

2016

Going Beyond the Data Management Plan: Services and Partnerships

Abigail Goben

University of Illinois at Chicago, agoben@uic.edu

Lisa Zilinski

Carnegie Mellon University, ldz@andrew.cmu.edu

Kristin Briney

University of Wisconsin - Milwaukee, kabriney@gmail.com

Follow this and additional works at: https://dc.uwm.edu/lib_staffart

 Part of the [Library and Information Science Commons](#)

Recommended Citation

Goben, A., Zilinski, L., & Briney, K. (2016). Going Beyond the Data Management Plan: Services and Partnerships. In Federer, Lisa (Ed.), *The Medical Library Association Guide to Data Management for Librarians*. MLA.

This Article is brought to you for free and open access by UWM Digital Commons. It has been accepted for inclusion in UWM Libraries Other Staff Publications by an authorized administrator of UWM Digital Commons. For more information, please contact open-access@uwm.edu.

The Medical Library Association Guide to Data Management for Librarians

Chapter 10: Going Beyond the Data Management Plan: Services and Partnerships

Abigail Goben, Library of the Health Sciences, University of Illinois at Chicago

Lisa Zilinski, Carnegie Mellon University

Kristin Briney, University of Wisconsin-Milwaukee

As reviewed in the previous chapter, data management plans (DMPs) have provided an excellent gateway for librarians seeking ways to collaborate with researchers. However, the DMPs frequently capture only initial plans for data at the end of a grant application, often written just hours before the submission deadline. Researchers may also have data captured from a previous pilot project, a project already underway, or a project that is already completed. Further, once the grant is received or a plan has been completed, the researcher may realize she needs something entirely different halfway through the process; may run into intellectual property, patent, or privacy issues; or may encounter an unexpected publisher or university mandate regarding data.

These myriad experiences provide a variety of points for librarian-researcher collaboration that require the librarian to consider a broader array of research data management activities in order to best meet the needs across the data lifecycle. Considering a panoply of services also allows librarians to anticipate questions that may be referred to them. Before launching any data management services, it is important for librarians to understand the purpose of each service in order to select the services they are prepared to offer, identify library staff and funding needed to support these efforts, and have advance communication with relevant campus partners. To assist with this, we have outlined services and grouped those that require

progressive levels of time and effort, as well as identified the primary campus partners with whom the library will wish to communicate when launching these services. By identifying appropriate service levels, librarians can expand their collaborations in stages without feeling overwhelmed or unable to appropriately direct researchers. When investigating research data services for your library, it is best to start small, investigate research data needs locally, and develop strategies to build services over time.

Background

Library literature on research data management activities recommends librarians to consider opportunities beyond the data management plan, although researchers are quick to qualify that one size won't fit all libraries. Tenopir surveyed college and research librarians in 2012 and asked them to identify services that they were currently offering or were planning to offer.¹ At that time, only helping users find datasets and creating guides to finding data had greater than 50% current or planned uptake in the following two years. In that same year, Raboin described three universities with different approaches to research data management.² All identified the need for a broader variety of activities, but each identified unique opportunities at their institutions. Speaking from experience at his institution, Westra outlines embedded approaches to offering data management guidance, support, and collaboration.³ He discusses the notion that subject specialist liaisons need to incorporate data management responsibilities into their roles, while data librarians need to improve their skills in areas of science research, research

¹ Tenopir, et al., "Academic Libraries and Research Data Services."

² Raboin, et al., "Forging New Service Paths: Institutional Approaches to Providing Research Data Management Services."

³ Westra, "Developing Data Management Services for Researchers at the University of Oregon."

and data analysis tools, metadata, and curation and preservation workflows to support the subject specialists. More recently, Henderson and Knott⁴ describe the process of establishing data services that starts with collaborating on and reviewing DMPs and then branching out. This paper is particularly useful as it shows what can be done by an institution that is only able to employ one person to focus on data services.

Researchers examined more emerging roles with science librarians and noted that the publications on research data management have focused on individual, institution-based case studies.⁵ Antell identify the relatively new nature of research data management work as the reasoning behind seeing this lower level of evidence.⁶ However, it is important that librarians continue to move their research beyond case studies to more authoritative research levels.

With the National Science Foundation (NSF) DMP mandate now in its fifth year and responses to the Office of Science and Technology Policy mandates having been released throughout 2015, librarians are better positioned to understand the emerging funder requirements surrounding data and target services at their institutions.⁷ This expertise will allow for more

⁴ Henderson, and Knott, “Starting a Research Data Management Program Based in a University Library.”

⁵ Westra, “Developing Data Management Services for Researchers at the University of Oregon.”; Henderson, and Knott, “Starting a Research Data Management Program Based in a University Library.”; Antell et al., “Dealing with Data: Science Librarians’ Participation in Data Management at Association of Research Libraries Institutions.”

⁶ Antell et al., “Dealing with Data: Science Librarians’ Participation in Data Management at Association of Research Libraries Institutions.”

⁷ National Science Foundation, “Dissemination and Sharing of Research Results.”; Burwell et al., “Memorandum for the Heads of Executive Departments and Agencies.”; Holdren “Memorandum for the Heads of Executive Departments and Agencies: Increasing Access to the Results of Federally Funded Scientific Research.”; Whitmire et al., “A Table Summarizing the Federal Public Access Policies Resulting from the US Office of Science and Technology Policy Memorandum of February 2013.”

targeted services as well as improved assessment and research in order to determine the effectiveness and sustainability of library research data management services.

Internal Preparation

Before embarking on the development of research data services in the library, it is important for libraries to assess their capabilities and outline their planned services, as well as how they will be staffed and evaluated. Adding data management into the library's strategic plan and creating a specific strategic plan for the identified data services provides a framework for structuring these services. Education of internal staff, including librarians and front-line support, is integral to successful and sustainable research data services.

Strategic Planning

Beyond individual services, librarians can collaborate in mapping out an institutional vision for the development of data services based on local researcher needs. These efforts will allow the librarian to work with administration to identify partnerships, be aware of particular challenges with unique areas of research, and create action plans towards meeting community needs. An excellent example comes from Oregon State University Libraries & Press. This Strategic Agenda for Research Data Services outlines different phases of goals and activities, and lists partners and timelines.⁸ Any library wishing to strategically build its data offerings over time is likely to benefit from at least an informal strategic plan.

⁸ Sutton et al., "Oregon State University Libraries and Press Strategic Agenda for Research Data Services."

Data Management Education (Internal)

In order to promote data services and fully support researchers, it is essential that liaison librarians and those working with researchers become familiar with research data management vocabulary. Preliminary education may provide data lifecycle overview, guidance in locating datasets for reuse, a review of current funding mandates, and explanations of resources being supported by the library and their uses. Further education may include performing initial data management interviews, providing external data management education, and collaborating with research teams. One example of librarian-focused data management education is the Data Scientist Training for Librarians program created and led by Christopher Erdmann in Boston.⁹ Data librarians or an internal committee focusing on improving library data management awareness are best positioned to lead this internal training, although support and time must be offered by library management. Chapter 7 of this book provides further information on internal education and self-education opportunities.

Library Data Management Services

The concept of offering “data management services” may seem nebulous and too large a target for a solo librarian or a library committee for whom data management is only one of many responsibilities. The following services are defined with key examples from libraries currently offering these services and from the case study literature, with suggestions of who could provide the service and the target audience. These services should be considered in addition to the DMP assistance described in the previous chapter.

⁹ Guillette and Damon, “Data Scientist Training for Librarians: A Course and a Community.”

As libraries are developing their research data management service portfolio, it would be unsustainable to attempt to provide all services immediately without bringing in multiple librarians and staff. Additionally, the cultures of different institutions will likely not allow for all of these library services to develop, nor relationships to grow. It will be up to the individuals involved at each library to determine what additional workload can be absorbed and which services will best meet the needs of their institution.

While reviewing these services, librarians should consider what resources may be allocated to research data management activities and how data management aligns with the library mission. New data services must also align with other library strategic goals and current services so as not to exist in a vacuum. A suggested level of staff support is included for many service descriptions; however, this may vary depending on the library and external campus partners who can lend their support, whether in terms of time, money, or personnel. Though the majority of services primarily need library staff time, several of them require greater financial commitment.¹⁰

Finding Data

Just as librarians are experts with assisting patrons in locating appropriate literature, so too can they become facile in locating datasets for use in coursework and research. Due to the increasing amount of data available, particularly from open government initiatives and funder sharing requirements, patrons will need increasing assistance in locating appropriate data and navigating reuse requirements. Columbia University Libraries' Digital and Social Science Center (DSCC) leads their "Services to Users" list with helping to identify datasets both within and

¹⁰ Jones et al., *How to Develop RDM Services - a Guide for HEIs*

externally to the library collections.¹¹ While a data librarian may be of assistance here, liaison subject knowledge is likely to come to the fore in assisting patrons because of their expertise. One specific resource to explore is Re3data.org, which is a research data repository registry that can be explored and shared with students, faculty, and researchers.¹² As of this publication, the registry held over 1400 repositories, with nearly 350 in medicine alone.

Purchasing Data

Allowing patrons to request datasets easily aligns with collection development responsibilities. However, it may come with some new challenges. Many datasets are purchased by individuals or labs, and the dataset owner may not have procedures for institutional purchase or access. As most libraries do not purchase materials solely for an individual or small group's use, this can cause challenges in licensing. Datasets may also include material with unique storage and security requirements. Consultation with library or campus information technology (IT) departments to ensure discoverability without privacy violations may be necessary prior to purchase. An example of a data purchasing program is available from the University of Illinois at Urbana-Champaign, where purchase requests were solicited from faculty, academic professionals, and graduate students with the goals of meeting smaller needs (>\$5000) and improving the datasets available to the campus.¹³

Data Citation

¹¹ Columbia University Libraries, <http://library.columbia.edu/locations/dssc/data/service.html>.

¹² re3data.org, "Registry of Research Data Repositories."

¹³ "Data Purchase Program." <http://www.library.illinois.edu/sc/datagis/purchase/description2014.html>

While matters of intellectual property and copyright surrounding data as an object continue to evolve, a best practice for those reusing or referring to another's data is to provide a citation. Data and liaison librarians have the opportunity to assist faculty, researchers, students, and their peers in teaching how to cite a dataset. In their article on creating a metadata scheme for the DataCite consortium, Starr and Gastl describe suggested core pieces of a data citation that are used by DataCite and other organizations.¹⁴

A library may also be interested in setting up a service to create digital object identifiers (DOI) for researchers who are releasing data and who would like it to be cited and accessible. The advantage of using a DOI is that it can remain persistent over time, where websites are more likely to change or go unmaintained. One service that libraries can work with to offer DOIs and other unique identifiers is EZID from the California Digital Library. This service is likely to be managed by a specific data librarian, though education and referrals from liaison librarians will be essential.

Data Management Education (external)

While it is unlikely that a librarian can be assigned to work one-on-one with every research team at an institution, the library can offer research data management education. Graduate and doctoral level students and faculty supervising them are often interested in this education, as these skills may not be taught elsewhere in the curriculum. Full curricula such as the New England Collaborative Data Management Curriculum (NECDMC) and the DataONE Curriculum, developed with grant support from the NIH and NSF respectively, are available for

¹⁴ Starr and Gastle, "A Metadata Scheme for DataCite."

librarians to use as a starting point to develop their instructional offerings.¹⁵ Muilenburg provides insight into using the NECDMC curriculum and identifies areas where the framework may need to be adapted for individual institutions.¹⁶ External data management instruction may take the form of workshops, video series, webinars, and tutorials and may be offered by the data librarian alone or in collaboration with the liaison librarians. *Data Information Literacy: Librarians, Data, and the Education of a New Generation of Researchers*, edited by Jake Carlson and Lisa R. Johnston provides additional information about planning instruction.¹⁷

Data Management Consultations

Individual or group consultations for research data management are a natural opportunity for data librarians and liaison librarians to offer assistance. Such consultations may either be the product of or impetus to create additional services focusing on education, workflow creation, embedded librarians, or other data-related services. With a target of reaching faculty, students, and lab groups in particular, these consults are an opportunity to provide guidance, support, and resources on implementing data management, as well as providing a preliminary assessment of current data workflows for managing, curating, and publishing data. Carlson describes the Data Curation Profile (DCP) Toolkit as “the means for librarians to conduct data interviews with an individual research or small lab group.”¹⁸ The DCP can be used as a starting point in conducting

¹⁵ “New England Collaborative Data Management Curriculum” <http://library.umassmed.edu/necdmc/index>; “Education Modules.” <https://www.dataone.org/education-modules>

¹⁶ University of Washington, Muilenburg et al., “Lessons Learned From a Research Data Management Pilot Course at an Academic Library.”

¹⁷ Carlson and Johnston, *Data Information Literacy: Librarians, Data and the Education of a New Generation of Researchers*.

¹⁸ Carlson, “Demystifying the Data Interview: Developing a Foundation for Reference Librarians to Talk with Researchers about Their Data.”, 8.

data management consultations and can be customized to serve a number of purposes or disciplines.

Developing Data Workflows

Data librarians may also provide more targeted consultations by collaborating with lab staff or a lab manager in order to improve data workflows; example workflows include intra-lab data transfer or the evolution of naming conventions across the analysis process. Workflows may be established at the grant, project, or lab level with the goal of ensuring that data is captured in process rather than in backfill. As outlined by Tina Griffin, developing workflows will entail finding the difficult points within a lab group—whether that be tied to individuals, processes, or equipment, and determining needs from there to improve behavior.¹⁹ This service relies on the librarians providing an outside perspective on information organization and metadata, which are areas of librarian expertise. Workflow improvement may include organization changes, developing new lab policies or procedures, and data curation touch points for quality deposits and preservation. As this process is time-intensive, cost recovery for the librarians' efforts may be required.

Embedded Librarian

Even more in depth than consultations and workflow development is being embedded within the lab itself. This entails being an information manager for aspects of an entire research project. Medical librarians have participated in other forms of embedded librarianship, engaging

¹⁹ Griffin, “How Data Management Issues Manifest: Using Concept Mapping to Discover the Information Network in a Laboratory Environment.”

in curriculum, clinical rounds, and other activities as described by Florance and Davidoff; here, however, “embedded” refers specifically to being in a research laboratory.²⁰ Several examples of the embedded librarian in a grant come from the NIH Administrative Supplements for Information Services in NIH-funded Research Projects.²¹ Supplement Awards, paired with R01 and other large grants, have been given in 2012, 2014, and are anticipated for 2016.

The goal of this grant funding is to enhance the research process by adding an information specialist as an active and ongoing member of the research team. Previous awards have been used to develop online communication tools, create databases and warehouses, develop input tools, introduce best practices at the laboratory level, improve collaboration between multiple research groups, assist with GIS data, and perform literature searches and facilitate bibliographic management. While embedding has allowed data and liaison librarians to engage deeply with a research team, uncertainty remains about the sustainability of the efforts after the end of the funding cycle and the difficulties in scaling this level of service to other researchers who do not have similar funding.

Metadata Services

Another area of expertise that may be offered by the library is metadata services. Metadata librarians in particular have the opportunity to assist in contextual description and documentation in standard formats. As metadata standards vary widely across disciplines, there are many schemas, ontologies, and taxonomies that researchers may need to refer to or work

²⁰ Davidoff and Florance, “The Informationist: A New Health Profession?”

²¹ “NLM Administrative Supplements for Informationist Services in NIH-Funded Research Projects (Admin Supp).” <http://www.nlm.nih.gov/ep/AdminSupp.html>

within. Medical librarians may be most familiar with the NLM Metadata Schema, which incorporates Medical Subject Headings (MeSH) into standard Dublin Core elements and is used for materials published by the National Library of Medicine.²² A metadata service, as defined by the CalTech Metadata Services Group, may include helping researchers use metadata in order to find data, apply descriptive and technical metadata to identify what a dataset is about, identify rights metadata to assist with legal reuse, and create preservation and administrative metadata to document the data and its workflows over time.²³

Promoting Open Data

In examining whether shared data will be reused, Wallis affirmed the gift culture of scholarship, where data is shared between colleagues and friends based on personal relationship.²⁴ This has a great potential effect on access to data across institutions. As arbiters of information access, librarians have a unique opportunity to promote open data access. In much medical research this must, of course, be balanced with protecting research subject privacy, but librarians can assist in finding ways to share data subsets or appropriate metadata to facilitate awareness that a dataset exists that allows other researchers to discover and it and consider going through their institutional review board to petition for access. This is best done in collaboration with the scholarly communications librarian or other open access advocates.

Preparing Datasets for Deposit

²² “NLM Metadata Schema,” <http://www.nlm.nih.gov/tsd/cataloging/metafilenew.html>.

²³ “Metadata Services Group | Caltech Library,” <https://library.caltech.edu/metadata/>

²⁴ Wallis et al., “If We Share Data, Will Anyone Use Them? Data Sharing and Reuse in the Long Tail of Science and Technology.”

While ideally a researcher would have consulted with the library early in the grant process and created a data management plan, researchers may approach the library only after the completion of data gathering and as they are preparing datasets for deposit. The degree to which the library is willing and able to provide this service is likely to vary based on factors such as time needed to prepare the dataset, homogeneity of the data, cleanliness, dataset complexity, and repository requirements. The data librarian should have some capacity to provide consultations for data deposit preparation in a subject or other non-institutional repository. The Inter-University Consortium for Political and Social Research provides a data preparation guide with overarching guidelines.²⁵

Collaborative Data Management Services

Some services and activities are more resource intensive than an individual campus unit can frequently support. In addition to librarian expertise, they may require technical systems, physical space, or further commitments from administration and faculty to ensure success. Despite requiring more resources, these services can meet a specific campus need and should be considered when strategizing the best data offerings.

Data Policy

Policies surrounding data, including ownership, access, and sharing are evolving quickly. Researchers may encounter policies from funders, institutions, and journals, and these policies may vary across disciplines. Until recently, the NIH required a data management plan only for

²⁵ “Guide to Social Science Data Preparation and Archiving: Introduction,” <http://www.icpsr.umich.edu/icpsrweb/content/deposit/guide/index.html>.

research with greater than \$500,000 requested in any grant year and expressed particular concern for maintaining protection of human subject data.²⁶ The February 2015 response to the OSTP mandate described a planned requirement that all NIH-funded researchers have data management plans for peer review. However, at the time of publication, this new policy has not been fully implemented.²⁷ “A table summarizing the Federal public access policies resulting from the US Office of Science and Technology Policy memorandum of February 2013”, compiled by a collaboration of data librarians around the country, is updated regularly and attempts to collect and consolidate the guidelines and policies from federal agencies related to their plans for facilitating public access to results from the research each agency funds.²⁸ Along with becoming familiar with external policies, librarians can collaborate to create or update institutional policies surrounding data.

Collaborating on policy development offers a different avenue than many data services for the library: the opportunity to work with administration and the research office. While education efforts to assist with compliance will continue to focus on students and faculty, policy creation provides an opportunity for a data librarian to work alongside campus research support personnel and administration. Briney reviews the current institutional policy environment and identifies further opportunities for librarians to engage with data policy.²⁹

²⁶ “NIH Data Sharing Information,” http://grants.nih.gov/grants/policy/data_sharing/.

²⁷ “National Institutes of Health Plan for Increasing Access to Scientific Publications and Digital Scientific Data from NIH Funded Scientific Research,” <http://grants.nih.gov/grants/NIH-Public-Access-Plan.pdf>.

²⁸ Whitmire et al., “A Table Summarizing the Federal Public Access Policies Resulting from the US Office of Science and Technology Policy Memorandum of February 2013.”

²⁹ Briney et al., “Do You Have an Institutional Data Policy? A Review of the Current Landscape of Library Data Services and Institutional Data Policies.”

Data Repositories

As part of promoting open data and preparing datasets for deposit, librarians may connect patrons with subject data repositories such as GenBank or general data repositories like figshare. However, a library may also wish to explore providing a data repository at their institution to meet the needs of their researchers, faculty, and students. The purpose of a repository is to provide preservation and access to datasets, allowing researchers to meet funding agency requirements or to gain further citations and awareness of the research. One example of institutional data repositories is the Purdue University Research Repository (PURR), which provides all faculty, students, and staff with a collaborative working space and data sharing platform, as well as DOIs for datasets.³⁰ PURR also allows inter-institutional collaboration, as long as one member of the team is at Purdue. Because of the hardware, software, and storage space required for a dedicated data repository, creating a repository will require collaboration with the library IT department and possibly the campus IT department, as well as potentially significant financial investment. Further details on creating an institutional data repository can be found in an article by Michael Witt and a chapter by D. Scott Brandt, which describes the Purdue experience.³¹

For those interested in focusing on small data, or the long tail as described by Borgman, adding datasets to what is accepted into existing institutional repositories—which typically focus primarily on articles—may prove a more reasonable option.³² Some libraries may also wish to explore collecting data corresponding to the electronic theses and dissertations in their

³⁰ “PURR – Home,” <https://purr.purdue.edu/>

³¹ Witt, “Defining and Deploying an Institutional Data Repository Service at Purdue (PURR)”;
Brandt, “Purdue University Research Repository: Collaborations in Data Management.”

³² Borgman, “Big Data and the Long Tail: Use and Reuse of Little Data.”

institutional repositories. Any such data collection should involve collaboration between the repository manager and the data librarian. Library policies should be established early to clarify data appropriate for deposit in either a dedicated data repository or an institutional repository, as well as the formats accepted and how data must be prepared for ingest.

Data Analysis Support

Data analysis support should not be taken to mean that the library is taking on the responsibilities of the researcher to perform research data analysis. Rather, the focus of this service is to teach the tools for performing analysis. These tools may fall into a number of categories: statistical analysis such as SPSS or SAS, qualitative software such as NVivo, programming languages such as Python, documentation software such as Colectica, or other analysis tools such as OpenRefine. A data librarian will need awareness of these tools in order to understand where processed data may be coming from.

Library offerings may include teaching introductory webinars, coordinating access to webinars on these tools offered by professional organizations or the vendors, or going further to establish a digital scholarship center in the physical library. More expensive software may be loaded on dedicated computers in the library for greater access among students and researchers who are not able to purchase individual subscriptions for a small or unfunded project. One example of such a center can be found at Notre Dame, which provides assistance with GIS, data use, text mining, digitization, data management, and referral.³³ If such services are offered, it is important that more than one member of the library staff be able to perform basic troubleshooting on the software.

³³ “Home | Center for Digital Scholarship,” <http://library.nd.edu/cds/>

Data Visualization

As with data analysis, the purpose of library data services in visualization comes more in terms of instruction, opportunities for experimentation and collaboration than in providing end services themselves. Data visualization is defined here as transforming data into visual representations including a map, bar chart, timeline or an artistic rendering and these services have a target audience of faculty, researchers, and students. The library could specialize in teaching a couple of specific tools, such as GeoCommons or Tableau, which allow for easy interactive maps that could be used for public health projects. A visualization space, with dedicated workstations and loaded software such as Duke University offers, also allows for collaboration and experimentation outside of dedicated program-specific computer labs.³⁴ The library also has an opportunity to serve as a location where visualizations may be shared. If there is a Biomedical Visualization program at the institution, class samples may be displayed with permission in the library on a dedicated computer or in a gallery space. Chapter 12 provides further discussion of library-based data visualization services.

Secure Data Room

Institutions and university security offices are faced with the challenge of allowing researchers to use sensitive secondary data while maintaining appropriate precautions. This may include educational, personal, health, financial, or other kinds of data from outside sources. The library has the opportunity to become a location for secure data access on campus by setting aside a locked space containing a secure, non-networked computer. The library can also provide

³⁴ “The Brandaleone Lab for Data and Visualization Services,” <http://library.duke.edu/data/about/lab>

boilerplate information on secure access for grant applications. Such a space allows security checks to be centralized and prevents duplication of secure rooms in multiple departments.

Creating a secure space in the library will require collaboration with the IT department and may require extra staff to monitor the room. Though this solution is not necessarily scalable for the entire campus, there may be opportunities to partner with high performance computing centers. For example, the Pittsburgh Supercomputing Center (PSC) is a joint effort of Carnegie Mellon University and the University of Pittsburgh.³⁵ PSC provides researchers with access to their Data Supercell to enable storage, transfer, and access services for all sizes of datasets.

Identifying Campus Partners

In order to best develop these services, it will be important for libraries to develop relationships across the institution.³⁶ Partners can provide services that the library cannot presently tackle, as well as augmenting services the library does provide with their resources and expertise. Campus partners also have the opportunity to refer researchers, faculty, and students to the library for data management services at different points in their research. By collaborating, the library prevents service duplication across different units on campus. For the purposes of this chapter, campus partners will be considered external to the library, as it is assumed that all departments in the library have been engaged as partners and stakeholders in relevant services.

Campus IT

³⁵ “PSC Data Supercell,” <https://www.psc.edu/index.php/major-initiatives/data-supercell>

³⁶ Erway and OCLC Research, *Starting the Conversation University-Wide Research Data Management Policy*.

Researchers frequently work with campus IT when they are trying to find storage space for their research data. Campus IT often provides data storage and infrastructure during the life of the project and will also have an interest in data security. Their concerns are not necessarily focused on what happens after a research grant is done or in describing the data, and they may have limited interest in sharing the data. At the very least, data librarians should communicate with local IT to learn local best practices and storage offerings to recommend during consultations with researchers.

Office of Research and/or Sponsored Programs

The Office of Research and/or Sponsored Programs is likely to work with all researchers at some point, whether it be in an oversight or grant reviewing capacity. Researchers who are applying for external funding and those who have received funding will work with this office on fulfilling the terms of application and compliance for their grant. As such, this office has a strong interest in ensuring that researchers have data management plans where required and that these meet minimum expectations of grant review boards. They further have an interest in ensuring compliance with data requirements and that any required data sharing is done to prevent delay in receiving renewal or new grants. This office can refer researchers who are struggling with research data management as it relates to the grant process and would be integral partners in determining needs of the campus for the strategic plan of the library.

In addition, this office is likely able to contact all graduate students and to encourage them to attend educational programs. Some offices even run programming themselves, such as a Responsible Conduct of Research series, in which data librarians can take part as core instructors. The Association of Research Libraries' SPEC Kit on 336: Responsible Conduct of

Research Training provides examples of what these workshops may entail and opportunities for librarians to offer instruction.³⁷ The research office is also likely to know about campus funding opportunities, as well as more significant research that may not be receiving external funding, particularly in subspecialties like History of Medicine.

Institutional Review Board (IRB) / Animal Care Committee (ACC)

Anyone working with human subjects or animals should have contact with the IRB or ACC to ensure compliance with federal regulations for safety and ethics. In this capacity, the review boards are particularly well placed to refer questions about workflow and data sharing plans back to the library for consultation, workflow review, and similar services. Where possible, a librarian sitting on these committees could serve as a contact point.

Faculty Senate

As a body, the faculty senate has a particular interest in the intellectual property produced by the university researchers and will likely have opinions about research data. As libraries are advocating for policy development or changes, the senate can provide insight and direction and serve as a conduit to the faculty as a whole. Because their focus is different from other administrations, they are likely to identify other services and needs that may not be readily apparent from the infrastructure side.

Office of Technology Management/Intellectual Property/Business Development

³⁷ Leonard and Bennett, “Responsible Conduct of Research Training, SPEC Kit 336.”

While faculty are often primarily focused on the generation of new knowledge and moving science, medicine, and scholarship forward, saleable intellectual property may also be an outcome of research. This office has the responsibility of taking ideas created by researchers at the institution to sale for private business. They will be engaged in filing for patents, developing businesses, or marketing patents to companies. Because patent filing has specific data retention requirements and sharing restrictions, this office may influence data storage and sharing policies and practices that the library is implementing.

Selecting the Right Services

Selecting the right data services for a library depends primarily on three factors: staffing, funding, and the library's goals. On the staffing side, the library must decide if data support will be the purview of a single librarian or if it will take a group approach. While a skilled librarian may be able to deliver very specialized services, it is often more sustainable to have several individuals providing at least low-level data support. A group approach also enables data support to be a normalized part of the total library offerings. Additionally, partnering with other campus offices could allow for resource sharing of both personnel and budget.

Each of the services described in this chapter also requires a different level of staff support. This is exemplified by the three consultation services – data management consultations, developing workflows for a project, and embedded librarians – which require increasing levels of staff time and therefore may be more or less feasible, depending on the library. Staff time devoted for a project should be balanced against staff skills. For example, metadata services and data repositories both require specialized skills or coordination with librarians who work in these areas. Finally, librarians' interests should be taken into consideration, especially when services

are provided by a group. Leveraging individual librarians' existing interests is a great way to more quickly establish services in a new area.

The other consideration for staffing services is identifying personnel outside the library who can act as partners in support. As has been discussed, a wide range of potential campus partners may be able to assist libraries. While not critical for every potential service, working with campus stakeholders can make services stronger and more far-reaching. Wherever they exist, partnerships with other campus groups should be leveraged and new relationships cultivated to prove the best data support.

Staff time will be a major resource needed for new services, but additional funding is often necessary, to a greater or lesser degree depending on the service. For example, data purchase programs and providing a secure data room will require more money than advocating for data policy or providing training. The library may also need to dedicate resources to technical solutions, such as when building a data repository or providing DOIs. As several of the described services are not scalable for the whole campus, the library should consider the scale of the project as well as the tradeoff between staff time and money when deciding on offerings. At least one of the two resources will be required for any service.

The ultimate goal is to find the right service portfolio to meet the library's mission. For example, a library that is a strong advocate for openness may wish to support open data and a data repository, whereas a library that is more focused on information literacy might adopt a data management training curriculum. Reviewing the library's strategic plan and developing a coordinating strategic plan for data services will be of benefit.

Conclusion

Libraries have a considerable number of options for providing data support beyond consulting on data management plans. This chapter describes an array of possible services that different libraries are already offering. While it is rarely feasible to provide every service listed, libraries should take advantage of their staff skills, their goals, and existing partnerships on campus to provide a level of data support that is right for their campus. When in doubt, start small, examine local researcher data needs, and create a plan for how to grow data support over time.

Pearls

- A wide variety of research data management services are needed beyond data management plans for grants.
- Services can be targeted at different audiences and rolled out over time or as staffing is available
- Funding for resources or staff will be necessary for successful service launch
- Collaborating with other campus partners will help with alerting researchers to the services available and prevent duplication of effort.

Bibliography

- Antell, K., J. B. Foote, J. Turner, and B. Shults. 2014. "Dealing with Data: Science Librarians' Participation in Data Management at Association of Research Libraries Institutions." *College & Research Libraries* 75 (4): 557–74. doi:10.5860/crl.75.4.557.
- Borgman, Christine. 2013. "Big Data and the Long Tail: Use and Reuse of Little Data." *Oxford eResearch Centre Seminar*, March. <http://works.bepress.com/borgman/269>.
- Brandt, D. Scott. 2014. "Purdue University Research Repository: Collaborations in Data Management." In *Research Data Management: Practical Strategies for Information Professionals*, edited by Joyce M. Ray, 325–45. West Lafayette, Indiana: Purdue University Press.
- Briney, Kristin, Abigail Gobin, and Lisa Zilinski. 2015. "Do You Have an Institutional Data Policy? A Review of the Current Landscape of Library Data Services and Institutional Data Policies." *Journal of Librarianship and Scholarly Communication*.
- Burwell, Sylvia, VanRoekel, Steven, Park, Todd, and Mancini, Dominic. 2013. "Memorandum for the Heads of Executive Departments and Agencies." <http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf>.
- Carlson, Jake. 2012. "Demystifying the Data Interview: Developing a Foundation for Reference Librarians to Talk with Researchers about Their Data." *Reference Services Review* 40 (1): 7–23.
- Carlson, Jake, and Lisa R. Johnston, eds. 2015. *Data Information Literacy: Librarians, Data and the Education of a New Generation of Researchers*. West Lafayette, Indiana: Purdue University Press.

“Data Purchase Program.” 2015. Accessed June 1.

<http://www.library.illinois.edu/sc/datagis/purchase/description2014.html>.

“DSSC Data Service | Columbia University Libraries.” 2015. Accessed May 31.

<http://library.columbia.edu/locations/dssc/data/service.html>.

“Education Modules | DataONE.” 2015. Accessed June 12. <https://www.dataone.org/education-modules>.

Erway, Ricky., and OCLC Research. 2013. *Starting the Conversation University-Wide Research Data Management Policy*. Dublin, Ohio: OCLC Research.

<http://www.oclc.org/content/dam/research/publications/library/2013/2013-08.pdf>.

“EZID: EZID Home.” 2015. Accessed June 16. <http://ezid.cdlib.org/>.

Davidoff, Frank, and Valerie Florance. 2000. “The Informationist: A New Health Profession?”

Annals of Internal Medicine 132 (12): 996–98.

Griffin, Tina. 2015. “How Data Management Issues Manifest: Using Concept Mapping to Discover the Information Network in a Laboratory Environment.” presented at the Research Data Access & Preservation Summit, Minneapolis, MN, April 23.

“Guide to Social Science Data Preparation and Archiving: Introduction.” 2015. Accessed June 1.

<http://www.icpsr.umich.edu/icpsrweb/content/deposit/guide/index.html>.

Guillette, Jeremy, and James Damon. 2014. “Data Scientist Training for Librarians: A Course and a Community | Library Connect.” Newsletter. *Library Connect*. November 13.

<http://libraryconnect.elsevier.com/articles/2014-11/data-scientist-training-librarians-course-and-community>.

- Henderson, Margaret E., and Teresa L. Knott. 2015. "Starting a Research Data Management Program Based in a University Library." *Medical Reference Services Quarterly* 34 (1): 47–59. doi:10.1080/02763869.2015.986783.
- Holdren, John P. 2013. "Memorandum for the Heads of Executive Departments and Agencies: Increasing Access to the Results of Federally Funded Scientific Research." *Office of Science and Technology Policy, Executive Office of the [US] President, The White House*, February.
http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf.
- "Home | Center for Digital Scholarship." 2015. Accessed June 1. <http://library.nd.edu/cds/>.
- Jones, Sarah, Graham Pryor, and Angus Whyte. 2013. *How to Develop RDM Services - a Guide for HEIs*. Edinburgh: Digital Curation Centre. <http://www.dcc.ac.uk/resources/how-guides/how-develop-rdm-services>.
- Leonard, Michelle, and Denise Beaubien Bennett. 2013. "Responsible Conduct of Research Training, SPEC Kit 336," ARL Digital Publications, September.
<http://publications.arl.org/Responsible-Conduct-Research-Training-SPEC-Kit-336/>.
- "Metadata Services Group | Caltech Library." 2015. Accessed May 31.
<https://library.caltech.edu/metadata/>.
- "National Institutes of Health Plan for Increasing Access to Scientific Publications and Digital Scientific Data from NIH Funded Scientific Research." 2015, February.
<http://grants.nih.gov/grants/NIH-Public-Access-Plan.pdf>.
- National Science Foundation. 2010. "Dissemination and Sharing of Research Results." *US NSF - About*. November 30. <http://www.nsf.gov/bfa/dias/policy/dmp.jsp>.

“New England Collaborative Data Management Curriculum | Lamar Soutter Library - University of Massachusetts Medical School.” 2015. Accessed May 31.

<http://library.umassmed.edu/necdmc/index>.

“NIH Data Sharing Information - Main Page.” 2014. Accessed October 21.

http://grants.nih.gov/grants/policy/data_sharing/.

“NLM Administrative Supplements for Informationist Services in NIH-Funded Research Projects (Admin Supp).” 2015. Accessed May 31.

<http://www.nlm.nih.gov/ep/AdminSupp.html>.

“NLM Metadata Schema.” 2015. Accessed May 31.

<http://www.nlm.nih.gov/tsd/cataloging/metafilenew.html>.

“PSC Data Supercell”. 2015. Accessed May 31. <https://www.psc.edu/index.php/major-initiatives/data-supercell>.

“PURR - Home.” 2015. Accessed May 31. <https://purr.purdue.edu/>.

Raboin, Regina, Rebecca Reznik-Zellen, and Dorothea Salo. 2012. “Forging New Service Paths: Institutional Approaches to Providing Research Data Management Services.” *Journal of eScience Librarianship*. doi:10.7191/jeslib.2012.1021.

“re3data.org | Registry of Research Data Repositories.” 2015. Accessed June 2.

<http://www.re3data.org/>.

Starr, Joan, and Angela Gastle. 2011. “isCitedBy: A Metadata Scheme for DataCite.” *D-Lib Magazine* 17 (1/2). doi:10.1045/january2011-starr.

Sutton, Shan, David Barber, and Amanda L. Whitmire. 2013. “Oregon State University Libraries and Press Strategic Agenda for Research Data Services.”

<http://ir.library.oregonstate.edu/xmlui/handle/1957/38794>.

- Tenopir, Carol, Birch, Ben, and Suzie Allard. 2012. "Academic Libraries and Research Data Services." ACRL.
http://www.ala.org/acrl/sites/ala.org.acrl/files/content/publications/whitepapers/Tenopir_Birch_Allard.pdf.
- "The Brandeone Lab for Data and Visualization Services | Duke University Libraries." 2015. Accessed June 1. <http://library.duke.edu/data/about/lab>.
- University of Washington, Jennifer Muilenburg, Mahria Lebow, and Joanne Rich. 2014. "Lessons Learned From a Research Data Management Pilot Course at an Academic Library." *Journal of eScience Librarianship* 3 (1): 67–73. doi:10.7191/jeslib.2014.1058.
- Wallis, Jillian C., Elizabeth Rolando, and Christine L. Borgman. 2013. "If We Share Data, Will Anyone Use Them? Data Sharing and Reuse in the Long Tail of Science and Technology." Edited by Luís A. Nunes Amaral. *PLoS ONE* 8 (7): e67332. doi:10.1371/journal.pone.0067332.
- Westra, Brian. 2014. "Developing Data Management Services for Researchers at the University of Oregon." In *Research Data Management: Practical Strategies for Information Professionals*, edited by Joyce M. Ray, 375–91. West Lafayette, Indiana: Purdue University Press.
- Whitmire, Amanda, Kristin Briney, Amy Nummerger, Margaret Henderson, Thea Atwood, Margaret Janz, Wendy Kozlowski, Sherry Lake, Micah Vandegrift, and Lisa Zilinski. 2015. "A Table Summarizing the Federal Public Access Policies Resulting from the US Office of Science and Technology Policy Memorandum of February 2013." http://figshare.com/articles/A_table_summarizing_the_Federal_public_access_policies_r

esulting_from_the_US_Office_of_Science_and_Technology_Policy_memorandum_of_F
ebruary_2013/1372041.

Witt, Michael. 2014. "Defining and Deploying an Institutional Data Repository Service at
Purdue (PURR)." http://docs.lib.purdue.edu/lib_fspres/44/.