
Wayne State University
Western University
University of Wisconsin–Madison
Yale University
York University