# Libraries and the Research Data Management Landscape

*Jodi Reeves Flores, Jason J. Brodeur, Morgan G. Daniels, Natsuko Nicholls, and Ece Turnator*

Across the world, organizations, institutions, and governments are increasingly recognizing the importance of research data management (RDM): the documentation, curation, and preservation of research data. RDM activities ensure long-term value and utility of research data for new analyses and replication of study findings. Stakeholders include research funding agencies, research institutions, and individual researchers. Because of the numerous requirements, mandates, techniques, and tools that make up the RDM landscape, it is difficult for stakeholders to carve out their niche. One of these stakeholders is the university research library. Research libraries have always offered a variety of research services, but as digital data became more prevalent and the need to manage them more pressing, some libraries began incorporating RDM into the research services offered. These RDM services result from the demands of government agencies or university administration; a perceived need to stay relevant in a changing, digital research world; or a thoughtful assessment of the needs of researchers. Often, libraries deal with a combination of all three of these influences and myriad other motivations, making it important to examine the possibilities for incorporating the library as a critical stakeholder in the RDM landscape.

When considering the library's role in RDM development, common themes include activities associated with conducting RDM needs assessment in user communities (Corrall et al. 2013, 646); policy development; advocacy, awareness, and training; advisory services; data repository development (Cox and Pinfield 2013; Jones et al. 2013); helpdesk services; and data management plan (DMP) development (Corrall et al. 2013, 646; Pinfield et al. 2014, 4). In addition to these activities, the RDM pyramid proposed by Lewis (2010) suggests a broader role for libraries and librarians, including integration of RDM into teaching at the undergraduate level and in schools

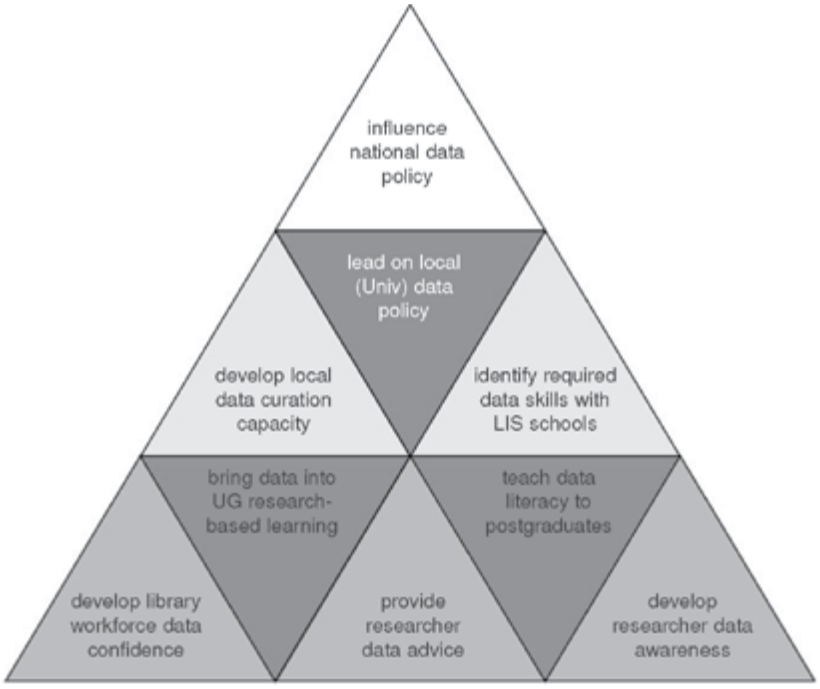


*Fig. 1: Research data management pyramid for libraries, as presented by Lewis (2010)*

of library and information science, as well as influence and participation in national policy development (figure 1).

As fellows of the Council on Library and Information Resources (CLIR) and the Digital Library Federation (DLF) who have experience within the researcher community and the library/RDM community, we are keenly interested in the role that libraries should assume in building coordination between funding agencies, institutions, and researchers. Research libraries moving into this space increasingly see themselves as major contributors to RDM activity in general and in the design of research data services in particular (Pinfield et al. 2014). However, libraries operate with finite resources, mandates, and limited researcher buy-in.

Taking into consideration these issues, the experiences of library staff from multiple institutions, and our hybrid research/library experiences, we advocate that libraries work to situate themselves in the wider RDM landscape so that they can make strategic decisions about their activities in RDM support development and work with those parties outside of the library best suited to address research needs. In this way, libraries can leverage both their relationship with university leadership and research support units, and their ability to disseminate knowledge regarding requirements, standards, and tools, to assume a leadership role in fostering a more collaborative and navigable RDM landscape for researchers.

## The RDM Landscape

Whether viewed at an institutional, national, or international level, RDM development relies on the collaborative and coordinated work of many engaged partners. Considering the role of the academic library in activities at any of these levels requires a general consideration of the current RDM landscape. Establishing the various stakeholders involved in RDM activities and characterizing their interests, roles, and responsibilities makes it possible to identify activities where the library is well situated to facilitate and coordinate RDM development.

### RDM Stakeholders

Although the spectrum of RDM stakeholders has been variously categorized in the literature (e.g., Erway 2013, 7; Jones et al. 2013, 3; Pinfield et al. 2014, 4), they can be assembled into four main categories (figure 2). This structure is not intended to denote or prescribe segregation between groups, but rather to align them according to similar interests, roles, and responsibilities in RDM.

As the primary funders of academic research, governments and funding agencies have an interest in maximizing the return on their investments. Properly managed and shared data have the potential to yield manifold benefits when reused in primary research,

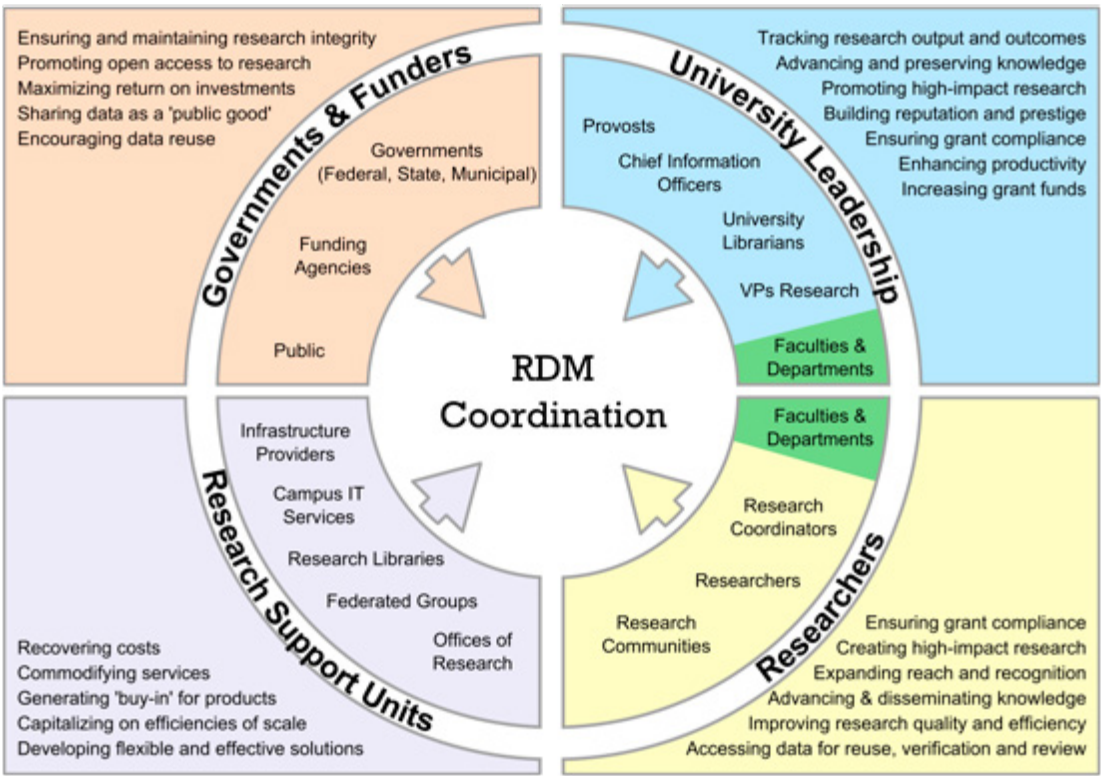


*Fig. 2: Partners in RDM development, categorized into four general stakeholder groups. Individual stakeholder units are identified in the central ring, with general group interests listed in the adjacent boxes.*

follow-up, and synthesis studies, as well as in interdisciplinary and data-intensive research (Heidorn 2011, 662; Pryor 2012, 1). With varying degrees of response, funders encourage RDM activities to ensure that appropriate data stewardship and sharing are embedded in the research process through one or a combination of high-level government recommendations, requirements for submitting RDM plans with grant proposals, and the sharing of research data products where applicable and appropriate.

The RDM interests of university leadership—university provosts, chief information officers, vice presidents of research, and university librarians—reflect their responsibilities to jurisdictional government agencies and funders, as well as to their institutional researchers, students, and community. These interests include ensuring compliance, advancing the creation and preservation of knowledge, tracking research output, and building the institution's reputation and prestige. Additionally, members of the university leadership group may be researchers themselves and, as such, identify strongly with the research culture in their specific discipline. In view of these factors, the RDM approaches and activities initiated at the level of university administration reflect a combination of requirements imposed by national governments and funding agencies, the awareness and perceived importance of RDM by those in leadership positions, and developments occurring within various departments and service units at the institution (Erway 2013, 7).

As the producers of data and disseminators of knowledge, researchers and their associated communities, departments, and faculties regard RDM as a means of ensuring compliance with funders, increasing the efficiency and quality of their research, and advancing knowledge in their field of study (Erway 2013, 10). The interests, challenges, and needs of researchers with respect to RDM vary by discipline, as well as by institutional and national context.

Addressing the diverse interests and requirements of the other RDM stakeholders takes the combined efforts of several research support units at institutional, national, and international levels. RDM partners at the institutional level commonly include the library, information technology (IT) services, and the office of research, as well as other relevant internal and third-party service providers (Jones et al. 2013, 3; Pinfield et al. 2014, 4). At national and international levels, RDM includes the collaboration of infrastructure providers and is both discipline-specific and cross-disciplinary. These parties have an interest in developing effective RDM solutions and services that are highly used, scalable, and sustainable. Such development needs to commodify or ensure long-term support for services and requires a clear definition of RDM roles and responsibilities among units.

### The Need for RDM Coordination at Varying Levels

RDM development takes place over many levels, ranging from international collaborations to national and institutional policy development to efforts within individual research groups. Although each of the previously introduced stakeholder groups has a significant role in RDM development activities, their interests, involvement, and contributions vary with the level. Furthermore, the divergent interests and expertise among these groups present obstacles to the creation of a comprehensive, cohesive data stewardship and sharing ecosystem. Where these differences result in substantial challenges to processes or practices, it is necessary for one or more of the partner groups to assume a coordinating and mediating role in RDM development (i.e., the central position in figure 2). Given the various levels of development and their diverse circumstances and dynamics, opportunities for groups to lead coordination efforts flourish.

For instance, the development of government funding agency RDM policies and requirements is uneven, both within and between nations. Variation in these policies has consequences for all stakeholder groups and presents an important area for RDM coordination at a number of levels.

Additionally, in the United States, data management and sharing policies have been implemented to some extent for all major federal grant funders, including the National Science Foundation (NSF), the National Institutes of Health (NIH) and, most recently, the Department of Energy (Dietrich et al. 2012; U.S. Department of Energy 2014). Although most divisions and programs require DMPs for grant proposals and data access upon study completion, a number of inconsistencies remain among agencies (Dietrich et al. 2012). As stated in the U.S. Department of Health and Human Services Public Access Plan (2015), there is a lack of common standards for data management and archiving, as well as a lack of common requirements and enforcement practices for data sharing across agencies. In addition, a policy comparison by Dietrich and colleagues (2012) highlighted inconsistencies in metadata standards used among NSF directorates and programs, which have led to confusion for the researcher.

Like federal funding agencies, institutions differ in their development and implementation of RDM policies. Although nationally driven efforts have led to widespread institutional policy implementation in countries such as Australia, development in other countries—including the United States and the United Kingdom—is ongoing and often uneven (Horton and DCC 2014). The precise nature of the policy development process differs between institutions because their stakeholder interactions, characteristics, and interests provide a unique context, but recent explorations of these efforts have revealed a number of commonalities.

### Supporting Researcher RDM Needs

A great number of coordination opportunities for enhancing data sharing and stewardship activities emerge because, as researchers

commonly identify more closely with their research community than with their institutions, RDM support needs continue to vary among disciplines and within institutions (Akers and Doty 2013, 14; Cox and Pinfield 2013, 19).

Perhaps the most significant opportunity for coordination in managing and sharing data occurs in addressing the considerable variation among academic disciplines in their treatment of and approaches to data organization, documentation, sharing, and preservation. The dimensions of these differences include the quantity, structure, and format of data; the accepted metadata standards in the field; the researchers' interests and requirements to manage and share data; and the discipline-specific norms for sharing data (Cox and Pinfield 2013, 19; Harley et. al. 2010, 4). Disciplines such as astronomy, genomics, ecology, and quantitative social sciences operate within a well-developed culture of data stewardship and sharing, with established metadata standards, tools, and data repositories to support these activities. Conversely, the development of research data standards, tools, and norms has been slower for other disciplines, providing significant opportunities for RDM stakeholder groups to facilitate and coordinate such efforts. For these fields, libraries and library staff—particularly subject specialists—can play an essential role by preparing scholars for new research requirements, such as DMPs, and providing tools and services to support data stewardship activities. Indeed, these actions could help avoid regretful statements such as "had the librarians been involved earlier in the life cycle of the pilot data . . . data preparation and workflows could have been adjusted to accommodate eventual data deposit" (Newton et al. 2011, 15).

Despite the fact that many researchers do understand the importance and academic value of data sharing, there remain several barriers to providing access to data, and this is where many coordination possibilities for enhancing data sharing and stewardship activities lie. By addressing challenges and working to remove barriers, RDM services can enable individual investigators to easily, quickly, and effectively share their primary research data. Such an advancement has the potential to greatly enhance transparency and efficiency, and to foster positive impacts on knowledge advancement in all fields of study (California Digital Library 2014).

One challenge involves determining an appropriate repository for long-term data preservation and sharing. Another is imposed on the data stewards who assist researchers in choosing the right repository. Baker and Yarmey (2009) discuss data stewardship as the tending of multiple related repositories from a big-picture perspective, requiring a broad knowledge and solid understanding of repositories' different features, functionality, fees, and any limits on the number or size of data sets that can be deposited in each repository (MetaArchive Cooperative Outreach Committee 2015). A consolidated registry, re3Data, contains information for more than 1,130 data repositories (re3data.org team 2015).

Leading science, technology, engineering, and mathematics

publishers are increasingly adopting the practice of publishing data(set) papers in conventional journals and data journals (Candela et al. 2015); therefore, publishers provide authors with instructions for a data set deposit into select trusted/approved repositories. For instance, the *Journal of Environmental Quality*, published by the American Society of Agronomy, currently states in its author guidelines that "dataset papers or collections of datasets integral to a dataset paper can be hosted on the ACSESS internal servers or an appropriate external repository (institutional repositories or another acceptable repository such as Dryad)" (ACSESS Digital Library 2015). The question for both researchers and data consultants is which repository option is better to allow for data accuracy, discoverability, and usefulness.

Some researchers rely on multidisciplinary institutional repositories, which provide publication-related materials from multiple subject areas within a single organization. Researchers in other fields recognize and use common discipline-specific repositories that collect and preserve intrinsically domain- or discipline-oriented research. Prominent disciplinary repositories include the Dryad Digital Repository for scientific and medical publications, The Archive of the Indigenous Languages of Latin America, the Inter-University Consortium for Political and Social Research, and The Digital Archaeological Record (tDAR). In some cases, a consortium of institutions hosts these repositories, while in other cases, a single institution supports the repository. These varied approaches demand an advanced RDM coordination effort.

## The Academic Library as a Leader in RDM

Because of their expertise in research methodology and knowledge retention, academic librarians can offer relevant leadership in RDM efforts within their universities. Although collaboration across the institution is key in developing an RDM program, research libraries play an essential coordination role in the process.

Campus stakeholder groups such as university administrators, researchers, and research support units all have an interest in how RDM services are designed for their institution. It is the library's unique position, as both a facility with staff who have expertise in many of the issues surrounding RDM and a campus-wide service with relationships among these many stakeholder groups, that favorably positions it to lead the RDM effort (Erway 2013; Shaffer 2013): "The library is well situated to be a key player in data management, curation, and preservation, given its extensive experience with selection, metadata, collections, institutional repositories, preservation, curation and access" (Erway 2013, 10–11).

Academic libraries have a history of provisioning data for research use, giving many librarians a familiarity with the reuse requirements and concerns surrounding research data. Humphrey notes that data services librarians "often assist with locating data, interpreting data documentation, retrieving data files, and providing

the data in a format that can be directly loaded into analytic software" (2014). Experience helping researchers use data sets can be leveraged in the provision of data management services. In addition, academic librarians are masterful at designing and delivering educational content tailored to the research practices of members of various disciplines, at varying levels of expertise. Their fluency and flexibility as instructors equips them to educate members of the university community in RDM issues.

Research library staff have existing relationships with researchers, other research support units, and leaders across the university, making them well situated to coordinate services, such as RDM services, that are offered horizontally across the institution (Humphrey 2014). As institution-wide resources, libraries can coordinate services across disciplines, helping researchers in many disciplines meet best practices in data management. Most researchers are amenable to receiving librarians' expert assistance with multiple aspects of data management, particularly given their own limited resources and the many other demands on their time and energy (McLure et al. 2014). With their connections to faculty and their disciplinary knowledge, subject specialists can raise awareness of RDM services across campus—once they themselves are provided with a background on RDM. Perhaps this is why many academic libraries, including those at the University of Michigan, Purdue University, Baylor University, and the University of Maryland, have focused on data education and training for their own subject specialists before they reach out to academic departments and researchers (Zilinski et al. 2013). With the help of data education and training to develop new skills and knowledge appropriate to data management responsibilities, ongoing communication between subject and data curation specialists in libraries can facilitate the flow of information about researcher needs and RDM capabilities throughout the university as a whole.

In addition, academic librarians can coordinate RDM efforts beyond their institution, sharing lessons learned in professional groups and building partnerships with other universities to develop and test RDM solutions. A number of cross-institutional partnerships have been developed via the E-Science Institute (sponsored by the Association for Research Libraries [ARL], Digital Library Federation [DLF], and DuraSpace); the DLF E-Research Network; the Association of College and Research Libraries (ACRL) Data Management Working Group; the New England Collaborative Data Management Curriculum; and the Virginia Data Management Bootcamp, to name only a few. Librarians are taking advantage of these and other interinstitutional collaboration opportunities as they develop RDM capabilities.

An important part of the coordination work needed to develop an RDM program is advocacy on behalf of researchers as key stakeholders and toward data stewardship as a goal. Advocacy is a multidirectional process: learning about researcher needs and requirements and taking them to university leaders to plan for RDM services, while communicating campus policy back to researchers and research support units. Librarians need to start discussions

about RDM across campus stakeholder groups, which may have their own preexisting goals regarding RDM. Conducted by ARL, a review of a group of library, IT, and university strategic plans among member institutions found that RDM goals cut across both library and IT plans (ARL 2014). Through outreach, libraries can leverage this mutual interest to build partnerships to develop policy and planning at the university level.

## Research Data Management Services

Table 1 describes many of the services that can make up RDM offerings within an academic institution, lists the stakeholders concerned with each offering, and suggests activities potentially undertaken by libraries to coordinate these services. This list is neither exhaustive nor prescriptive. It does not capture the full range of services considered RDM. It will expand and change as the RDM landscape changes. It is not intended to tell research libraries what services they *should* offer. The suite of RDM services offered to a given campus community should be tailored to the needs of that community's researchers, in consideration of the organizational and technological resources available.

One major question that arises as RDM teams develop their suites of services on campus is whether to provide an institutional data repository. Although some libraries find that data sets fit easily into the infrastructure of an existing institutional repository, others consider building or licensing a standalone data repository for research data sets produced by the campus community. Some institutions have taken this route, but survey results published by the Digital Curation Centre (DCC) in 2014 underscored many institutions' preference for collaborating with other organizations to provide a research data repository (Whyte 2014). In alignment with this sentiment, some institutions and support groups opt to create federated, shared repositories. Examples include development led by the DCC in the United Kingdom and work under way in the United States by the California Digital Library and Texas Digital Library. There are, in addition, initiatives among the library and the research community for sharing metadata across institutional and disciplinary repositories. For instance, the SHARE project, cofounded in 2013 by ARL, the Association of American Universities, and the Association of Public and Land-grant Universities, is building notification tools and services to make research outcomes and outputs widely accessible, discoverable, and reusable across repositories.

Although institutional data repositories may be part of RDM services offered on campus, we believe they should be offered primarily as mechanisms to preserve and publish data that do not already have a natural disciplinary home. Institutional repositories focusing on more traditional, text-based scholarly output may seek to be comprehensive, collecting all research publications generated at their institution. However, they cannot offer the features and visibility to researchers that disciplinary repositories specializing in data sets for

| Service | Key Stakeholders | Library's Coordination Role |
|---|---|---|
| Access control | Researchers, research support units | Advise on data embargoing and access control issues |
| Awareness of RDM mandates and services | University leadership, researchers, research support units | Coordinate with research office staff and administrators across campus to raise awareness of RDM mandates and services |
| Data citation | Researchers, research support units | Provide persistent identifiers, including digital object identifiers (DOIs) for data sets |
| Data documentation | Researchers, research support units | Help researchers determine how best to document their data at the beginning of a project, following disciplinary standards |
| Data management planning | Governments and funders, researchers, research support units, university leadership | Provide outreach to university leadership and research support units to develop data management plan assistance processes on campus; connect researchers with local and disciplinary resources to meet funding agency requirements |
| Hosting data | University leadership, researchers, research support units | Work with university leadership and research support units to provide infrastructure for hosting data (institutional data repositories), or helping connect researchers with available infrastructure (disciplinary repositories) |
| Intellectual property and copyright | Researchers, research support units | Provide guidance on intellectual property and copyright matters surrounding research data |
| Preservation | Researchers, research support units | Advise on appropriate data formats for preservation, preparing data sets for long-term preservation |
| Privacy and confidentiality | Researchers, research support units | Advise researchers and research office staff on privacy and confidentiality issues in data management |
| Repository selection | Researchers, research support units | Help individuals select trusted digital repositories for preserving data sets, whether those are disciplinary repositories or institutionally managed repositories |
| RDM workshops | Researchers, research support units | Communicate best practices developed by the RDM community to groups of researchers |
| Scholarly impact | Researchers, research support units | Promote mechanisms to track the impact of data sharing—downloads, citations, etc. |
| Scholarly output | Researchers, research support units | Help connect data sets to other scholarly output through linked data and citation mechanisms |

*Table 1: Research data management (RDM) service offerings*

a specific audience are able to provide (e.g., tDAR and the Archaeology Data Service for archaeological data). On the other hand, institutional data repositories perform an important service by archiving materials related to a research project that are not within the collecting focus of a disciplinary data repository, providing important context to research data (Strasser 2014). Although a library may decide to offer an institutional data repository as a core RDM tool in a particular university community, it is unwise to make that tool comprehensive of the entire output of research data sets on campus.

Academic libraries may choose to configure RDM services in any number of ways, given the array of tools and services that support RDM. Most campus libraries today are structured by discipline to support academic departments. However, RDM support requires activities that cut across this departmentally aligned organizational structure. Just as many academics find themselves challenged to adjust to new requirements in order to continue funding their research (Akmon et al. 2011, 330), libraries are challenged to develop RDM support that cuts across their own organizational structure. As they establish partnerships to offer RDM services in collaboration with other campus stakeholders—including offices of research, offices of sponsored programs, technology service units, research compliance offices, and academic departments—they must work broadly across disciplinary and functional units (Humphrey 2014). This breadth renders the development of RDM services an organizational challenge within research institutions.

## Activities Supporting the Development of RDM Service Offerings

Given the assortment of potential and ongoing RDM activities, many libraries are currently in a state of redefinition with a reduced budget (Lewis 2010; Lyon 2012; Shaffer 2013). Needing to create services that will be truly useful to the campus community, libraries must carefully consider their role in developing and offering RDM services. Many RDM teams undertake a number of activities to structure their service offerings. The order in which these structuring activities take place varies a great deal among universities. Some may choose to complete each of these activities prior to rolling out RDM services to the campus community, while others do them concurrently. The "right" way to combine these structuring activities with service offerings can only be decided contextually, by considering the needs of researchers, readiness and capabilities of on-campus partners, and available infrastructure on a particular campus.

As figure 3 (adapted from Akers et al. 2014) illustrates, universities have taken different paths to an RDM program, including such milestones as providing data services to researchers, building an institutional repository, performing assessment activities, offering RDM services, and providing a data repository. These components came at different times at the universities surveyed by Akers and colleagues, and not all universities completed each of these milestones

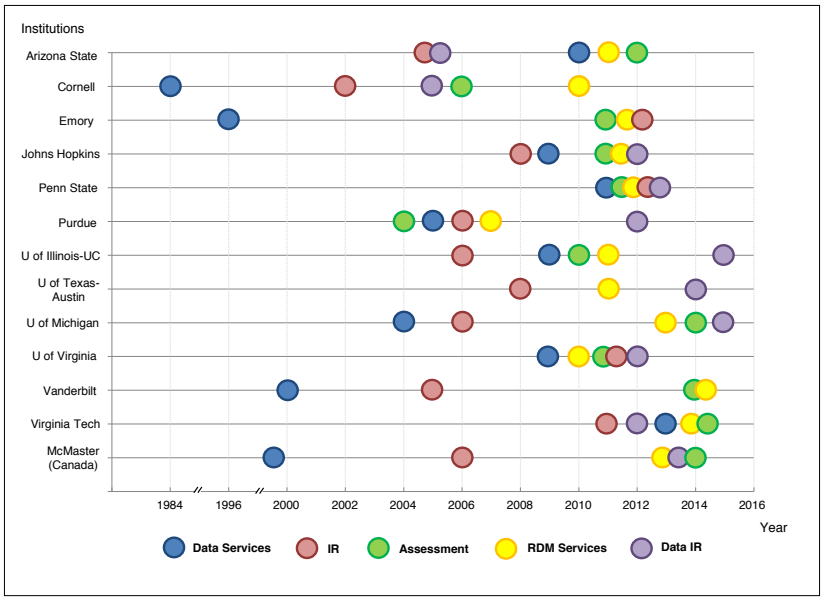


*Fig. 3: Institutional timelines for building RDM services (adapted from Akers et al. 2014; timelines in five criteria for the eight institutions are revised, and five institutions are added where authors are currently affiliated)*

in their development of RDM services. Each path is unique to the institution itself.

Many library RDM teams tend to undergo a number of common activities in the process of developing RDM services; the activities are often iterative, with the team revisiting them multiple times. These common milestones include building partnerships, conducting an environmental scan, assessing needs, and forming policy. Although many activities are inherent in building and providing RDM services, there is no one linear path.

### *Building Partnerships*

Within a campus community, many stakeholders play important roles in planning RDM services. In their discussion of eight research universities' approaches to RDM, Akers and colleagues note that RDM activities have been initiated by reaching out to numerous groups on campus, "with university research offices, advanced research computing facilities, and campus information technology departments being prominent library partners" (2014, 184). Campus research offices such as an office of sponsored programs are essential partners, as they support grant proposal preparation and submission. Such partnerships are likely to be new, but very important, relationships for libraries building RDM services. Partnerships with campus information technology and high-performance computing centers are also vital to establishing RDM services, as these support units provide data storage and computational resources that enable data capture and analysis for research. RDM librarians should work with these units to establish mechanisms for data transfer.

As libraries develop and propose RDM services, identifying and

creating relationships with these and other stakeholders become essential. Partners' ongoing roles may range from advisory to participatory, and they can include working to develop campus policy regarding the disposition of research data sets, advising the library's RDM team, provisioning computing infrastructure needed for data management, reviewing DMPs, and joining the RDM team as active members supporting specific projects.

### *Conducting an Environmental Scan*

Often an ongoing process, an environmental scan that supports RDM services may take place both internally and outside the institution. On campus, an internal environmental scan of the resources available and in use to support RDM in various departments serves several functions. Not only can it help identify potential partners across the university that are already considering or supporting RDM to a subset of those on campus, but also it can help map existing and potential RDM-related services and resources on campus. An external environmental scan helps RDM service providers keep current on the topic, learn from peers through reading the data management literature, participate in online discussion forums, and attend conferences organized on the subject of data management. Regular environmental scans outside the institution are essential to keep RDM librarians aware of new developments in RDM and opportunities for collaboration. Librarians can then convey these developments to campus stakeholders through ongoing education, training, and outreach.

### *Assessing Needs*

Assessments of needs, undertaken primarily through surveys and interviews with researchers on campus, have taken place at many universities to help RDM librarians determine how campus constituents manage research data and how RDM services might help. Resources such as the DCC's Data Asset Framework and the Data Curation Profiles Toolkit developed at Purdue University are openly available to help RDM service developers collect information through interviews with researchers about the data sets they produce and the resources they currently use to manage them. RDM assessment surveys reveal researchers' awareness of DMPs and identify individual and community practices of data documentation, sharing, and preservation. Surveys and interviews may also suggest departments or disciplines on campus that are potential partners for an RDM pilot project, often because of distinct data support needs discovered through the needs assessment process (e.g., Nicholls et al. 2014).

Assessment can identify and document underlying concerns of researchers, such as the degree to which DMP quality affects funding decisions and the way in which post-award compliance monitoring of data management would be carried out (Lalwani 2015). Some researchers have had NSF proposals conditionally accepted until their DMPs are revised and resubmitted, or have been informed by reviewers that their DMPs must be improved before the proposal can

be funded (personal communication, Sayeed Choudhury, 10 March 2015). In response to the varied requirements of DMPs, a number of libraries have proactively developed discipline- and NSF directorate-specific resources, effectively coordinating the RDM interests of funding agencies and researchers by offering more tailored RDM support to principal investigators (Nicholls et al. 2014).

### *Forming Policy*

Policy development is a cross-institutional process, but by initiating the conversation about RDM policy, libraries can ensure that they have a voice in the discussion (Erway 2013). Although many RDM providers feel that policy development must have a top-down component involving the university leadership, they observe that bottom-up engagement is also essential to build buy-in among stakeholders (Pinfield et al. 2014). With their cross-institutional connections throughout campus, libraries are in a good position to help manage and represent stakeholder interests to a policy development group.

Qualitative analysis conducted by Pinfield and associates (2014) showed that the policy development process typically involves consultation with RDM stakeholders across the institution, that it is often iterative in nature, and that policymakers commonly adapt elements from other institutions to suit their own needs. In many of the U.K. institutions investigated in the study, libraries and librarians actively participated and facilitated RDM policy development (Pinfield et al. 2014). These findings support recommendations made by previous authors (Erway 2014; Jones et al. 2013; Lewis 2010) that libraries should be active leaders and coordinators in these efforts. The University of Alberta's Research Policy (2015) is an excellent example of the library's potential role in institutional data policymaking and RDM support. As part of the policy framework, the library assumed a leadership role as one of the offices of administrative responsibility supporting the Research Records Management and Preservation Guidelines (University of Alberta 2013).

Deciding which of these RDM activities to undertake first presents a daunting task to university library leaders and librarians. Many are taking advantage of federated RDM support groups and other resources beyond the institution for guidance in their efforts. Two such federated approaches are the E-Science Institute (ARL/DLF/Duraspace) and the DLF E-Research Peer Network Program. More than 50 college and research universities have participated in either or both of these two programs over the last few years. Through participation in these programs, librarians built connections with peers, including CLIR fellows, and took inspiration from advances in RDM made by other academic libraries. More importantly, though, participation helped build dialog between campus leaders and librarians and often resulted in new hires designated to work on RDM service development and implementation. The E-Science Institute, in particular, required representation from library leadership, librarians, and campus IT from each institution. Participation also served as a training opportunity for liaison librarians with new data responsibilities and offered

a way to identify potential partnerships and investigate relationships among stakeholders through interviews of researchers and campus administrators. Librarians at Montana State University and the University of Manitoba mentioned their attendance at the E-Research Network (DLF) specifically as playing an important role in helping them bolster their support for RDM (Clark 2015; Ishida 2014).

## How Can RDM Services Help Libraries Enlarge Their Role?

Academic libraries find more and more opportunities to provide services throughout the different phases of the research life cycle: RDM is one of these areas, as are digital humanities, digital projects support, and others, where libraries can help academics as they produce and disseminate research. We believe that RDM offers an opportunity for libraries to reformulate their role in the life of the university.

Libraries offering RDM services can have a great impact on their campus communities by supporting communication among researchers, enhancing knowledge of the data life cycle, providing disciplinary and institutional resources, and emphasizing the importance of documentation of data sharing (McLure et al. 2014, 158). Within and beyond the institution, many libraries and associated entities work to bring together the various RDM stakeholder groups to create collaborative and cooperative solutions (Tenopir et al. 2012). At the institutional level, the establishment of hybrid positions for researchers in the library has removed barriers between the librarians and the researchers when they are developing RDM services and solutions, as illustrated by the Sloan Foundation-supported Data Curation Fellows for the Sciences and Social Sciences program that places academics in research libraries (CLIR 2013). Through a cohort of postdoctoral fellows placed at universities around the United States and Canada, more than 20 participating universities are working to develop their RDM programs while taking advantage of the ongoing learning activities of fellows.

Beyond the institution, libraries play an active role in developing national and international federated RDM support groups, which have been formed to encourage data stewardship and to share efficiencies of scale. National groups such as the Australian National Data Service and the United Kingdom's DCC seek to support and enhance national research data environments by providing a range of resources, services, and tools that facilitate data curation, connection, discovery, and reuse. At an international level, communities such as the Committee on Data for Science and Technology and the recently formed Research Data Alliance seek to improve the quality and accessibility of data across technologies and countries in science and across all disciplines, respectively.

Without doubt, the library is well situated to be a key player in data management, curation, and preservation. Because best practice in RDM dictates that research data be actively curated, not just

stored or backed up, librarians are positioned to train and assist researchers in long-term curation of data (Erway 2013, 10–12). Recent studies show that academic researchers, however, are uncertain of their responsibilities regarding data management and unsure where to seek help (Mischo et al. 2014; Parham et al. 2012; Steinhart et al. 2012). Although RDM outreach efforts can alleviate this problem over time, several questions remain. Are academics ready and open to being trained by librarians? What are the barriers to further engaging researchers at the faculty level? Can programs like CLIR's Postdoctoral Fellowship Program help bridge the library-researcher gap where it exists? Perhaps the biggest challenge in all this is to change the perceptions of overworked academics who have no time or desire to undergo any kind of training and view depositing their work or data in a repository as nothing more than an administrative function (Jones 2007, 9, 16–17; Pinfield et al. 2014). The library can change these perceptions by actively helping researchers navigate the requirements, demands, and tools that make up the RDM landscape, particularly when it comes to the organization, preservation, and sharing of research data.

The library can play a key role in the move toward research data stewardship, one of many changes to research practice enabled by digital technologies. However, although "many research processes have transitioned from print to digital, the standards and training used to ensure research integrity have not" (Coates 2014, 598). Librarians can help researchers navigate these "changing cultures of research." As Coates argues, "culture change is complex and slow, so we first need to understand which research practices are effective in promoting integrity and then determine how to encourage and reward those practices" (599).

In this same context, it is important to note the power and influence of established networks of field-specific social influence among peers, mentors, and senior scholars that often determine the amount of trust given to a certain repository and the research data that the repository accumulates (Roland and Lee 2013; Yakel, et al. 2013; Yoon 2014; Zimmerman 2007). Despite entrenched challenges in a changing landscape, libraries play a role—promoting the principle that data sharing enhances the integrity of research by permitting results to be reproduced and reexamined, directly supporting the academic enterprise. By offering data management training and services to researchers, particularly early career researchers and graduate students, libraries can help encourage a cultural shift toward effective data stewardship and value to data sets, making them meaningful and useful digital objects into the future.

Librarians do need to recognize that they are not the only group increasing their involvement in the research life cycle. Funding agencies, through DMP requirements, and publishers, through software managing the publication process, are other stakeholders recognizing and acting on parts of the research life cycle beyond the point of publication. Librarians, however, are offering their services from the position of a trusted institution and from an embedded understanding of

the university context. If libraries actively take advantage of this position, they can possibly transform their relationships with the larger institution they serve and with researchers, tying RDM to the changing role of the academic library. In addition to the important role of educating researchers about the RDM landscape, libraries can, and should, take on the role of advocate. Building on knowledge gained of researcher needs and requirements, libraries can take these issues directly to campus administrators in order to plan for RDM services. And they can do this while communicating campus policy back to members of academic departments, helping to complete the RDM feedback loop. Because of their ability to help coordinate between different stakeholders and foster collaborations, no matter what RDM or other research services individual libraries choose to offer, academic libraries should remain a vocal and critical part of the discussion.

## References

*All URLs are current as of September 1, 2015*

ACSESS Digital Library. 2015. *Dataset Paper Author and Reviewer Instructions*. Available at https://dl.sciencesocieties.org/publications/jeq/author-instructions-datasets.

Akers, Katherine G., and Jennifer Doty. 2013 Disciplinary differences in faculty research data management practices and perspectives. *International Journal of Digital Curation* 8 (2): 5–26. Available at http://ijdc.net/index.php/ijdc/article/view/263.

Akers, Katherine G., Fe C. Sferdean, Natsuko H. Nicholls, and Jennifer A. Green. 2014. Building Support for Research Data Management: Biographies of Eight Research Universities. *International Journal of Digital Curation* 9 (2): 171–191. doi:10.2218/ijdc.v9i2.327.

Akmon, Dharma, Ann Zimmerman, Morgan Daniels, and Margaret Hedstrom. 2011. The Application of Archival Concepts to a Data-Intensive Environment: Working with Scientists to Understand Data Management and Preservation Needs. *Archival Science* 11 (3–4): 329–348. Available at http://link.springer.com/article/10.1007%2Fs10502-011-9151-4.

ARL (Association of Research Libraries). 2014. *Report of the Association of Research Libraries Strategic Thinking and Design Initiative*. Available at http://www.arl.org/storage/documents/publications/strategic-thinking-design-full-report-aug2014.pdf.

Baker, Karen S., and Lynn Yarmey. 2009. Data Stewardship: Environmental Data Curation and a Web-of-Repositories. Remote Repositories—Distant Origin. *International Journal of Digital Curation* 4 (2): 12–27.

California Digital Library. 2014. Dash: About Dash. University of California Curation Center. Available at https://dash.library.ucsc.edu/xtf/search?smode=aboutPage.

Candela, Leonardo, Donatella Castelli, Paolo Manghi, and Alice Tani. 2015. Data Journals: A survey. *Journal of the Association for Information Science and Technology*. Available at http://onlinelibrary.wiley.com/doi/10.1002/asi.23358/abstract.

Clark, Jason. 2015. About the DLF E-Research Network. DLF Events, Blog post, March 2, 2015. Available at http://www.diglib.org/archives/8010/.

Coates, Heather. 2014. Ensuring Research Integrity: The Role of Data Management in Current Crises. *College & Research Libraries News*, 75 (11): 598–601. Available at http://crln.acrl.org/content/75/11/598.full.

Corrall, Sheila, Mary Anne Kennan, and Waseem Afzal. 2013. Bibliometrics and Research Data Management Services: Emerging Trends in Library Support for Research. *Library Trends* 61 (3): 636–674. Available at http://muse.jhu.edu/login?auth=0&type=summary&url=/journals/library_trends/v061/61.3.corrall02.html.

Council on Library and Information Resources (CLIR). 2013. CLIR Receives Sloan Foundation Grant for Data Curation Fellows. News release, April 1, 2013. Available at http://www.clir.org/about/news/pressrelease/sloan-data-curation-award.

Cox, Andrew, and Stephen Pinfield. 2013. Research Data Management and Libraries: Current Activities and Future Priorities. *Journal of Librarianship and Information Science* 46 (4): 299–316. Available at http://lis.sagepub.com/content/46/4/299.

Data Asset Framework. Available at http://www.data-audit.eu/.

Data Curation Profiles Toolkit. Available at http://datacurationprofiles.org/

Dietrich, Dianne, Trisha Adamus, Alison Miner, and Gail Steinhart. 2012. De-mystifying the Data Management Requirements of Research Funders. *Issues in Science and Technology Librarianship* 70.

Erway, Ricky. 2013. *Starting the Conversation: University-wide Research Data Management Policy*. Dublin, OH: OCLC Research. Available at http://www.oclc.org/content/dam/research/publications/library/2013/2013-08.pdf.

Harley, Diane, Sophia Krzys Acord, Sara Earl-Novell, Shannon Lawrence, and C. Judson King. 2010. *Assessing the Future Landscape of Scholarly Communication: An Exploration of Faculty Values and Needs in Seven Disciplines*. Berkeley, CA: The Center for Studies in Higher Education. Available at http://escholarship.org/uc/cshe_fsc.

Heidorn, P. Bryan. 2011. The Emerging Role of Libraries in Data Curation and E-science. *Journal of Library Administration* 51 (7–8): 662–672. Available at http://www.tandfonline.com/doi/abs/10.1080/01930826.2011.601269#.VbpjALcwdaQ.

Horton, Laurence, and DCC. 2014. *Overview of UK Institution RDM Data Policies*. Digital Curation Centre. Available at http://www.dcc.ac.uk/resources/policy-and-legal/institutional-data-policies.

Humphrey, Chuck. 2014. Are Libraries Organized to Provide Research Data Management Services? Preserving Research Data in Canada, Blog post, December 10, 2014. Available at http://preservingresearchdataincanada.net/2014/12/04/are-libraries-organized-to-provide-research-data-management-services/.

Ishida, Mayu. 2014. Building a Community of Practice for Research Data Services. Blog post, October 9, 2014. Available at http://connect.clir.org/blogs/mayu-ishida/2014/10/09/building-a-community-of-practice-for-research-data-services.

Jones, Catherine. 2007. *Institutional Repositories: Content and Culture in an Open Access Environment*. Oxford, England: Chandos Publishing.

Jones, Sarah, Graham Pryor, and Angus Whyte. 2013. *How to Develop Research Data Management Services—A Guide for HEIs*. DCC How-to Guides. Edinburgh: Digital Curation Centre. Available at http://www.dcc.ac.uk/resources/how-guides.

Lalwani, Leena. 2015. NSF DMP Review Cases at University of Michigan. E-mail interview. March 2, 2015.

Lewis, Martin. 2010. Libraries and the Management of Research Data. In *Envisioning Future Academic Library Services*, edited by Sue McKnight, 145–168. London: Facet Publishing.

Lyon, Liz. 2012. The Informatics Transform: Re-Engineering Libraries for the Data Decade. *International Journal of Digital Curation* 7 (1): 126–138. Available at http://ijdc.net/index.php/ijdc/article/view/210.

McLure, Merinda, Allison V. Level, Catherine L. Cranston, and Beth Oehlerts. 2014. Data Curation: A Study of Researcher Practices and Needs. *portal: Libraries and the Academy* 14 (2): 139–164. Available at https://muse.jhu.edu/login?auth=0&type=summary&url=/journals/portal_libraries_and_the_academy/v014/14.2.mclure.pdf.

MetaArchive Cooperative Outreach Committee. 2015. *Getting to the Bottom Line: 20 Cost Questions for Digital Preservation*. Available at http://www.metaarchive.org/cost-questions.

Mischo, William H., Mary C. Schlembach, and Megan N. O'Donnell. 2014. An Analysis of Data Management Plans in University of Illinois National Science Foundation Grant Proposals. *Journal of eScience Librarianship* 3 (1): 31–43.

Newton, Mark P., Christopher C. Miller, Marianne S. Bracke. 2011. Librarian Roles in Institutional Repository Data Set Collecting: Outcomes of a Research Library Task Force. *Libraries Research Publications*. Paper 122. http://docs.lib.purdue.edu/lib_research/122.

Nicholls, Natsuko, Sara M. Samuel, Leena N. Lalwani, Paul F. Grochowski, and Jennifer A. Green. 2014. Resources to Support Faculty Writing Data Management Plans: Lessons Learned from an Engineering Pilot. *International Journal of Digital Curation* 9 (1): 242–252. Available at http://www.ijdc.net/index.php/ijdc/article/view/315.

Parham, Susan W., Jon Bodnar, and Sara Fuchs. 2012. Supporting Tomorrow's Research: Assessing Faculty Data Curation Needs at Georgia Tech. *College & Research Libraries News* 78 (1): 10–13.

Pinfield, Stephen, Andrew M. Cox, and Jen Smith. 2014. Research Data Management and Libraries: Relationships, Activities, Drivers and Influences. *PLOS ONE.* Available at http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0114734.

Pryor, Graham. 2012. Why Manage Research Data? In *Managing Research Data,* edited by Graham Pryor, 1–16. London: Facet Publishing.

Re3data.org team. 2015. DataCite to Manage and Develop re3data.org. Blog post, May 4, 2015. Available at http://www.re3data.org/2015/05/datacite-to-manage-and-develop-re3data-org/.

Roland, Betsy, and Charlotte P. Lee. 2013. Beyond Trust and Reliability: Reusing Data in Collaborative Cancer Epidemiology Research. *CSCW '13 Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, 435–444. New York: Association for Computing Machinery.

Shaffer, Christopher. 2013. The Role of the Library in the Research Enterprise. *Journal of EScience Librarianship* 2 (1): 8–15. doi:10.7191/jeslib.2013.1043.

Steinhart, Gail, Eric Chen, Florio Arguillas, Dianne Dietrich, and Stefan Kramer. 2012. Prepared to Plan? A Snapshot of Researcher Readiness to Address Data Management Planning Requirements. *Journal of eScience Librarianship* 1 (2): 63–78. Available at http://dx.doi.org/10.7191/jeslib.2012.1008.

Strasser, Carly. 2014. *Institutional Repositories: Part 2*. Datapub blog post, Feb. 20, 2014. California Digital Library. Available at http://datapub.cdlib.org/2014/02/20/institutional-repositories-part-2/.

Tenopir, Carol, Ben Birch, and Suzie Allard. 2012. *Academic Libraries and Research Data Services*. Chicago, IL: Association of College and Research Libraries. Available at http://www.ala.org/acrl/sites/ala.org.acrl/files/content/publications/whitepapers/Tenopir_Birch_Allard.pdf.

University of Alberta. 2013. *Research Records Stewardship Guidance Procedure. Appendix A: Research Records Management and Preservation Guidelines*. Available at https://policiesonline.ualberta.ca/PoliciesProcedures/Procedures/Research-Records-Stewardship-Guidance-Procedure-Appendix-A-Research-Records-Management-and-Preservation-Guidelines.pdf.

University of Alberta. 2015. *Research Policy*. Available at https://policiesonline.ualberta.ca/PoliciesProcedures/Policies/Research-Policy.pdf.

U.S. Department of Energy. 2014. *Statement on Digital Data Management*. Available at http://science.energy.gov/funding-opportunities/digital-data-management/.

U.S. Department of Health and Human Services. 2015. *Public Access Plans Cover Letter*. Available at http://www.hhs.gov/open/public-access/public-access-plans-cover-letter.html.

Whyte, Angus. 2014. *Final Results of DCC RDM 2014 Survey*. Available at http://www.dcc.ac.uk/blog/rdm-2014-survey.

Yakel, Elizabeth, Ixchel M. Faniel, Adam Kriesberg, and Ayoung Yoon. 2013. Trust in Digital Repositories. *International Journal of Digital Curation* 8 (1): 143–156. Available at http://dx.doi.org/10.2218/ijdc.v8i1.251.

Yoon, Ayoung. 2014. End Users' Trust in Data Repositories: Definition and Influences on Trust Development. *Archaeological Science* 14: 17–34. Available at http://link.springer.com/article/10.1007/s10502-013-9207-8?no-access=true.

Zilinski, Lisa, Christina Chan-Park, Robin Dasler, and Natsuko Nicholls. 2013. Carpe Data: Data Curation Services at Four Different Institutions. Presented at the Digital Library Federation Forum, Austin, Texas, Nov. 4, 2013. Available at http://docs.lib.purdue.edu/lib_fspres/39/.

Zimmerman, Ann. 2007. Not by Metadata Alone: The Use of Diverse Forms of Knowledge to Locate Data for Reuse. *International Journal of Digital Libraries* 7: 5–16. Available at http://link.springer.com/article/10.1007%2Fs00799-007-0015-8.