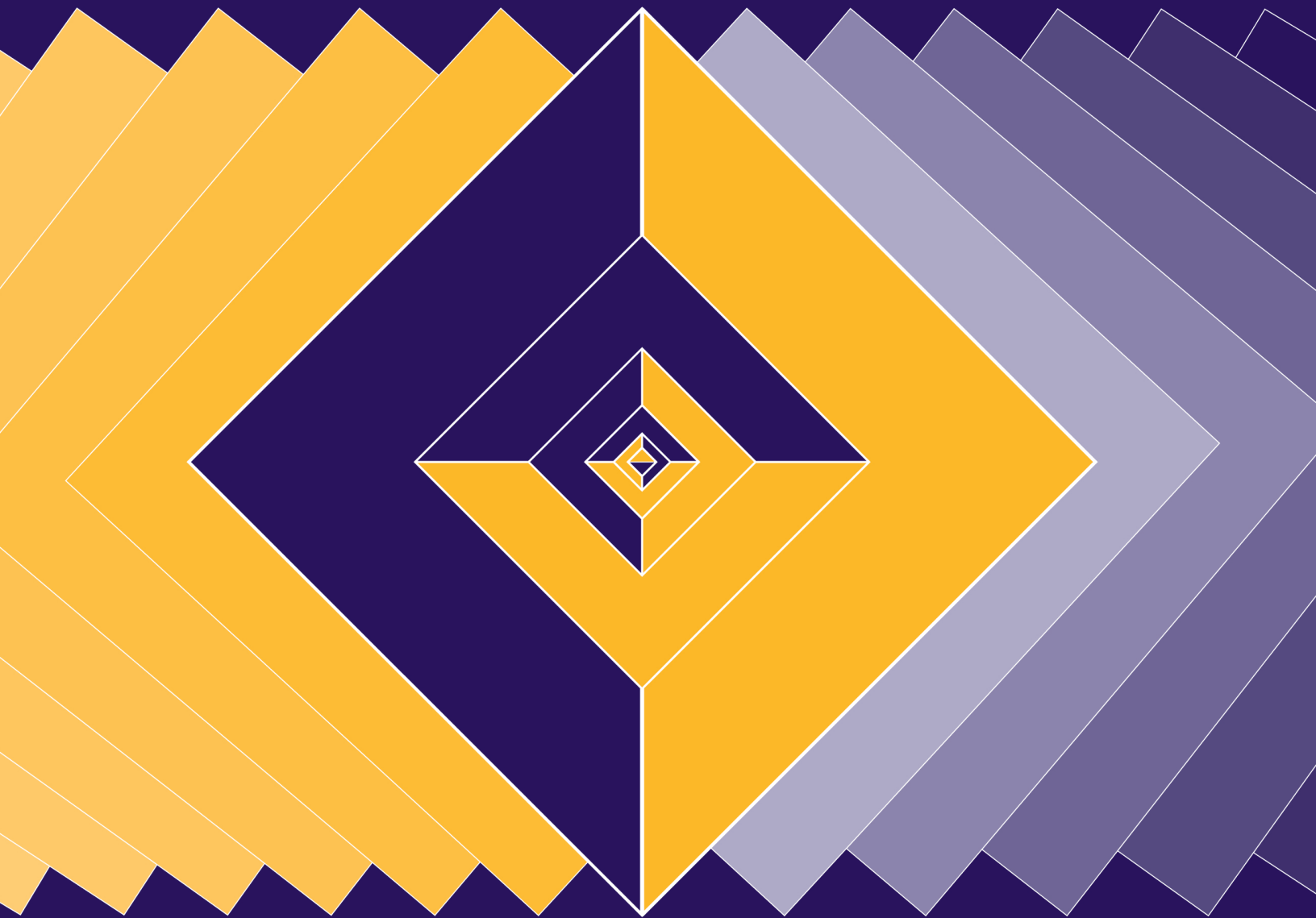




Kit 334

Research Data Management Services

July 2013



ASSOCIATION OF RESEARCH LIBRARIES

# SPEC KITS

Supporting Effective Library Management for Over 40 Years

Committed to assisting research and academic libraries in the continuous improvement of management systems, ARL has worked since 1970 to gather and disseminate the best practices for library needs. As part of its commitment, ARL maintains an active publications program best known for its SPEC Kits. Through the Collaborative Research/Writing Program, librarians work with ARL staff to design SPEC surveys and write publications. Originally established as an information source for ARL member libraries, the SPEC Kit series has grown to serve the needs of the library community worldwide.

## What are SPEC Kits?

Published six times per year, SPEC Kits contain the most valuable, up-to-date information on the latest issues of concern to libraries and librarians today. They are the result of a systematic survey of ARL member libraries on a particular topic related to current practice in the field. Each SPEC Kit contains an executive summary of the survey results; survey questions with tallies and selected comments; the best representative documents from survey participants, such as policies, procedures, handbooks, guidelines, Web sites, records, brochures, and statements; and a selected reading list—both print and online sources—containing the most current literature available on the topic for further study.

## Subscribe to SPEC Kits

Subscribers tell us that the information contained in SPEC Kits is valuable to a variety of users, both inside and outside the library. SPEC Kit purchasers use the documentation found in SPEC Kits as a point of departure for research and problem solving because they lend immediate authority to proposals and set standards for designing programs or writing procedure statements. SPEC Kits also function as an important reference tool for library administrators, staff, students, and professionals in allied disciplines who may not have access to this kind of information.

SPEC Kits are available in print and online. The executive summary for each kit after December 1993 can be accessed online free of charge. For more information visit: <http://www.arl.org/publications-resources>.



# SPEC Kit 334

---

Research Data Management Services

---

July 2013

**David Fearon, Jr., Betsy Gunia, and Barbara E. Pralle**

---

Johns Hopkins University

**Sherry Lake and Andrew L. Sallans**

---

University of Virginia



ASSOCIATION OF RESEARCH LIBRARIES

---

Series Editor: Lee Anne George

SPEC Kits are published by the

Association of Research Libraries

21 Dupont Circle, NW, Suite 800

Washington, DC 20036-1118

P (202) 296-2296 F (202) 872-0884

<http://www.arl.org/publications-resources>  
[pubs@arl.org](mailto:pubs@arl.org)

ISSN 0160 3582

ISBN 1-59407-901-3 / 978-1-59407-901-6 print

ISBN 1-59407-902-1 / 978-1-59407-902-3 online

Copyright © 2013

This compilation is copyrighted by the Association of Research Libraries. ARL grants blanket permission to reproduce and distribute copies of this work for nonprofit, educational, or library purposes, provided that copies are distributed at or below cost and that ARL, the source, and copyright notice are included on each copy. This permission is in addition to rights of reproduction granted under Sections 107, 108, and other provisions of the US Copyright Act.



The paper used in this publication meets the requirements of ANSI/NISO Z39.48-1992 (R1997) Permanence of Paper for Publications and Documents in Libraries and Archives.



## SURVEY RESULTS

<b>Executive Summary</b> .....	11
<b>Survey Questions and Responses</b> .....	23
<b>Responding Institutions</b> .....	105

## REPRESENTATIVE DOCUMENTS

### RDM Web Pages

University of California, San Diego <b>Research Cyberinfrastructure</b> .....	110
Emory University <b>Research Data Management</b> .....	111
University of Wisconsin-Madison <b>Research Data Services</b> .....	112

### IRs and Data Archives

Columbia University <b>Academic Commons</b> .....	114
University of Florida <b>IR@UF: The Institutional Repository at the University of Florida</b> .....	115
Harvard University <b>Harvard Dataverse Network</b> .....	116
University of Nebraska-Lincoln <b>UNL Data Repository</b> .....	117
University of New Mexico <b>Lobo Vault – UNM Research Data</b> .....	118
University of Oregon <b>Scholars' Bank</b> .....	119
Pennsylvania State University <b>ScholarSphere</b> .....	120

Purdue University	
Purdue University Research Repository (PURR).....	121
University of Virginia	
LIBRA: Online Archive of University of Virginia Scholarship.....	122

### **Data Management Plan Tools**

Indiana University	
Indiana University Guidance on NSF Data Management Plans .....	124
Johns Hopkins University	
Questionnaire to Help with the Creation of a Data Management Plan.....	128
University of Minnesota	
Data Management Plan .....	131
University of North Carolina	
Research Data Toolkit .....	134

### **Job Descriptions**

Johns Hopkins University	
Data Management Consultant .....	136
University of Maryland	
Research Data Librarian .....	138
University of Ottawa	
Social Sciences Research Liaison Librarian.....	141
Purdue University	
Digital Library Software Developer .....	144
Digital Data Repository Specialist .....	148
University of Tennessee	
Data Curation Librarian .....	152
University of Virginia	
Data Consultant.....	154

### **Organization Charts**

Brown University	
Center for Digital Scholarship .....	156
Cornell University	
Research Data Management Service Group Organizational Structure .....	157
Duke University	
Information Technology Services.....	158
Massachusetts Institute of Technology	
Specialized Content and Services .....	159
University of Rochester	
River Campus Libraries Organizational Chart.....	160

## **Research Data Policies**

University of Iowa	
Data Management: Research Records .....	162
Johns Hopkins University	
Policy on Access and Retention of Research Data and Materials .....	164
McGill University	
Regulation on the Conduct of Research excerpt .....	167
University of Tennessee	
Research Data Policy .....	169

## **Data Retention Policies**

Duke University	
Policy on Research Records: Sharing, Retention and Ownership .....	174
Emory University	
Policy 7.9: Guidelines for Responsible Conduct of Scholarship and Research .....	175
Section 7.9.02C: Access to and Retention of Scientific Research Protocols and Data .....	176
Johns Hopkins University	
Appendix V from Policy on Access and Retention of Research Data and Materials .....	177
Northwestern University	
Research Data: Ownership, Retention and Access .....	178
University of Virginia	
Policy: Laboratory Notebook and Recordkeeping .....	182
University of Wisconsin – Madison	
Policy on Data Stewardship, Access, and Retention .....	183

## **Data Needs Assessment**

Cornell University	
Research Data Management Group Survey of NSF Principal Investigators at Cornell University .....	188
Emory University	
Research Data Management. Faculty Survey Results .....	197
Massachusetts Institute of Technology	
A Data Planning Checklist .....	198
Purdue University	
Data Curation Profiles Toolkit .....	199
University of Virginia	
Data Interview Protocol .....	200

## **RDM Staff Resources**

University of Florida	
Data Management/Curation Task Force .....	206

Purdue University	
Supporting Information for Data Services.....	207
Yale University	
Data & eScience Group.....	208

**Outreach Materials**

University of Colorado at Boulder	
Research Data Services Outreach Card.....	210
University of Florida	
Speaker Series .....	211
Indiana University Bloomington	
Workshops .....	212

**SELECTED RESOURCES**

<b>Research Data Management Services (RDMS) and the Library.....</b>	<b>215</b>
<b>Supporting Data Management Planning .....</b>	<b>216</b>
<b>Data Archiving.....</b>	<b>216</b>
<b>Staffing and Staff Training .....</b>	<b>217</b>
<b>RDMS Policies .....</b>	<b>218</b>
<b>Scanning and Assessment of the Landscape for RDMS .....</b>	<b>218</b>
<b>Key Papers in the Development of RDMS.....</b>	<b>219</b>
<b>ARL/DLF/Duraspace eScience Institute .....</b>	<b>219</b>



## **SURVEY RESULTS**



---

## EXECUTIVE SUMMARY

---

### Introduction

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

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

Our survey also addresses contextual requirements for planning, developing, staffing, and managing new research data management services. Finding necessary expertise and funding for new positions is challenging, but creative new models of service provision are emerging. Collaboration across institutional units is one route toward unearthing expertise

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

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

## Background

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

One of the areas that we recognized as important in building upon the 2010 report was further refinement of terminology, which may have varying interpretations or ambiguity among respondents and more broadly in this field of service. Examples of such terms include “eScience,” “cyberinfrastructure,” “data archiving,” and distinguishing research data management from data services more broadly, the latter sometimes including institutional records management. Perhaps the most ambiguous term still emanating through these conversations is “digital services.” As the survey results show, in application this can range from digitization to research support to intellectual property to repository management. While we attempted to specify particular meanings and control response variables for the purpose of a more systematic analysis, this process and some of the responses we received highlight the fact that there is still great variation in maturity and definition of

services throughout the ARL community. Based on anecdotes, we believe that there is also even broader variation and understanding of the meaning of these terms and services beyond the ARL community, such as with the consumers of the services. Among other goals, we hope that this type of study may shed some light on the variations of meaning, and provide some opportunity for further maturation and convergence of terminology.

## Broad Data Support Services

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

## Research Data Management Services

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



### *Origin of Research Data Management Services*

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

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

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

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

who are planning to attend a future institute for the first time also already provide RDM services.

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

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

### *Data Management Plans*

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

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

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

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

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

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

number of participants in group sessions; the next highest number was 96 sessions.

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

### *RDM Services Other than DMP Preparation*

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

### **Data Archiving Services**

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

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

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

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

### *Data Archive Characteristics*

From prior literature and informal observations, we expected current archiving solutions among libraries to fall into two categories: repositories designed specifically for research data, and those repositories built for other purposes that can contain datasets. For the latter, we expected institutional repositories (IRs) used primarily for publications to be most common, as well as datasets archived in digital repositories that are used for institutional special collections such as historical photos. We expected data-specific archives built specifically for retaining and accessing research data to be less prevalent because of the higher resource

requirements for setting up a system, the lack of ready-made data repository platforms (in contrast to the more mature enterprise of institutional repository software), and the relative lack of demand for libraries to initiate such platforms at this time.

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

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

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

### *Funding Data Archives*

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

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

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

### *Data Archive Infrastructure*

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

To assess the use of repositories for data, the survey asked for estimates of the number of researchers currently depositing datasets in the archives (Q23), the typical sources of archived data (Q25), and total deposit size (Q26). Twenty-eight of those with IRs reported that zero to 1000+ researchers have deposited data (a median of 10 and an average of about 91 researchers). Four of those with data archives reported that between two and 100 researchers have deposited data. Twenty-two of the respondents (66%) reported that data deposits are in the gigabyte range; all but three are under 100GB. Eleven others reported total deposits between 1 and 75 terabytes. Follow-up with respondents might yield more precise numbers

and distinguish among archives with single large and many small deposits. Clearly, however, these are early days for both data-specific archives and IRs with data, and possibly also for researcher’s awareness and adoption of these archiving options.

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

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

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

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



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

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

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

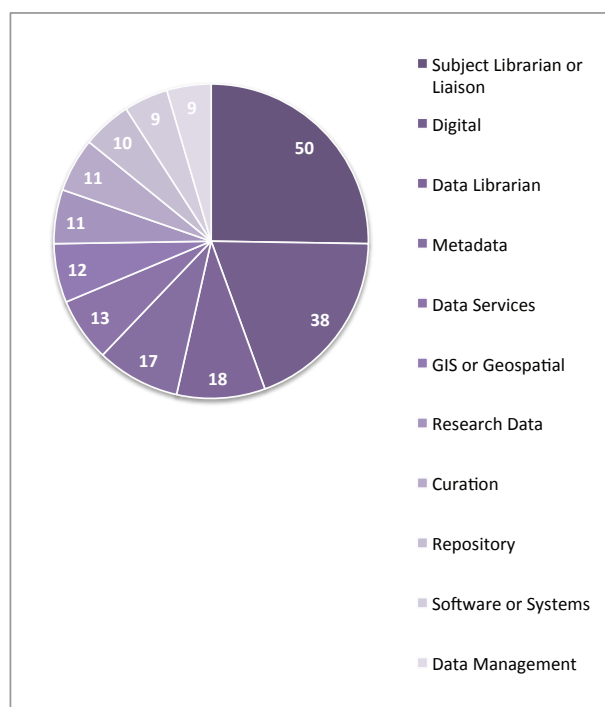
It will be increasingly relevant to follow up on libraries' forays into data archiving, since 30 institutions indicated they plan to offer data archiving within two years (Q56), but nearly a quarter of the respondents included providing data archiving in their top three challenges for RDM services (Q54),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

## RDM Service Staffing

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

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

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



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

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

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

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

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

less than half their time on RDM activities; nearly a quarter (47) spends less than 10%.

## Education and Skills

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

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

## Training Needs

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

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

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

### **Funding RDM Services**

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

### **Partnerships**

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

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

### **Conclusion**

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

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

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

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

Finally, the third biggest challenge reported is faculty (non)engagement due to a lack of awareness of services that the library provides, low perceived value of services, and resistance to data sharing. Respondents stated that the most effective marketing techniques were through workshops and presentation to researchers, referrals from research project (grants) administration, and direct emails to researchers (Q50).

The low perceived value of services is a more difficult issue to overcome. Some researchers do not view the library as a resource for data management, and as one respondent commented, “preservation and data sharing are a hard sell when the researcher only need[s] to write a plan.” Also, some respondents noted that funders are not taking mandates seriously, giving the researcher a low incentive to care about data management and sharing. As noted, the push for open data access for publicly funded research and compliance for data sharing policies, both in North America and around the world, may change the environment significantly. Our survey results suggest that many ARL libraries have at least a start toward growing services to meet an increasing demand.

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

### **Limitations of Survey**

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

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



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

### Acknowledgements

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

### Endnotes

- 1 For a sample see the references in the Selected Resources section of this SPEC Kit.
- 2 <http://www.arl.org/storage/documents/publications/escience-report-2010.pdf>
- 3 See heading "Key Papers in the Development of RDMS" in the Selected Resources section of this SPEC Kit.
- 4 <http://www.arl.org/focus-areas/e-research/e-science-institute>
- 5 See [White House Office of Science and Technology Policy \(OSTP\)](#) Feb. 22, 2013 memorandum on open access to funded research data and publications.
- 6 Dataverse is in relatively widespread use as a repository for specific disciplines and research centers, but only one library reported being directly involved with a Dataverse implementation.
- 7 Attendance at an E-Science Institute workshop noted earlier in responses to Question 3 was another influential training resources for many respondents.



## SURVEY QUESTIONS AND RESPONSES

---

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

---

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

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

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

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

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

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

## BACKGROUND QUESTIONS

---

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

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

Comments N=21

### **Yes, there is a policy now**

Current policy only covers IP & technology transfer.

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

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

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

University Records Retention Schedule.

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

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

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

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

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

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

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

**No**

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

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

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

Nothing beyond regular records retention policy.

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

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

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

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

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

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

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

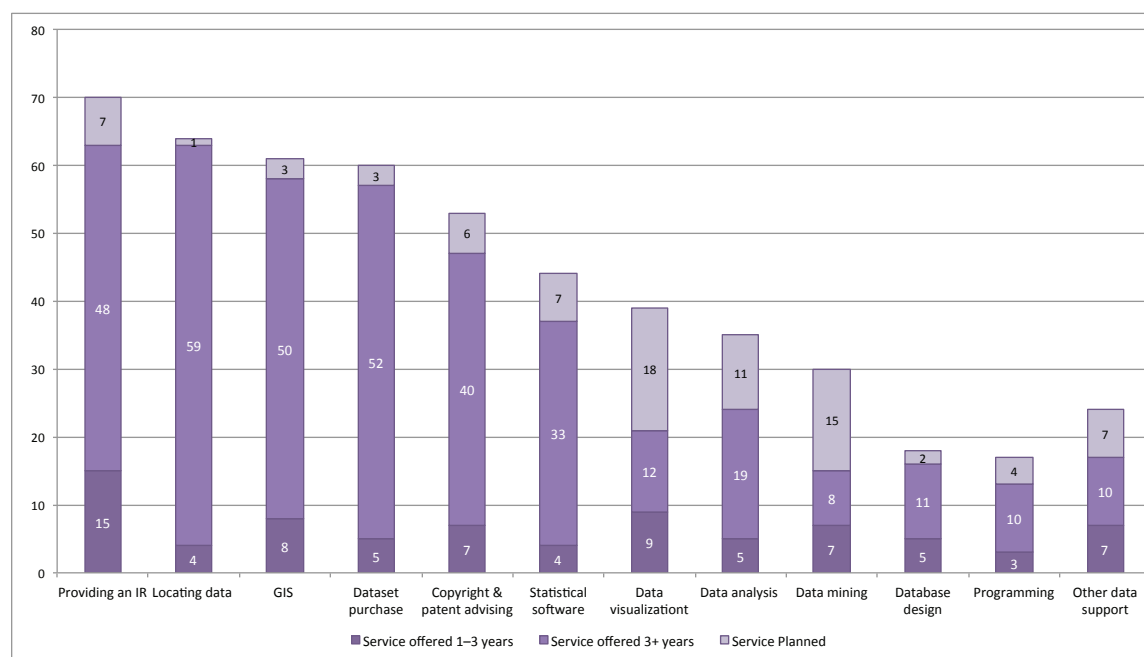
2011–12 ARL/DLF E-Science Institute	43	81%
2012–13 ARL/DLF E-Science Institute	7	13%
Future ARL/DLF/DuraSpace E-Science Institute	7	13%

## BROAD DATA SUPPORT SERVICES

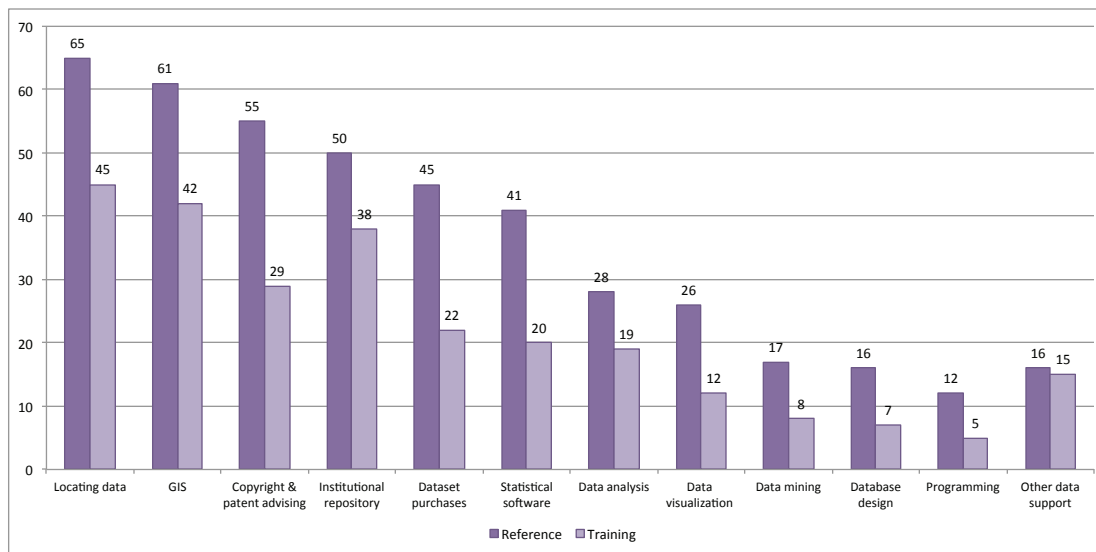
Most of this survey will focus on research data management (RDM) services as described in the introduction. Our first question, however, asks about the broader range of services provided at your institution that support research and data access.

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

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



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



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

Data preparation support: geocoding, georeferencing, distance/area/density calculations, converting formats, merging/joining, extracting/subsetting, and digitizing.

Data access: maintain a data extractor and data fileserver.

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

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

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

Metadata consulting, Persistent Identifiers/data citation.

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

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

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

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

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

Web scraping/harvesting, metadata management.

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

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

## **RDM Services**

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

Advice for data management plans.

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

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

Data management plan training and support, DMP tool

Data Management Planning (DMPTool)

Data management plans

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



## RESEARCH DATA MANAGEMENT SERVICES

---

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

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

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

### Comments N=11

#### Yes

Minimal. Information only through a LibGuide.

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

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

We are starting up these services now.

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

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

We provide basic level support of research data management planning.

#### No, but we plan to

Planned LibGuide on website.

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

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

### No, and we have no plans to

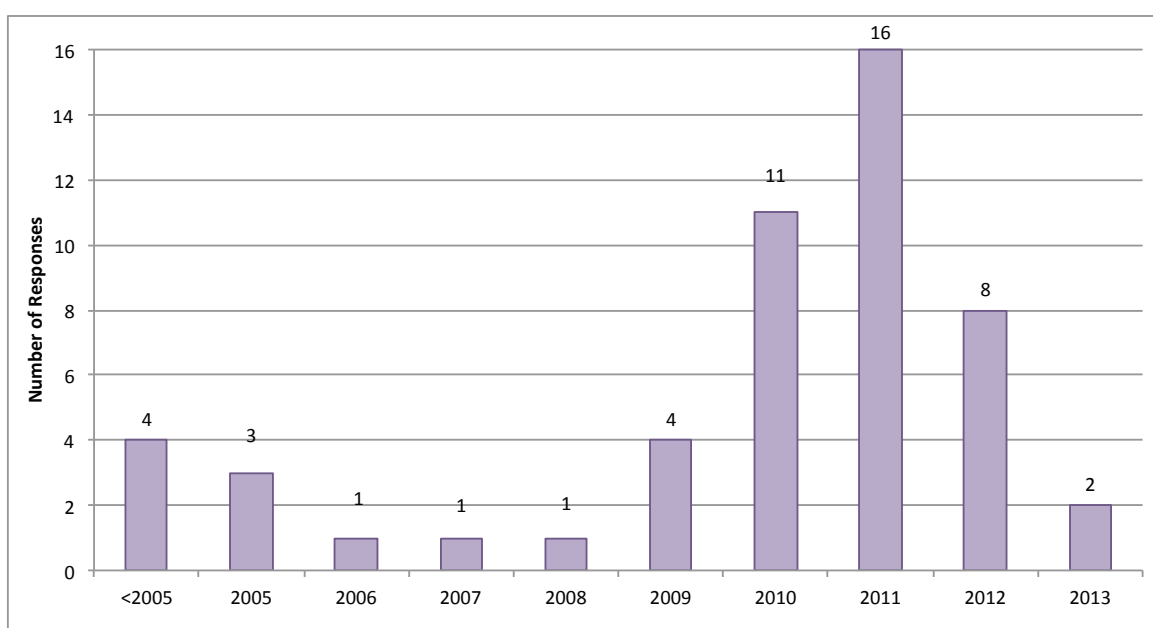
No, we wish we could, but we have no formal plan to do so in the short term because of lack of resources.

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

## ORIGIN OF RDM SERVICES

---

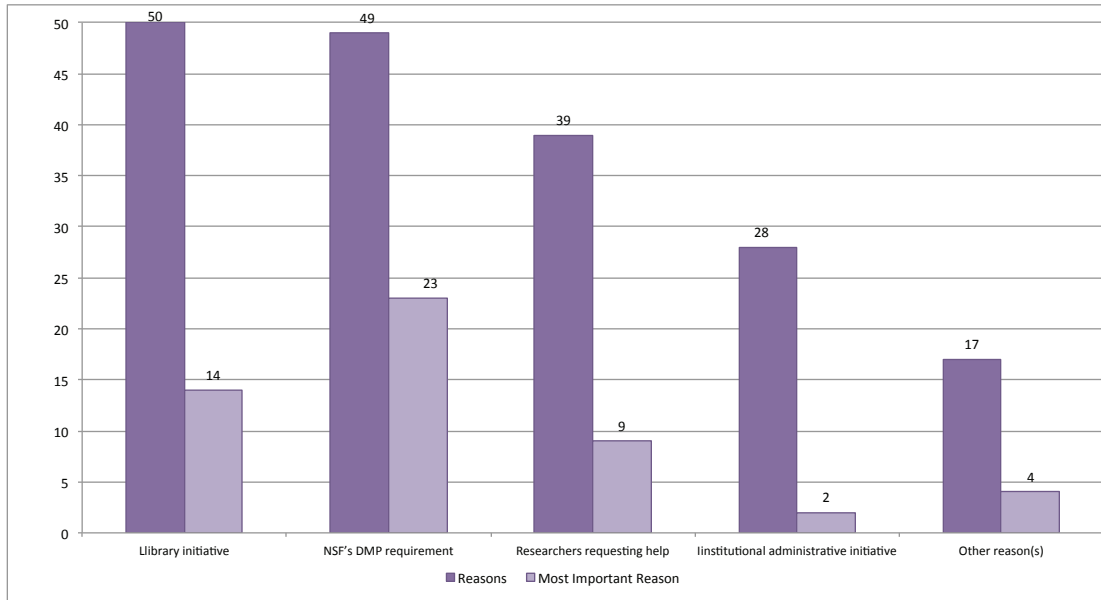
5. Please enter the year when your library initiated RDM services. (This can include simply adding references to your webpage if this is the extent of your RDM services). N=51



Range: 1966 to 2013

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

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



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

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

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

Developing trend among peer institutions.

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

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

Institutional history for archiving research data.

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

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

Library strategic direction.

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

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

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

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

Recognition of the growth of e-Science.

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

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

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

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

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

Addressing a need on campus.

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

Filling gap identified by faculty.

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

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

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

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

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

Our library director also made this effort a priority here.

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

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

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

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

We deploy a diverse, multidisciplinary support team.

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

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

## **ONLINE DATA MANAGEMENT PLAN (DMP) SERVICES**

---

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

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

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

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

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

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

A data planning checklist.

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

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

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

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

Information about digital repository services.

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

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

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

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

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

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

### **Offered by the Library AND Elsewhere**

Central IT, IRB

Central IT, Texas Advanced Computing Center (TACC)

Health Sciences, Office of Sponsored Programs

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

Office of Research

Office of Sponsored Research Projects

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

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

Research center

Research computing (central IT)

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

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

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

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

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

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

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

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

### **Only Offered Elsewhere**

Central IT, Research Administration, Office of Sponsored Programs

Research Computing/Office of Research and Technology Management

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

Yes	41	76%
No	13	24%

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

Yes	29	54%
No	25	46%

## DMP TRAINING & CONSULTATION SERVICES

---

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

Yes	33	61%
No	21	39%

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

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

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

### Offered by the Library AND Elsewhere

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



Center for Computational Science and possibly the Office of Research  
Central IT, TACC  
Clinical and Translational Science Institute, Data Coordinating Center  
College of Arts & Sciences: Office of Research Funding and Support. Research and Health Sciences IT.  
Colleges, research & sponsored projects.  
Health Sciences  
High performance computing and research center  
Institute for Policy and Social Research, Center for Research Methods & Data Analysis  
IRB, Sponsored programs, Research Administration, but this is ad hoc and not a centralized or organized effort.  
Office of Research  
Office of Research, Office of Sponsored Programs  
Office of Research and Engagement  
Research administration  
Research center, Research administration/sponsored projects  
Research Computing (central IT) Research Computing links and directs these to the Libraries, but provides support specific to High Performance Computing as well.  
Some research institutes offer limited DMP services to researchers affiliated with the Institute.  
Sponsored Program Services/Pre-Awards (mostly done in conjunction with the Libraries).  
Sponsored programs proposal development office  
The Odum Institute also offers such consultations. Other campus groups that provide DM solutions (for which they would help a researcher write a DMP) include a division of campus IT called Research Computing; a semi-independent computing institute based on campus called RENCI; and the Data Intensive Cyber-Environment (DICE) group based in the library school.

### **Only Offered Elsewhere**

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

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

## DMP CONSULTATION SERVICES

---

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

2010	2011	2012	2013
10	23	12	2

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

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

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

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

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

Faculty survey

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

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

No faculty has contacted the library and asked for help.

Online training and videos

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

Researchers contact me directly via email most of the time.

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

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

Tutorials, training, collaborations with Office of Contracts and grants

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

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

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

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

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

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

Whatever the researcher needs.

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

Workshops

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

Yes      28   58%  
 No        20   42%

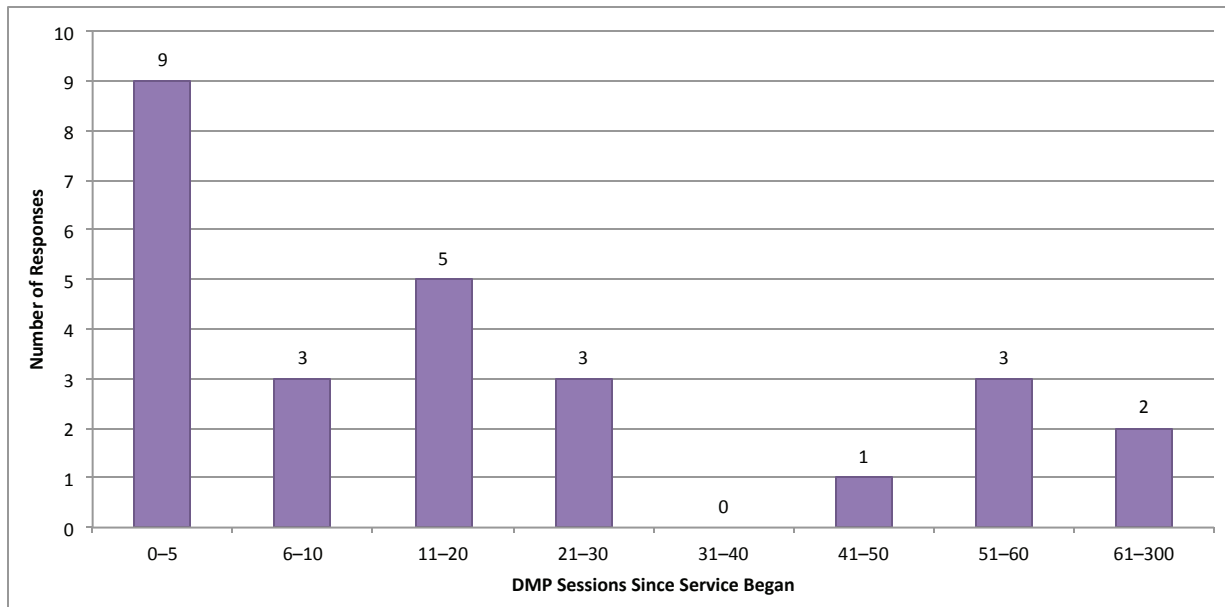
If yes, please enter the number of data management plans consulted on per month, on average. If possible, please also enter the number of DMP sessions since this service began. N=26

**Average number of plans per month**

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

**DMP sessions since service began**

Minimum	Maximum	Mean	Median	Std Dev
0	300	31.62	11.50	59.72



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

Yes 37 80%  
 No 9 20%

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

Transactional Only (in Minutes) N=7

Minimum	Maximum	Mean	Median	Std Dev
30	60	45.0	45.0	15.00

Iterative Only (in Hours) N=12

Minimum	Maximum	Mean	Median	Std Dev
1	10	4.6	3.0	3.26

Both Transactional & Iterative N=5

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

Transactional, Iterative, & Project N=6

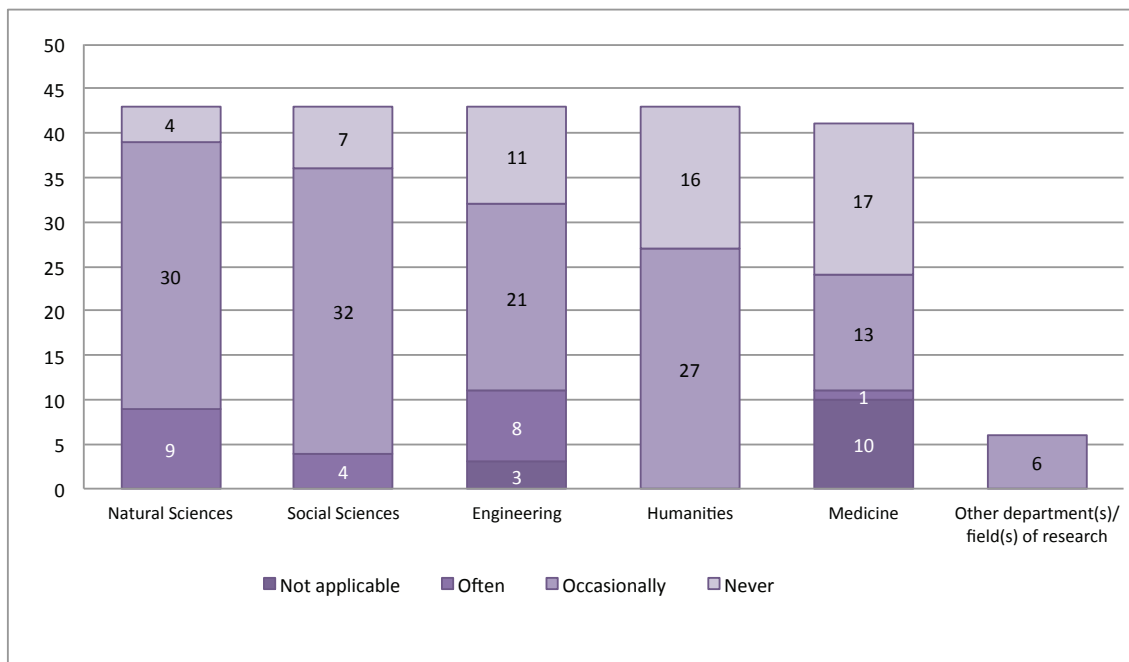
Duration	Minimum	Maximum	Mean	Median	Std Dev
Minutes	30	90	60.0	60.0	18.97
Hours	1	3	1.8	2.0	0.75
Days	.8	10	3.6	2.5	3.48

Additional Comment

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

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

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



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

Agriculture

Business

Libraries

Other areas include the arts and agriculture.

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

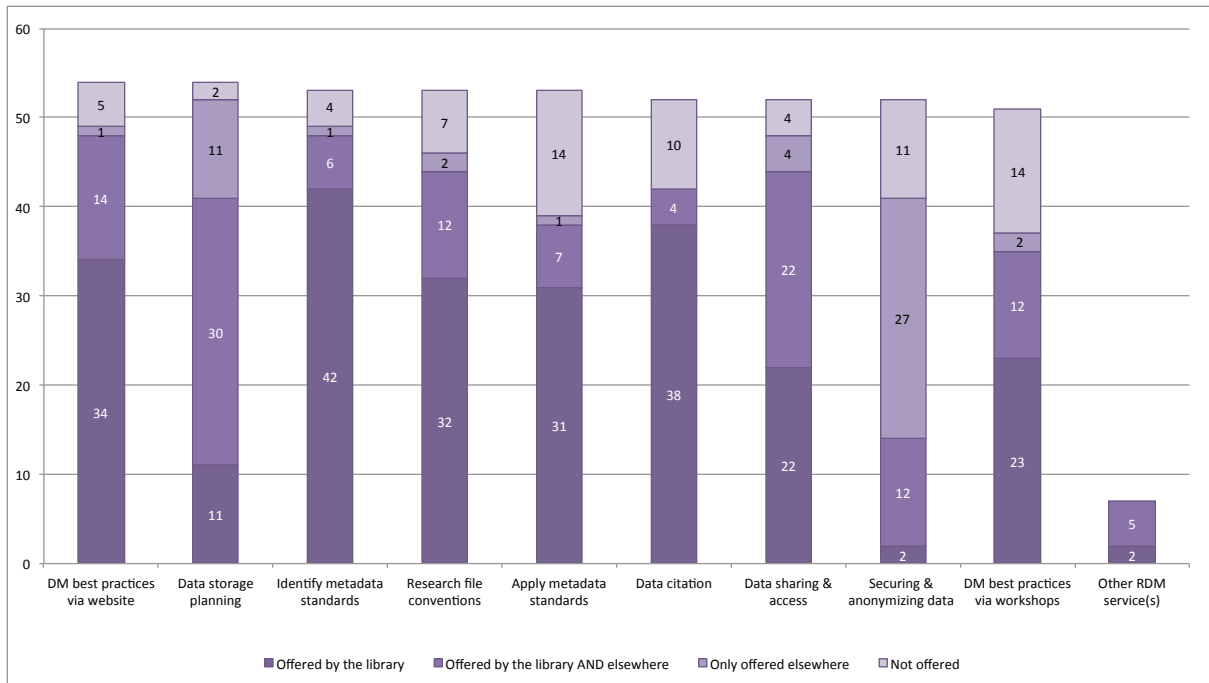
Physical sciences and other various unspecified.

## **RDM SERVICES OTHER THAN DMP SUPPORT**

---

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

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



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

### Offered by the library

Conversion of analog image data to digital data.

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

### Offered by the library AND elsewhere

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

Data publication, data rights management.

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

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

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

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

Campus cyber-infrastructure organization

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

Central IT (2 responses)

Central IT, Research and Health Sciences IT

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

Central IT, college-level

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

Central IT, human subjects protection, IRB

Central IT, IRB, Office of Research

Central IT, research centers, IRB

Central IT, TACC

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

Data storage and backup planning: Central IT

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

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

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

Health Sciences, Office of Sponsored Programs, University Archives

High performance computing and research center

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



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

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

Information Technology Services

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

IRB

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

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

Office of Research Services

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

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

Other academic departments and organizational units

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

Research Computing (central IT), IRB

Research computing and office of research

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

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

Sponsored Programs

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

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

## DATA ARCHIVING SERVICES

---

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

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

N=54

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

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

### Offered by the library AND elsewhere

Central IT (2 responses)

Central IT and TACC

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

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

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

Consortially through Scholars Portal

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

Many

Research center

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

Varies

### Only offered elsewhere

Center for Computational Science

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

College of Literature, Sciences and Arts

Colleges and departments

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

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

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

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

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

Dataverse

Deposit into Chronopolis, a campus-managed service.

We are building a solution for data only.

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

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

## DATA ARCHIVE CHARACTERISTICS

---

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

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

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

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

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

BePress Digital Commons (2 responses)

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

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

Chronopolis

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

DAITSS

Dataverse

Drupal

DSpace (11 responses)

DSpace (with a planned conversion to Hydra/Fedora)

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

DSpace, Dark Archive

DSpace, Dataverse

Fedora Commons (3 responses)

Fedora, Hydra (2 responses)

Fedora, iRODS to manage storage and preservation functions

Fedora, Data Conservancy software

Fedora, Merritt

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

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

HUBZero

Meta Archive Cooperative

Self developed software

SobekCM

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

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

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

#### IR with Data Comment

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

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

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

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

Please briefly describe the other method.

##### IRs with Data

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

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

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

Paid for consortially.

##### Data Archives

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

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

To be determined.

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

## IRs with Data

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

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

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

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

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

## Data Archives

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

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

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

Under review.

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

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

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

## IRs with Data

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

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

Directly from faculty, Archives and Special Collections.

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

Raw data, active storage.

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

### Data Archives

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

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

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

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

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

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

### IRs with Data

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

1 TB

2 GB per project.

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

50 MB and then must ask permission.

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



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

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

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

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

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

We are still working on establishing appropriate limits.

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

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

### Data Archives

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

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

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

### IRs with Data Answered No

Have ability to override limits.

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

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

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

Comments N=7

### IRs with Data

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

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

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

Self-deposits are always mediated.

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

### Data Archives

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

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

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

Yes            38   97%  
No             1   3%

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

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

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

### IRs with Data

EZID

Fedora PID URI

Locally created UUID [Universally Unique Identifier]

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

Locally generated Persistent Identifier (PID)

Persistent URL (2 responses)

Resource Object ID

This will happen in our planned managed digital repository.

UUID

## Data Archives

Self-generated

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

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

Please briefly describe the other preservation capabilities. N=11

### IRs with Data

Academic Preservation Trust

Audit trail

Checksum (fixity)

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

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

Format migration is manual on upload and only if required.

Normalization of formats, characterization of the files

Periodic virus scanning

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

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

### Data Archives

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

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

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

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

#### IRs with Data

Archive service for faculty.

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

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

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

#### Data Archives

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

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

## STAFFING: ORGANIZATIONAL STRUCTURE

---

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

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

## LIBRARY COMMITTEE/GROUP

---

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

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

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

## **LIBRARY AND INSTITUTION COMMITTEE/GROUP**

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

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

## SINGLE LIBRARY POSITION

---

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

Center for Digital Research and Scholarship

Digital Library Services

Information Technology

Physical Sciences and Engineering Library

Scholarly Communication

Scholarly Communications and Digital Curation

Scholarly Publishing and Data Management

Systems Office

## SINGLE LIBRARY DEPARTMENT

---

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

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

## OTHER ORGANIZATIONAL STRUCTURE

---

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

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

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

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

## **STAFFING: POSITION DETAILS**

---

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

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

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

### **Single Library Position N=8**

Data Curation Librarian  
Digital Repository Coordinator  
Digital Services Librarian  
Director of Research Systems Development  
Research Data Librarian  
Research Data Manager  
Research Services Librarian  
Science Data Management Librarian

### **Two Positions N=4**

Digital Repository Librarian  
Metadata Librarian, Head of Digital Access Services  
  
Research Data & Metadata Librarian  
Research Data Specialist (Research Computing)  
  
Science Data Management Librarian  
CLIR Data Curation Fellow  
  
Web Support Librarian, Desktop Network Services  
Science & Technology Librarian



**Three Positions N=2**

Digital Curation Librarian

Data Services Librarian

Metadata Librarian

Faculty, Digital Repositories Librarian

Faculty, Digital Technologies Development Librarian

Faculty, Data Services and Informatics Librarian

**Four Positions N=9**

(e-)Science Librarian II

Science Librarian II

Digital Publishing Librarian I

Library Director

e-science librarian

Head, Digital Collections and Scholarly Communication Services

Biosciences Librarian

Associate Director for Academic and Research Technologies, Research Computing Services

Engineering Librarian

Health Sciences Librarian

Assistant Head of Reference

Knowledge Management Specialist

eScholarship, ePublishing & Digitization Coordinator

Data Management Librarian (to be hired in 2013)

Liaison librarian

Library assistant

Head of the RSMAS Library

E&O Librarian for Sciences

Metadata Librarian

RSMAS Librarian

Head, ScholarSphere User Services

Metadata Librarian

Data Management Specialist

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

Post Doctorate for Biology

Science and Health Librarians

Metadata Specialist

Associate University Librarian

Research Data Librarian

Assistant Dean for Digital Humanities Research

Engineering/Research Data Librarian

Data Curation Assistant (Student Position)

Research Liaison Librarians

Data Librarian

GIS Librarian

Scholarly Communications Librarian

**Five Positions N=7**

CLIR Data Curation Intern

CLIR Data Curation Intern

Head of Science, Engineering, and Data Team

Head of Institutional Repository

Physics and Astronomy Librarian

DACS Committee Chair

Data Curation Librarian

Data Librarian for Economics and Business

Research Librarian for Engineering

Fine Arts Librarian

Head, Data and GIS Services

GIS Librarian

Data Visualization Coordinator

Systems Librarian

Data Services Intern

Librarian

Librarian

Associate Director, Scientific Computing and Visualization Group

Institutional Repository Librarian

Assistant Director of Library Computing & Systems

Lorry I. Lokey Science Data Services Librarian

Social Science Data & Government Documents Librarian

Head, Digital Scholarship Center

Coordinator, Institutional Repository

Map/GIS Librarian

Research Data Curation Program Director

RDC Technical Manager

Faculty Liaison Project Manager

Faculty Liaison

Metadata analyst/consultant

Scientific Data Curator

Social Sciences Data Librarian

Digital Humanities Librarian

Digital Repository Manager

Head, Digital Production Services

**Six Positions N=23**

Associate Dean

Data Curation Librarian (search in progress)

Graduate research assistant / intern  
Development Programmer  
Digital Production and Publishing Manager  
Head, Digital Initiatives

Associate Dean for Library Technologies  
Senior Associate Dean for Digital Scholarship and Technical Services  
Director of Digital Scholarship  
Digital Library Architect  
Head of Metadata and Cataloging  
Information Technology Enterprise Architect

Associate Head, Collection Management, and Director of Research Data Services  
Head, Digital Library Initiatives and Digital Projects  
Data Services Librarian  
Director, Copyright and Digital Scholarship  
Digital Collections and Preservation Librarian  
Associate Head, Research and Graduate Studies

Data Library Coordinator  
GIS Librarian  
Digital Repository Services Librarian  
Digital Initiatives Coordinator  
Digital Preservation Officer  
Metadata Librarian(s)

Data Management Specialist  
e-Science Librarian/Postdoctoral Fellow  
Coordinator of Geospatial Data Services  
Coordinator of Numeric Data Services  
Copyright & Scholarly Communications Librarian  
Multiple other members

Digital Assets Librarian

Map Librarian

Data Librarian

Business Librarian

Science Librarian

Social Sciences Librarian

Digital Humanities Librarian

Reference & Liaison Librarian in the Health Science Center

Agricultural Sciences Librarian

Political Science Librarian

GIS Librarian

University Archivist and Head of Records Management

Digital Scholarship Librarian

GIS Specialist

Digital Librarian

Research Services Librarian(s)

Team Leader: Digital Learning & Scholarship

Team Leader: Acquisitions & Metadata Services

Digital Strategies Coordinator

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

Head, Systems and Web Management

Digital Repository Resident

Science Librarian for the Center for Hierarchical Manufacturing

University Archivist

Head of Maps & Data Initiatives

GIS Specialist

Technology Specialist

Statistical Computing Specialist

eSciences Librarian

Social Sciences Librarian

Head of reference

Director of collection development

Life sciences librarian

Physics/astronomy/chemistry/mathematics librarian

Engineering librarian

Geology librarian

Head of Specialized Content & Services

Social Science Data Services Librarian

Bioinformatics and Biosciences Librarian

Electrical Engineering and Computer Science Librarian

Head of Curation & Preservation Services

Digital Curation and Preservation Services Library Fellow

Head of Strategic Data Initiatives

Data Consultant

Health Sciences Data Consultant

Senior Data Consultant

Data Consultant

Intern

Interdisciplinary Research Librarian

Data Services Specialist (2)

Digital Data Repository Specialist

Software Developer (2)

Metadata Specialist

Digital Archivist

Librarian

Librarian

Academic Technologist

Academic Technologist

Information Architect

Information Architect

Librarian for Numeric Data Services and Data Management

(Health Sciences Library) Bioinformatics and Translational Science Librarian

(Health Sciences Library) Cataloging and Metadata Coordinator

Head, Kenan Science Library

GIS Librarian

(IR) Information Infrastructure Architect

Manager

Data Management Consultant

Data Management Consultant

Systems Administrator

Software Developer

Technical Consultant

Numeric Data Services Librarian

Scholarly Communication Librarian

GIS Librarian

Metadata/Cataloging Librarian

Research & Assessment Analyst

Director, Program Management Center

Research Data Librarian

Digital Projects Specialist

Metadata and Digital Resources Librarian

Acting Head, Collection Development

Research Librarian

Director, Information Technology

Research Data Management Service Coordinator, Science Data and Metadata Librarian

Consultant, Research Data and Life sciences Librarian

Consultant, Life sciences Librarian

Consultant, Physics and Astronomy Librarian

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

Consultant, Research Associate in Astronomy

Science and Social Science Data Librarian

Science and Social Science Data Librarian

Statistical Consultant

StatLab Manager

Director of Academic IT Solutions

Director of Science and Social Science Libraries

Science Librarian Team Leader and Librarian for Biology and Agriculture

Preservation and Data Management Services Librarian

Cataloging and Metadata Librarian

Librarian for Psychology, Sociology and Communications

Engineering Librarian

Collection Development Librarian (UC Health Center)

Software Engineer Lead

Software Engineer

Digital Projects Manager

Geospatial and Research Data Manager

Collections & Scholarly Communications Librarian

Metadata Librarian

### **Other**

All liaison librarians are responsible for data management advising.



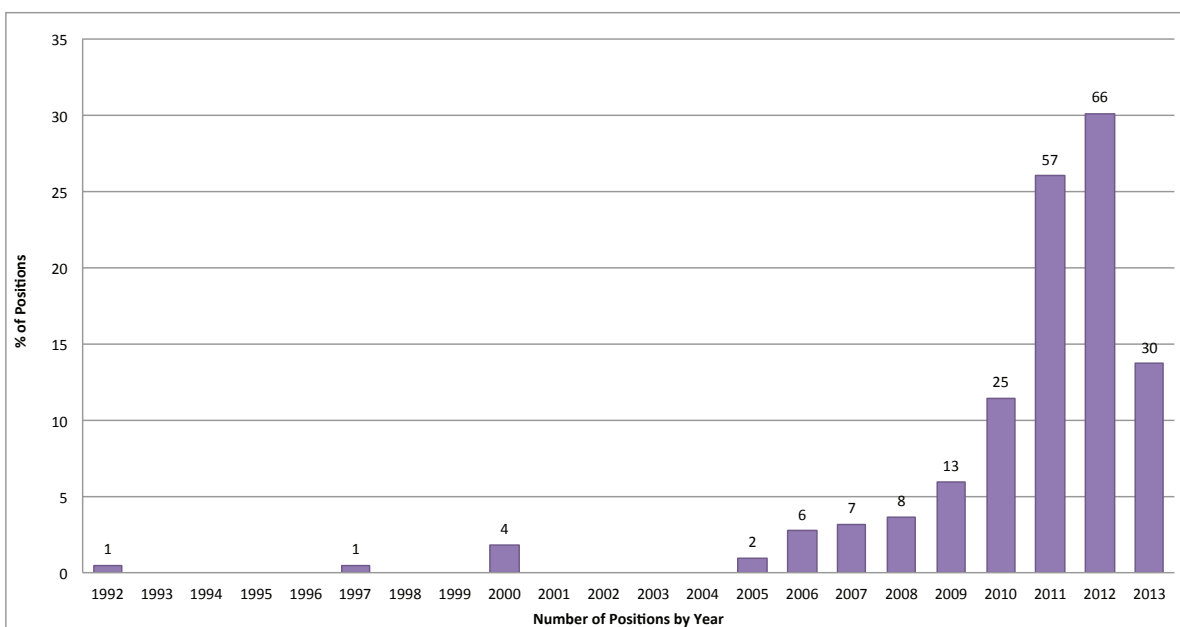
39. For each position, please indicate if it was created specifically for RDM services, substantially redesigned to focus on RDM, or had RDM activities added to existing responsibilities. N=53 respondents, 230 positions

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

**Comment**

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

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



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

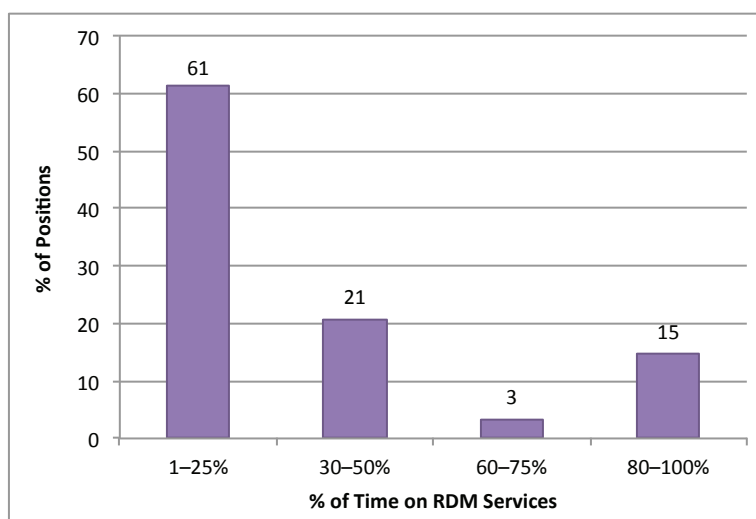
Permanent or Temporary	N	Positions	% of Positions
Full-time (permanent)	51	203	90%
Part-time (permanent)	4	8	4%
Temporary (full-time or part-time)	11	13	6%

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

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

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

Minimum	Maximum	Mean	Median	Std Dev
1	100	31.99	20.00	31.73



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

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

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

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

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

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

Analog Media preservation.

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

Copyright and digital scholarship guidance.

Copyright and Scholarly Communication

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

Digital projects, collection development

Digital Repository Coordination

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

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

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

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

Library instruction and faculty liaison work.

Manager of digital repository.

Other digital project related work.

Other tasks could include too many activities to detail here.

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

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

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

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

Position 6 works with archive system development.

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

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

Repository Administration.

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

Statistical analysis.

Teaching, scholarly communication; digital repository development.

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

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

**Please specify the other degree. N=4**

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

JD (2 responses)

PhD in Information Science

**Non-graduate Degrees**

Bachelor in computer and information technology

BA, economics

BS, Biochemistry

Geography, English, Curriculum Design

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

### Fields of Study

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

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

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

Biology

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

Biology, Bioinformatics, Computer Engineering

Computer and Information Sciences

Computer Science; Religion & Literature

Computing Science, Political Science

Digital Humanities

Earth sciences

Engineering

English/Digital Humanities; Political Science; GIS

Environmental Science

Geography

History

Masters in Instructional Technology

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

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

PhD anthropology/archaeology; MS Geographic Information Systems

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

Philosophy, Geosciences, Plant and Soil Science, Higher Education

Political Science and Biochemistry

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

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

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

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

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

Psychology, Neurobiology, English

Russian Language & Literature

Russian; Engineering

---

## STAFFING: SKILLS & TRAINING NEEDS

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

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

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

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

### Research Methods and Data Analysis

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

Data mining; database creation; data analysis

Research methods and data analysis

## Research Data Management

Research data management practices

Self-education in data management and big data

## Data Curation

Data curation

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

Metadata

## Scholarly Communication

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

Policy/copyright/IP management

Scholarly communications training and experience

## Other

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

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

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

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

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

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

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

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

Research methods and skills

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

Systems Administration

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

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

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

**Please specify other training methods. N=10**

Cross-training among librarians

Dorothea Salo's Introduction to Research Data Management course

ICPRS Training

Independent research on best practices

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

Participation in data curation pilot projects with related organizations.

Scientific Data Management course at a local MLS Program.

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



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

Webinars

## FUNDING

---

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

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

### Please specify any other source of funding. N=5

Contracts for services

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

IT internal budget

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

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

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

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

Government of Canada

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

National Library of Medicine

NSF-EPSCOR

## RDM SERVICE OUTREACH & ASSESSMENT

---

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

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

Please briefly describe any other outreach methods. N=13

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

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

Faculty outreach through a survey.

Held Campus Data Summit and other promotion/outreach events.

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

Posting flyers in campus buildings.

Print promotional materials

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

Social media

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

Webpages and LibGuides, communication with liaison librarians

Word of mouth

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

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

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

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

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

##### Answered Yes

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

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

Surveys, feedback forms/evaluations, tracking consultations.

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

##### No, but we plan to

As yet undetermined.

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

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

Faculty survey to be distributed in the very near future.

Follow up surveys to users; focus groups.

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

Not sure yet. (2 responses)

Number of consultations. Metrics are still under development.

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

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

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

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

assessment are being worked on.

Survey.

Survey and data curation profile interviews.

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

Uncertain at this time.

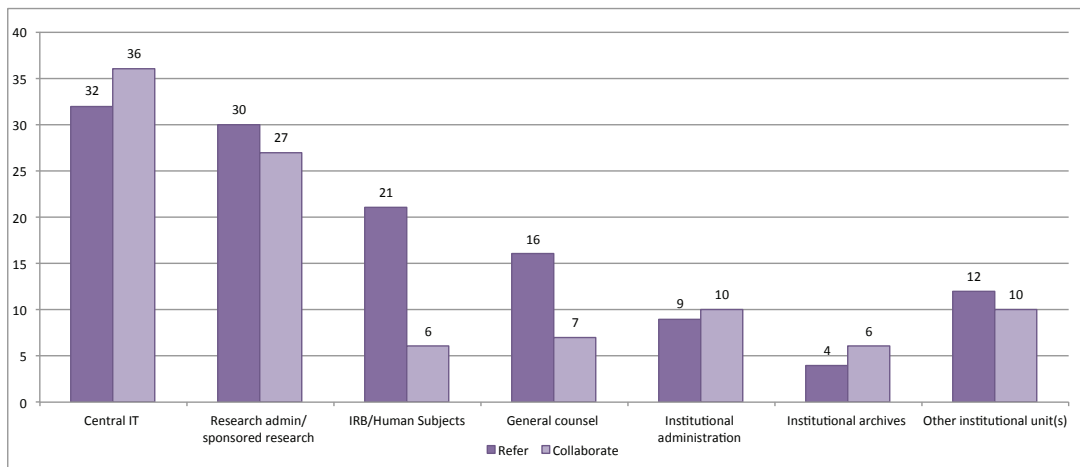
Usage of university sponsored DMP forms. Increased use of the institutional repository.

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

## PARTNERSHIPS

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

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



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

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

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

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

College and departmental IT groups

Data Coordinating Center, Clinical and Translational Science Institute

Departmental or school IT

Departments/institutes on campus

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

Institutes, laboratories, other data producers on campus

Institutional repository

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

School of Information Science

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

TACC

The Odum Institute, RENCI, DICE

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

Yes 13 26%

No 38 74%

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

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

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

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

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

DMPTool.

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

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

Participate in consortial committees through Ontario Council of University Libraries.

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

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

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

We want to.

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

## RDM SERVICE CHALLENGES

---

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

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

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

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

### Collaboration Campus-wide

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

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

Collaboration with campus units.

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

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

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

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

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

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

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

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

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

### Funding

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

admin buy in).

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

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

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

Describing the need and justifying a budget.

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

Financial support for service

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

Funding and support for program and services.

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

Implementing sustainable services.

Insufficient resources to significantly move this forward as a service.

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

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

## **Faculty Engagement**

Attracting participants in training sessions.

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

Convincing researchers to speak to the library about data management.

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

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

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



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

Reaching grant writers well before their deadlines.

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

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

### **Technology Infrastructure**

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

Developing the technology infrastructure.

Digital preservation for research data.

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

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

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

Providing archiving and storage options.

Providing support while lacking a data repository.

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

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

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

We lack an appropriately configured repository for optimal data storage.

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

### **Limited Staffing**

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

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

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

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

Lack of staffing devoted to RDM.

Limited staffing.

No dedicated staff to data management.

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

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

Stretching existing staff capacity.

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

We do not have current staffing to support RDM services.

## **Marketing Services**

Awareness of services.

Building awareness among library colleagues and campus administrators.

Campus awareness of services continues to be low.

Communication of services offered.

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

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

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

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

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

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

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

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

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

## **Staff Training**

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

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

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

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

Educating librarians on research methods.

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

Education and training of liaison librarians.

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

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

## Scoping Services

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

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

Challenge: defining needs. Measures: faculty survey.

Creating a clear services statement.

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

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

Defining and scoping services.

Getting the service off the ground, planning, etc.

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

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

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

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

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

## Staff Roles

Balancing traditional library services with RDM.

Challenges on understanding and accepting roles for data management.

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

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

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

Getting outreach librarian supervisors to make RDM a priority.

Motivating outreach librarians to have research conversations with faculty.

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

### **Institutional Commitment**

Institution-wide leadership on RDM issues is lacking.

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

Lack of institutional and administrative support

Lack of institutional priority

### **Faculty Education**

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

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

Educating/knowledge of faculty in understanding data management

Finding time for the boot camp was during spring break.

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

### **Evaluating Demand**

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

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

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

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

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

### **Policy**

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

### **Scaling Services**

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

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

Scaling for consultation services [build campus service catalog]

**Agency Ambiguity**

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

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

**Other**

Creating local infrastructure.

Finding participants for pilot projects.

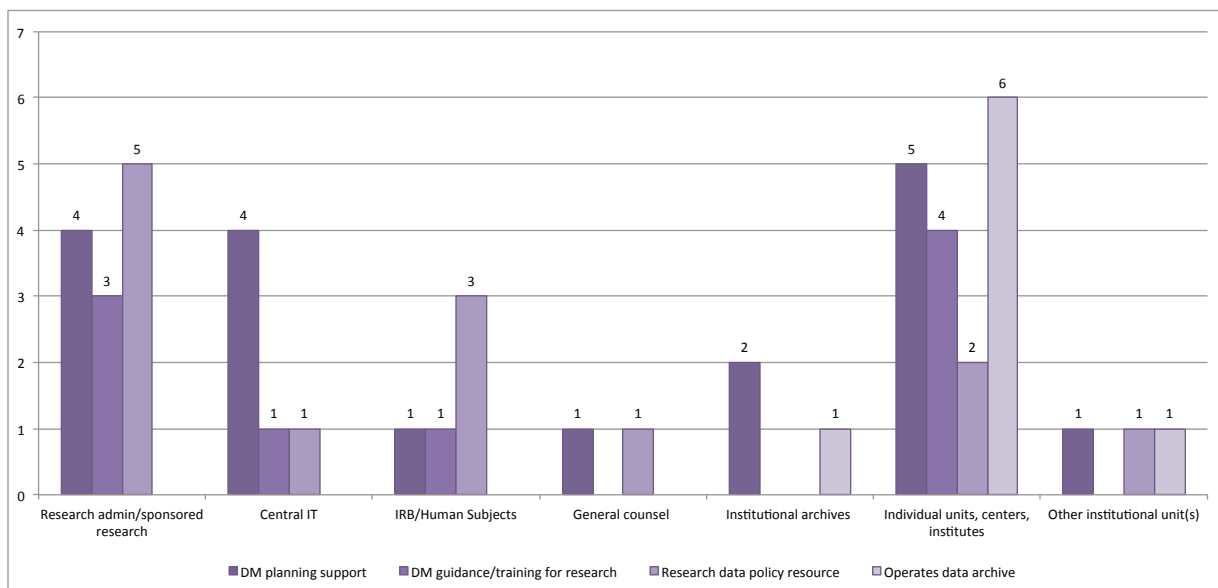
The library did not launch a focused effort to leverage resources effectively and efficiently for entire campus.

**OTHER UNITS OFFERING RDM SERVICES**

---

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

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



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

### DM Planning Support

Research Computing Center: Data Management Planning/Guidance/Training

### Research Data Policy Resources

University Libraries and the Department of High Performance Computing

### Operates a Data Archive

Special Collections & Archives, within the Library

#### Additional Comments

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

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

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

## FUTURE PLANS FOR RDM SERVICES

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

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

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

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

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

### Within 1 year

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

Dataviz tools and guidance (tentative)

Expand archiving in the local data repository

Metadata creation and related quality control

More robust data curation services

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



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

Training opportunities

### **Within 2 years**

Expansion of identifier assignment, assistance with sensitive/restricted data

Statistical reference services; database creation

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

### **Additional Comments**

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

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

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

### **Staffing**

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

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

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

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

### **Assessment**

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

management planning at this point.

## Planning

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

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

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

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

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

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

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

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

## Please briefly describe the other repository model. N=9

Chronopolis

Dark replicated archives, library-hosted

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

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

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

This process is still under discussion.

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

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

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

## **FUTURE PLANS FOR RDM STAFFING**

---

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

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

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

### **New Positions and/or Postdocs Being Added**

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

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

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

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

Research Data Coordinator

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

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

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

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

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

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

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

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

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

### **New Positions Proposed/Hoped For**

Possibility of hiring Post Docs.

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

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

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

We hope to add one librarian position.

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

### **No New Positions/Repurposing Existing**

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

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

### **Position Planning/Assessment**

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

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

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

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

We haven't given this any thought, yet.

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

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

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

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

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

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

Data visualization support

Digital arts and humanities

Digital Collections creation and management

Research using the data

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

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

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

Education/training in statistics

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

Informatics

Legal training for the Copyright Services Librarian role

Metadata

Project management

Some experience working with research data.

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

Statistics; R

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

## FUTURE PLANS FOR RDM FUNDING

---

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

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

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

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

Contract for services

ITS internal budget, special projects budget

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

university units, central administration, or perhaps the government.

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

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

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

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

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

This is currently unknown.

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

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

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

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

We haven't given this any thought, yet.

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

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

---

**NO LIBRARY RDM SERVICES**

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

Little to no demand.

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

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

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

Comment

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

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

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

## ROLE OF RESEARCH LIBRARIES

---

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

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

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

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

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



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

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

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

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

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

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

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

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

I think it will only increase.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

## **ADDITIONAL COMMENTS**

---

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

## RESPONDING INSTITUTIONS

---

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





## **REPRESENTATIVE DOCUMENTS**





## RDM Web Pages

---

HOME



UC San Diego  
Research Cyberinfrastructure

**RCI, a campus-wide initiative, frees researchers to do research, and is:**

- **Reliable and state of the art**
- **Easy to use**
- **Environmentally green**
- **Cost effective**

## RCI: Research Cyberinfrastructure

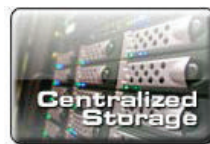
### Create... Share... Discover

RCI is a UCSD-sponsored program that offers campus researchers facilities, storage, data curation, computing, and networking to facilitate their research using shared cyberinfrastructure services across campus.

The RCI program is designed to provide cost-effective, reliable services which can be utilized by UCSD principal investigators in their current research efforts and incorporated in proposals for future research. In general, these services are available to researchers at a reduced cost, supplemented by the RCI program.

RCI has a number of services which are described in more detail via the links below. Some of these services are already available now in production, while other services are in pilot phase to best determine researcher requirements and appropriate business models.

### RCI Services



#### **CENTRALIZED STORAGE**

Centrally-administered disk storage featuring high performance, accessibility, reliability, and scalability



#### **COLOCATION SERVICES**

Energy-efficient, centrally managed datacenter space for hosting computer equipment and related components



#### **COMPUTING**

High-performance computing with fast interconnect, large memory options, and high I/O bandwidth for data analysis



#### **DATA CURATION**

Consulting services that help researchers with data management plans and long-term curation of research data




#### **NETWORKING**

An uncongested, leading-edge network that facilitates research collaborations, data exchanges, and access to the colocation facility



#### **TECHNICAL EXPERTISE**

Human expertise to optimize utilization of RCI services in the context of individual research projects



**EMORY LIBRARIES** | Robert W. Woodruff Library

Main Library » Guides » Research Data Management Admin Sign In

## Research Data Management

Tags: data, e-research, electronic\_data\_center, research data management

Research Data Management at Emory University

Last Updated: Jun 18, 2013 | URL: <http://guides.main.library.emory.edu/datamgmt> | [Print Guide](#) | [RSS Updates](#) | [SHARE](#) [f](#) [t](#) [e](#)

**Overview** | Data Management Plans | Data Storage and Preservation | Data Sharing and Re-Use | Data Documentation | Data Publication

Campus Resources | More Information | Faculty Survey Results | Researcher Interviews

Overview [Comments\(0\)](#) [Print Page](#) Search:  This Guide [Search](#)

### Need Help?

Need assistance with preparing a data management plan for a grant application? Looking for guidance on how to manage your research data? Research Data Management services at Emory University can help. Please contact us at [dataplans@emory.edu](mailto:dataplans@emory.edu).

Receive information and updates about developments in research data management by requesting to be added to the [RDM-L mailing list](#).

[Comments \(0\)](#)


### Research Data Management

Many funding agencies, including the National Science Foundation (NSF), the National Institutes of Health (NIH), and the National Endowment for the Humanities (NEH), require a data management plan as a component of grant applications. This requirement encourages researchers to consider in greater detail how their data will be preserved and shared.

Depending on the particular research community, data can include spreadsheets, images, videos, audio files, text files, models, computer software and code, patient records, interview transcripts, survey results, field/lab notes, and physical objects such as artifacts and samples.

[Comments \(0\)](#)

### Data Management Specialist




**Jennifer Doty**

**Contact Info**  
Electronic Data Center  
217 Woodruff Library  
404-727-0498 (P)  
[jennifer.doty@emory.edu](mailto:jennifer.doty@emory.edu)  
[Send Email](#)

**Links:**  
[Website / Blog](#)  
[Profile & Guides](#)

**Subjects:**  
[Research Data Management](#)

### Creating Data Management Plans



The **DMP Tool** can be used to prepare data management plans required by specific funding agencies.

Click the *Get Started!* button, select *Emory University* from the pull-down menu, enter your Emory Network ID and password, and create a new plan. The tool will walk you through each section, allowing you to save and revisit your plans.

[Comments \(0\)](#)

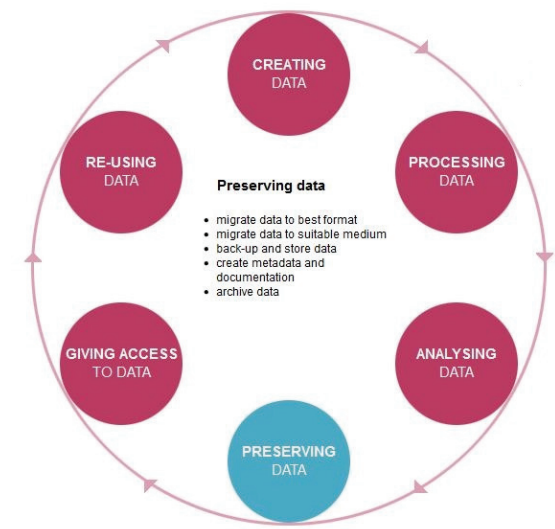
### Benefits of Research Data Management

Organizing, preserving, and sharing data will . . .

- improve data integrity.
- prevent data loss due to workforce turnover or hardware/software transitions.
- avoid unnecessary duplication of research efforts.
- help validate research findings.
- enhance the visibility of a researcher's work.
- lead to repurposing of data beyond its original intended use.
- ensure that the results of publicly-funded research become public property.

[Comments \(0\)](#)

### Research Data Lifecycle





The diagram illustrates the Research Data Lifecycle as a circular process with six stages: CREATING DATA, PROCESSING DATA, ANALYSING DATA, PRESERVING DATA, GIVING ACCESS TO DATA, and RE-USING DATA. The 'PRESERVING DATA' stage is highlighted in blue and includes the following sub-points:

- migrate data to best format
- migrate data to suitable medium
- back-up and store data
- create metadata and documentation
- archive data

### Research Data Management at Emory

Documentation of developments supporting research data management at Emory University Libraries:

-  [Exploring Research Data Management](#)
-  [Developments in Research Data Management Support](#)

[Comments \(0\)](#)

**Research Data Services**  
DIGITAL CURATION AT THE UNIVERSITY OF WISCONSIN-MADISON

Writing a Data Plan | Managing Data | Sharing Data | Events | Blog | About

**Datapoints: The RDS Blog**

**DMPTool Webinar Series Continues** DMPTool Webinar Series Brown Bag Join us for a ~15 part webinar series on the Data Management Planning Tool, DMPTool, from the California Digital Library. This series will introduce the tool, discuss ...

**VIVO Webinar Series** Overview of VIVO What is VIVO with Brian Lowe, Cornell University Implementation with Jon Corson-Rikert, Cornell University Future Directions with Dean Kraft, Cornell University Slides from the present ...

**LabKey Server** LabKey Server is an open source data management platform designed for organizing and managing data from large-scale research; for example, data from thousands of samples and/or subjects. It provides a ...

**Electronic Lab Notebooks** What are they? Electronic Lab Notebooks (ELNs) are software counterparts to paper lab notebooks. Although the name suggests a physical notebook device, ELNs are actually just software that runs on a c ...

[view full blog](#)

**News**

- ICPSR Updates its Data Management & Curation Site
- Google Alerts RSS delivery is temporarily not available. To keep receiving Google Alerts in the meantime, you can change to email delivery.
- 2013 Biennial ICPSR Meeting

[More news >>](#)

**Events**

Today [Navigation] Loading...

Events shown in time zone: Central Time [Google Calendar](#)

**too.** **Data Management Plan**  
For NSF, NIH, IMLS grant application  
Or for any research project  
Contact us now!

1 2 3 4 5 [Stop](#)

**Campus Links**  
UW-Madison  
MyUW  
UW-Madison Libraries  
DoIT

**Campus Resources**  
Center for High Throughput Computing  
Data & Information Services Center  
Digital Media Center  
Experts Guide  
Institute for Clinical & Translational Research

Records Management  
Research & Sponsored Programs  
Researcher Directory  
Scholarly Communication & Publishing  
UW Digital Collections Center

**Stay in Touch**  
[iTunes U](#)  
[YouTube](#)  
[Facebook](#)  
[Twitter](#)

© 2011-2013 The Board of Regents of the University of Wisconsin System  
Suffusion theme by Sayontan Sinha

## IRs and Data Archives



UF | George A Smathers Libraries University of Florida Digital Collections

UFDC Home myUFDC Home | Help | RSS

PRINT SEND ADD SHARE

# IR @ UF

The Institutional Repository at the University of Florida

Search Collection:  Go 1 2 3 4 5 6 7 8

HOME ADVANCED SEARCH TEXT SEARCH ALL ITEMS NEW ITEMS PROJECT-THESES SERIALS BY COLLEGE

The **Institutional Repository at the University of Florida** is the digital archive for the intellectual output of the University of Florida community, and includes research, news, outreach, and educational materials.

[Publish](#) Submit your research ([FAQ](#) and video of the process)

The University of Florida Libraries established and supports the *IR@UF* in order to offer a central location for the collection, preservation, and dissemination of scholarly, research, and creative production alongside historical materials from the University of Florida. The historical materials provide context for research and researchers, enabling insight into the history, nature, and culture of the University. The *IR@UF* includes the following [open access](#) materials from UF authors and UF colleges:

- Journal Articles
- Conference Papers and Proceedings
- Monographs and Monograph Series
- Technical Reports
- Theses and Dissertations
- White Papers
- Data and data sets (standalone or with publications)
- Journals and Other Publications of UF Colleges
- Grant Proposals
- Materials from the University Archives, such as graduation programs, photographs, audio and video of recent and historic campus events and people, campus directories and some yearbooks

The *IR@UF* encourages university units to contribute their [open access](#) research, reports and other materials to the *IR@UF* for archiving and dissemination free of commercial cost.

If you don't have a GatorLink account yet, you can create an account from the sign on page of the myUFL Portal. [Create a new account >](#)

An RSS feed from the *IR@UF* keeps subscribers up to date on all new submissions. To subscribe, click [here](#).

### COLLECTIONS

Chemical Engineering Documents

Florida Entomologist

Historic Preservation Documents


Preserving History: Projects from Internships in Historical Archives



Sustainability @ UF

University of Florida Institutional Repository

[Contact Us](#) | [Permissions](#) | [Preferences](#) | [Technical Aspects](#) | [Statistics](#) | [Internal](#) | [Privacy Policy](#)

© 2004 - 2011 University of Florida George A. Smathers Libraries.  
All rights reserved.  
[Terms of Use for Electronic Resources and Copyright Information](#)



*Share, Cite, Reuse, Archive Research Data*  
 Scientific data for reproducible research

## Harvard Dataverse Network

POWERED BY THE **Dataverse Network** PROJECT v. 3.5.1

[Create Account](#) [Log In](#)

[Advanced Search](#) [Tips](#)

The Harvard Dataverse Network is open to all scientific data from all disciplines worldwide. It includes the world's largest collection of social science research data. If you would like to upload your research data, first create a dataverse and then create a study. If you already have a dataverse, log in to add new studies. Learn more about the [Dataverse Network](#).

### Dataverses


551 Dataverses

i A **Dataverse** is a container for research data studies, customized and managed by its owner.

RECENTLY RELEASED DATERVERSES

Kimlong Chheng	Jul 19, 2013
Urban Institute Data Repository	Jul 16, 2013
Kyle Cranmer	Jul 12, 2013
Greg Snyder	Jul 11, 2013
Dan Pemstein	Jul 10, 2013

[View More >](#)



**CfA Dataverses**  
Dataverses: 17  
 Harvard-Smithsonian Center for Astrophysics (CfA) and affiliated Dataverses

### Studies

52,204 Studies, 725,463 Files, 835,277 Downloads

i A **study** is a container for a research data set. It includes cataloging information, data files and complementary files.

RECENTLY RELEASED STUDIES

<b>Trees for Food Security Project</b> by Muthuri, Catherine; Iiyama, Miyuki ; Betemariam, Ermias; Kindt, Roeland; Gyau, Amos; Kiptot, Evelyn; Kuria, Anne; Luedeling, Eike; Mohan, Sid	Jul 22, 2013
<b>Replication data for: Los mitos de la redistribución. Parte 1: Malapportionment</b> by Javier Marquez	Jul 20, 2013
<b>Replication data for: Regime Legacies and Levels of Democracy</b> by Perez-Linan, Anibal; Mainwaring, Scott	Jul 19, 2013
<b>Replication data for: Relying on the Ground Game: The Placement and Effects of Campaign Field Offices</b> by Darr, Joshua; Levendusky, Matthew	Jul 19, 2013
<b>The Spanish Sovereign Debt Crisis: The Impact of Politics on Fiscal Outcomes in Subnational Governments</b> by Haswell, Ethan	Jul 19, 2013

[View More >](#)

MOST DOWNLOADED STUDIES

<b>Replication data for: A Multivariate Model of Strategic Asset Allocation</b> by John Y. Campbell; Yeung L. Chan; and Luis Viceira	17535
<b>Replication data for: Asset Prices, Consumption, and the Business Cycle</b> by John Y. Campbell	17411
<b>10 Million International Dyadic Events</b> by Gary King; Will Lowe	16617
<b>Measuring the impact of microfinance in Hyderabad, India</b> by Abhijit Banerjee; Esther Duflo; Rachel Glennerster; Cynthia Kinnan	14934
<b>Textbooks and Test Scores</b> by Paul Glewwe; Michael Kremer; Sylvie Moulin	14666


[View More >](#)

PRIVACY POLICY

Copyright 1997-2013 The President and Fellows of Harvard College


116 · Representative Documents: IRs and Data Archives





# UNIVERSITY OF NEBRASKA-LINCOLN

## Libraries



UNL ▸ **Libraries** ▾ UNL Data Repository ▸ Home

- Home
- E-Resources
- Services
- Ask a Question
- Libraries @ UNL
- About

### Home

UNL Libraries and Information Services created the UNL Data Repository to provide for the growing requirements by external funding agencies for data management and data sharing. This repository, designed to provide researchers with a secure site for storage of data collections that are no longer actively in use, allows the researcher to stably retain data for future use and/or sharing with other interested parties. The UNLDR exists to manage data in a manner that facilitates research and scholarly activities, and that simplifies access to vital and unique research data.

The data and project information deposited by researchers maintains its value over time with information (known as metadata) that outlines its importance, its long-term usability, and the dedicated efforts of those who were involved.

Questions about depositing data may be directed to DeeAnn Allison, Professor, University Libraries at [dallison@unl.edu](mailto:dallison@unl.edu) or 402-472-3944

### Costs

Proposers should include a line item in their project budget to cover the costs of data storage. This one-time charge should be requested during the final year of project support to ensure data storage and maintenance for a minimum of five years beyond the award period. Please use the following figures in your budget:


Gigabytes	Storage Costs
100*	\$500
200	\$1,250
500	\$2,500
1000	\$5,000

\* This is the minimum amount of storage that can be requested.

For more information regarding data storage, contact Kevin Murphy, Information Services, at 402-472-6466 / [kmurphy@unl.edu](mailto:kmurphy@unl.edu) or DeeAnn Allison, Professor, University Libraries at 402-472-3944 / [dallison@unl.edu](mailto:dallison@unl.edu)

Begin depositing your data.

Search Through Existing Public Data.



### YOUR FEEDBACK

Please rate this page: ★★★★★

Comments for this page

### RELATED LINKS

[Parking and Visiting UNL](#)  
[University of Nebraska Foundation](#)



### CONTACT US

UNL Libraries  
318 Love Library  
P.O. Box 884100  
University of Nebraska-Lincoln  
Lincoln, NE 68588-4100

### SHARE THIS PAGE


GET A GoURL

© 2011 University of Nebraska-Lincoln | Lincoln, NE 68588 | 402-472-7211 | [About UNL](#) | [comments?](#)  
UNL web templates and quality assurance provided by the [Web Developer Network](#) | [W3C](#) | [CSS](#)



SPEC Kit 334: Research Data Management Services · 117

The University of New Mexico UNM A-Z StudentInfo FastInfo myUNM Directory more ..



LoboVault Home UNM Research Data

UNM Libraries

- University Libraries
- Law Library
- Health Sciences Library

Search LoboVault

Search LoboVault  
 This Collection

[Advanced Search](#)

Browse

- All of LoboVault
  - Communities & Collections
    - Date
    - Authors
    - Titles
    - Subjects
- This Community
  - Date
  - Authors
  - Titles
  - Subjects

My Account

- Login
- Register

## UNM Research Data

Search within this community:

[Advanced Search](#)

This community is interdisciplinary in scope and contains research data sets produced by UNM faculty.

### Collections in this community

- [Bureau of Business and Economic Research Datasets](#) [1]
- [Business Faculty and Staff Datasets](#) [2]
- [Earth and Planetary Sciences Datasets](#) [1]
- [Physics and Astronomy Faculty & Staff Datasets](#) [2]
- [Spanish & Portuguese Faculty and Staff Datasets](#) [1]
- [University Libraries Faculty and Staff Datasets](#) [1]

### Recent Additions

[2013-07-05] <a href="#">Troublesome Concepts and Information Literacy [dataset]</a> Townsend, Lori; Brunetti, Korey; Hofer, Amy R. (2013-07-05)
[2013-05-23] <a href="#">Ground Water Data Supporting the Riparian Evapotranspiration (ET) Study (SEON) along the Middle Rio Grande Bosque, New Mexico [dataset]</a> Thibault, Jim; Dahm, Clifford; Cleverly, James (2013-05-23)
[2013-04-11] <a href="#">Legacy Data from Astronomical Observations [dataset]</a> McGraw, John; Zimmer, Peter (2013-04-11)
[2013-02-26] <a href="#">Colonia Population and Socioeconomic and Housing Characteristic Estimates, Maps and Shape File Update: November 2012 [dataset]</a> Ruiz, Daren (2013-02-26)
[2012-11-06] <a href="#">National Highway Traffic Safety Administration Regulatory Notice and Comment Periods [dataset]</a> Cavazos, David (University of New Mexico Libraries, 2012-11-06)

University Libraries  
MSC05 3020  
1 University of New Mexico  
Albuquerque NM 87131  
505.277.9100

[Accessibility](#) | [Legal](#) | [Contact Us](#) | [Send Feedback](#)

**UNIVERSITY OF OREGON** | UO Home | Dept Index

# UO Libraries

Faculty Students FAQ About Contact Login

Scholars' Bank Home

## Scholars' Bank

Purpose/Scope: Scholars' Bank is the open access repository for the intellectual work of faculty, students and staff at the University of Oregon. Open access journals, student projects, theses, dissertations, pre and post-print articles, instructional resources and university archival material are all candidates for deposit.

To contribute: To start depositing to Scholars' Bank please [send us a message](#).

### Communities in Scholars' Bank

Select a community to browse its collections.

- [Archives of President Lariviere's Dismissal](#)
- [Bicycle and Pedestrian Transportation Plans](#)
- [Data](#)
- [Dissociation](#)
- [Instructional Resources](#)
- [Local and Regional Documents Archive](#)
- [Renaissance Editions](#)
- [Scholarly Works](#)
- [University Archives](#)

### Search Scholars' Bank

Enter some text in the box below to search Scholars' Bank.

The screenshot shows the ScholarSphere website interface. At the top, there is a dark blue header with the Penn State logo and 'SCHOLARSHERE' text on the left, and navigation links 'Home', 'About', 'Help', 'Contact', and a 'Login' button on the right. Below the header is a search bar with the word 'SEARCH' in large letters. The main content area is divided into several sections: 'Browse By' with a list of filters (Resource Type, Creator, Keyword, Subject, Language, Location, Publisher, File Format); a featured article titled 'Participate in Usability Testing for ScholarSphere!' with a sub-header 'What is ScholarSphere?' and a paragraph describing the service; a 'Contribute' section with a 'Share Your Work' button and a 'Terms of Use' link; a 'My Latest' section with a login prompt; and a 'Recently Uploaded' section with an 'Open Access' tag and a table of three items. The footer contains the 'SCHOLARSHERE v1.7' logo, copyright information for 2012, and a list of service links: 'Penn State', 'University Libraries', 'Information Technology Services', and 'Accessibility'.

PENNSYLVANIA STATE UNIVERSITY  
SCHOLARSHERE

Home About Help Contact Login

SEARCH

**Browse By** **Participate in Usability Testing for ScholarSphere!**

Resource Type  
Creator  
Keyword  
Subject  
Language  
Location  
Publisher  
File Format

**Contribute**  
Share Your Work  
Terms of Use

**My Latest**  
Login to see your recently added documents

**What is ScholarSphere?**  
ScholarSphere is a secure repository service enabling the Penn State community to share its research and scholarly work with a worldwide audience. Faculty, staff, and students can use ScholarSphere to collect their work in one location and create a durable and citeable record of their papers, presentations, publications, data sets, or other scholarly creations. Through this service, Penn State researchers can also comply with grant-funding-agency requirements for sharing and managing research data.

**Recently Uploaded** Open Access

Nancy Ellen Adams	EBP and librarians... collection of audio re...
Mona Lee Ostrowski	SSHtransferExmpl.png SSHtransferExmpl.png TEST
Margaret Louise Signorella	Signorella, Hayes &... Signorella.Hayes.Li_Se... gender, single-sex schooling, meta-analysis

**SCHOLARSHERE** v1.7  
A service of the University Libraries and Information Technology Services.

Copyright © 2012 The Pennsylvania State University  
Penn State University Libraries Information Technology Services Accessibility

**PURDUE UNIVERSITY** | **Purdue University Research Repository** | PURR

Login Register Report a bug

Home Publications Projects Get Started Contact Us Search

## Publish Datasets with DOIs

Use PURR to publish datasets with Digital Object Identifiers (DOI) that make it easier for people to cite your data and give you credit. Purdue is a founding member of DataCite, the international agency that registers DOIs for data.

[Learn More](#)



```
<header>
<identifier> http://dx.doi.org/10.4231/D39P2W550
<datestamp>2012-10-24T12:28:53-04:00</datestamp>
<setSpec>Dataset</setSpec>
</header>
<metadata>
<dc:identifier>
<dc:creator>David F. Gleason
<dc:subject>Internet</dc:subject>
```

### Start Your Research Project



#### Create a Data Management Plan

Learn about the detailed requirements for your data management plan (DMP). Funding agency requirements are very specific and our DMP resources can help you to clear up any confusion. [Get Started](#) >



#### Upload Research Data to Your Project

Create a project to upload and share your data with collaborators using our step-by-step form to guide you through the process. Invite collaborators from other institutions to join your project. [Create a Project](#) >



#### Publish your Dataset

Package, describe, and publish your dataset with a DataCite DOI. Publishing will ensure your dataset is citable, reusable, and archived for the long-term. [See Published Datasets](#) >

### Featured Dataset

#### Evaluation of Function Predictions for Moonlighting Proteins

By Ishita K. Khan, Meghana Chitale, Catherine Rayon, Daisuke Kihara  
*Purdue University, Université de Picardie Jules Verne*



Supplemental datasets used for evaluation of function predictions for moonlighting proteins.

### Do you have a question?

Ask a Librarian

Your Question/Message

You will be prompted and required to enter an email address before your chat session begins.

[Send](#)

UNIVERSITY OF VIRGINIA LIBRARY Sign In

---

**LIBRA** Online Archive of University of Virginia Scholarship

Please Note: U.Va. users may [log in](#) to view items restricted to the University community.

Libra makes publications available to the world and provides safe and secure storage for the scholarly output of the U.Va. community.

  
  
[Terms of Use](#)

Limit results by:


**Type of Work**

[Article](#) (838)  
[Doctoral Dissertation](#) (213)  
[Master's Thesis](#) (85)  
[more +](#)

**Department or Academic Plan**

[Department of Computer Science](#) (862)  
[Department of Electrical and Computer Engineering](#) (35)  
[Institute for Parallel Computation](#) (35)  
[more +](#)

**Peer Reviewed** (32)



**Open Access Works**

Faculty scholarly works available for research, scholarship, teaching and learning in a central, stable location. [Learn more...](#)

**Electronic Theses & Dissertations**

Find current theses and dissertations from departments and schools around Grounds. [Learn more...](#)

**Datasets**

Datasets may now be deposited into Libra. Additional features will become available in the future. Please send us your [feedback](#). [Learn more...](#)



---

University of Virginia Library  
PO Box 400113, Charlottesville, VA 22904-4113  
ph: (434) 924-3021, fax: (434) 924-1431, [library@virginia.edu](mailto:library@virginia.edu)

[Depts./Contacts](#) | [U.Va. Home](#) | [ITC](#)  
[Libra Feedback](#) | [Questions? Ask a Librarian](#) | [Search](#) | [Hours](#) | [Map](#) | [Policies](#) | [Press](#) | [Jobs](#)  
© 2011 by the Rector and Visitors of the University of Virginia

---

This library is a Congressionally designated depository for [U.S. Government documents](#). Public access to the Government documents is guaranteed by public law.



## Data Management Plan Tools

---

## Indiana University Guidance on NSF Data Management Plans

September 28, 2012

Effective for proposals submitted on or after January 18, 2011, the National Science Foundation (NSF) requires the inclusion of a supplementary document of no more than two pages entitled "Data Management Plan" (DMP). The plan should describe how the proposal will conform to NSF policy on the dissemination and sharing of research results (see AAG Chapter VI.D.4 [1]).

Following is a guide to writing your DMP, consisting of the following sections:

- **Section 1** gives a template and consideration points for completing a data management plan.
- **Section 2** is a short set of boilerplate language for your use when composing your DMP.
- Because no single template works for every discipline and community, **Section 3** lists additional resources that could be helpful in figuring out what works for your needs.
- **Section 4** explains the Indiana University (IU) resources that are available for your use. It is useful if you want to use UITS storage and one of the institutional repositories (IUScholarWorks or IUPUIDataWorks) as your data preservation solution.

This document is prepared with fonts and margins consistent with the NSF Grant Program Guide, so researchers may cut and paste directly from this document when preparing data management plans.

*This document is derived, in part, from a report by a blue-ribbon panel of IU experts led by Professor Beth Plale, of the School of Informatics and Computing (SOIC). As such, this guidance reflects the combined effort and consensus thinking of IU's top experts in data management and reflects IU guidance for NSF Data Management plans endorsed by the Office of the Vice President for Research and offered in a manner that is consistent with the Indiana University Information Technology Strategic Plan [6].*

---

### 1. Data Management Plan Template

A data management plan meeting the general NSF requirement can be organized by the below template, though not all items will be relevant for all disciplines, Directorates, or solicitations. See [2] for discipline specific advice. It may be helpful to begin your DMP with a few sentences describing the research project in general, to provide general context for the detailed information in each section. In each section, you should describe your reasoning, particularly if you are deviating from common practice or standards used in your discipline or community of practice. Identifying a particular person or role to carry out these activities is also vital.

1. Describe the **types of data and products** that will be generated in the research such as samples, physical collections, software, and curriculum materials. Characterize the data with details such as the types of data (text, numeric, images, audio, video, etc.) and an approximate number and size of files to be generated or used. Provide a brief description of the data collection process, including instruments or tools, sites, and process for getting data into a secure location. In addition, briefly describe the storage/backup plan and the IU cyberinfrastructure to be used.
2. Describe the **format in which the data, metadata, and other products are stored**. Describe the formats in which the data will be stored, preferably using a common or open file format standard. Include a description of the metadata that will make the actual data products useful to the general researcher. Some examples of discipline-based metadata standards include the NanoParticle Ontology [3] and Ecology Metadata Language (EML) [4].
3. Describe the **policies for general access** to data including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements. "Access to data" refers to data made accessible without explicit request from the interested party. Policies for access and reuse should clearly when, how, and to whom the data will be made available. Describe the policies and mechanisms for access to the data and other products, including specific provisions (described in the next section) for appropriate protection of privacy, confidentiality, security,



intellectual property, or other rights. Mechanisms should provide for access beyond the life span of the project, preferably via institutional or community infrastructure (i.e., institutional or subject repository).

Reminder: NSF allows grantees to retain principal legal right to intellectual property developed under NSF grants.

4. Describe **policies and provisions for re-use, re-distribution, and the production of derivatives**. Clearly describe the proposed policies and rationale for limitations on others' ability to re-use, re-distribute, and produce derivatives of the data and other products. These policies may be developed in response to the ethical and legal issues identified in the previous question (#3).
5. Describe **plans for archiving data**, samples, and other research products, and for preservation of access to them. Identify the data that is appropriate for preservation and the means through which preservation of digital and physical materials will be ensured. If the data will be preserved by a third party, refer to their preservation plans. If the data will be preserved at your institution, describe the cyberinfrastructure that will be used.

Depositing data into an institutional or subject repository ensures access to the data beyond the life span of the project. If you are interested in using an Indiana University repository (IUScholarWorks, IUPUIDataWorks) for your data, see Section 4 below.

---

## 2. Boilerplate Language

### Introduction

This plan describes the management, dissemination, retention, and archiving of the research data produced during the proposed project. The staff of [INSERT YOUR DEPARTMENT OR LAB NAME HERE], with the assistance of the [IU Libraries-Bloomington/IUPUI University Library] and University Information Technology Services (UITS), will provide for sustainable discovery, access to, and preservation of these data for use by other researchers, instructors, and interested members of the public for the length of this project and at least three years beyond. This will be facilitated through data and publication deposits in existing open-access disciplinary and/or institutional repositories.

### Data Formats and Description

We will utilize the Dublin Core metadata scheme to capture information about the data collected during the course of our research. We will work with a metadata expert from the [IU Libraries/IUPUI University Library] to create a working template that captures each dataset's metadata throughout the research process. Upon completion, we will export this data to Dublin Core format, which conforms to the data submission requirements of the IUScholarWorks and many other relevant museums/repositories.

### Access, Re-Use, Re-Distribution, and Derivative Works Policies

[If no sensitive or personally-identifiable information is used, include this:]

All data produced during this research will be available freely to the public; we anticipate no sensitive or confidential data. Under the terms of the Creative Commons Zero Universal 1.0 Public Domain Dedication (CC0 1.0; <http://creativecommons.org/publicdomain/zero/1.0/>), users may share, create, and/or adapt these data/databases.<sup>1</sup>

---

<sup>1</sup> If you wish to retain attribution rights so that anyone who uses your data must credit you as the creator, IU recommends you apply the Open Data Commons Attribution License (ODC-BY; <http://opendatacommons.org/licenses/by/>) to your data instead of CC0. In your plan, replace the noted sentence with the following: "Under the terms of the Open Data Commons Attribution License (ODC-BY; <http://opendatacommons.org/licenses/by/>), users may share, create, and/or adapt these data/databases with proper attribution."

[If sensitive or personally-identifiable information is used, include this instead:]

Results, data, and collections will be made available to other researchers in a timely basis with [EXAMPLE] limitations. Sensitive and confidential data collected will be treated following [HIPAA/IRB] regulations, and an added layer of security will be implemented using [STRATEGIES SUCH AS DATA ENCRYPTION, RESTRICTED ACCESS, OR THE SEPARATION OF IDENTIFIABLE DATA]. Under the terms of the Creative Commons Zero Universal 1.0 Public Domain Dedication (CC0 1.0; <http://creativecommons.org/publicdomain/zero/1.0/>), users may share, create, and/or adapt data/databases made freely available.<sup>2</sup>

#### **Data Archiving and Preservation**

To increase access to the published research that has been funded, the research collaborators will deposit peer-reviewed or pre-print manuscripts (with linked supporting data where possible) in the [IUScholarWorks/IUPUIScholarWorks]<sup>3</sup> institutional repository. Other works, including presentations and white papers, will also be made accessible via the [IUScholarWorks/IUPUIScholarWorks] institutional repository.

Digital data will be stored using the Indiana University Scholarly Data Archive (SDA; <https://pti.iu.edu/storage/sda>), a distributed storage service that is centrally supported across mirrored tape silos in Bloomington and Indianapolis. Data stored on the SDA that will be made freely available will be archived in the [IUScholarWorks/IUPUIDataWorks] repository, which will provide a user-friendly interface for the organization, context, and discoverability of data. This combination of [IUScholarWorks/IUPUIDataWorks] and the SDA provides mirroring, redundancy, media migration, access control, file integrity validation, embargoes, and other security-based services that ensure the data are appropriately archived for the life of the project and beyond the project if necessary.

---

### **3. Resources**

NSF funds a wide range of research. Some directorates and programs have provided specific guidance, which can be found at Dissemination and Sharing of Research Results [2]. In the absence of specific guidance, the Award & Administration Guide (AAG) Chapter VI.D.4 [1] applies.

Data management plan examples spanning a range of disciplines are available from the Inter-University Consortium for Political and Social Research [6]. Additionally, a Data Planning Checklist [7] can be helpful in preparation.

**For more help:** Skilled Librarians and grant writers are available to assist you in developing a data management plan, identifying appropriate data and metadata standards, finding resources on developing policies for sharing and reuse of data, locating community- or discipline-based data repositories, and finding resources on data management and preservation. To arrange a consultation to meet your needs, contact the data services program for your campus<sup>4</sup>:

---

<sup>2</sup> If you wish to retain attribution rights so that anyone who uses your data must credit you as the creator, IU recommends you apply the Open Data Commons Attribution License (ODC-BY; <http://opendatacommons.org/licenses/by/>) to your data instead of CC0. In your plan, replace the noted sentence with the following: "Under the terms of the Open Data Commons Attribution License (ODC-BY; <http://opendatacommons.org/licenses/by/>), users may share, create, and/or adapt these data/databases with proper attribution."

<sup>3</sup> IUPUIDataWorks is the data repository for the IUPUI campus. All other IU campuses should use IUScholarWorks as their institutional repository of choice.

<sup>4</sup> A list of regional campus research data specialists can be found on the [IUScholarWorks Data Management Service website](#).

**IUB**

**Stacy Konkiel, E-Science Librarian**

Email: [iuswdata@indiana.edu](mailto:iuswdata@indiana.edu)

Phone: 812-856-5295

**Meryl W. Bertenthal, Proposal Development Specialist**

Email: [mbertent@indiana.edu](mailto:mbertent@indiana.edu)

Phone: 812-856-5245

**IUPUI**

**Heather Coates, Digital Scholarship & Data Management Librarian**

Email: [dataserv@iupui.edu](mailto:dataserv@iupui.edu)

Phone: 317-278-7125

**Ann Kratz, Proposal Development Manager**

Email: [akratz@iupui.edu](mailto:akratz@iupui.edu)

Phone: 317-274-6732

---

## 4. IU Storage Systems and Institutional Repository

University Information Technology Services (UITs) maintains a large suite of storage systems. These are described in an extensive document that can be used (in whole or in part) in the Facilities section of an NSF proposal. This document is available online in a link accessible from: <http://kb.iu.edu/data/anwu.html>. It describes the storage systems operated by UITs and the backup facilities and plans for those storage systems. It also describes data security.

Indiana University has institutional repositories for archiving scholarly and scientific works called IUScholarWorks [8] at Bloomington and IUPUIScholarWorks [9] at Indianapolis. These repositories will accept digital data generated by IU researchers and from National Science Foundation funded efforts with PIs outside IU when there is a formal collaboration with an IU researcher (documented via a Memorandum of Understanding or via a Statement of Work associated with funding to an IU researcher as part of said project). IU, through IUScholarWorks and IUPUIDataWorks, will provide replicated storage of all data sets (as described in detail in the cyberinfrastructure facilities statement).

Both IUScholarWorks and IUPUIDataWorks accept data in all formats. Classified or confidential data requiring formal, contractual, or legal restrictions to access, such as HIPAA-designated Protected Health Information, will not be accepted for deposit, but may be stored on the SDA. In this case, the searching and metadata management facilities that help make these repositories so valuable in disseminating data are not available for use. However, de-identified datasets are eligible for deposit into the repositories. The PI is responsible for ensuring that datasets are appropriately and fully de-identified.

If you intend to use one of the institutional repositories, you should consider budgeting funds for data management:

1. Funding for a person to manage data and metadata. This may be funded within your own research group, or you may consider a consulting arrangement with the IU Bloomington Libraries (contact [iuswdata@indiana.edu](mailto:iuswdata@indiana.edu)) or the IUPUI University Library (contact [dataserv@iupui.edu](mailto:dataserv@iupui.edu)).
2. Funding for storage services for exceptionally large data storage needs (more than 50 TB). In this case, please contact [researchtechnologies@iu.edu](mailto:researchtechnologies@iu.edu) for more information. UITs may be able to offer storage above the default 50 TB limit as part of matching support for grant proposals.

NSF allows for adding data management costs to your proposal (typically Line G2).

## Questionnaire to Help with the Creation of a Data Management Plan

JHU Data Management Services of the Sheridan Libraries; [datamanagement@jhu.edu](mailto:datamanagement@jhu.edu)

### How to use this document

This questionnaire distills NSF’s guidelines for what to address in your data management plan. You can use the section headings in your own document. The questions can help you structure the content of each section of the plan. The table in section 1 facilitates listing the different data types for your study. Some researchers are including the table in their plan. Please note that you may not need to address all questions under a numbered category, and in some cases, you may not need to address each category, though any omissions should be justified in your plan. See endnotes for more tips, (view by placing mouse over the blue numbers in text). If you have any questions on the content of this questionnaire, please contact a JHU data management consultant at [datamanagement@jhu.edu](mailto:datamanagement@jhu.edu).

### 1. Data Products and Standards

#### Research Outputs

	Data Source	Data Product	Format(s)	Estimated Size or Amount
1				
2				
3				

You may include this table and use numbers to reference corresponding data types in your plan, or use numbered text paragraphs if needed to fit the 2-page format. [1](#)

#### Data and Metadata Standards [2](#)

1. Do the listed data products use standards for formats or metadata, and why are you using them? If not, will your project develop and maintain standardized formatting and metadata?
2. What details (metadata) are necessary for others to use your data?
3. How will metadata be generated (automatically, manually, or both)?
4. What naming conventions/schema will be used for your data, if any?
5. What data dictionaries/taxonomies/ontologies will you use for your data, if used within your field?
6. How will lineage/provenance of some or all of your data be documented (e.g., processing steps executed on raw source data)?
7. What tools will be required to read the data (e.g., software, instruments)?

### 2. Data Storing and Long-Term Preservation [3](#)

#### Storage during project

8. What digital and non-digital data will be retained during the project?
9. How (i.e., media) and where (i.e., location(s)) will the data be stored and who is responsible for it?
10. How and where will the data be backed up and who is responsible for it?
11. If data need to be secured through access controls (e.g. password-protected network space), how will they be applied? (e.g. local passwords, institutional LDAP or Shibboleth)
12. If data are stored in an unusual or not generally accessible format, will they be converted to a more common format for storage or sharing?

Revised 2/16/12

Preserving data after project

13. Which digital and non-digital data will be stored or archived after the project? Why will you preserve these data?
14. Will “raw data” (not processed, analyzed or associated with publications) be relevant to store for reuse in your or others’ future projects? If so, describe.
15. Where and for how long will data be stored or archived after the project? 4
16. Who will manage and administer the stored or archived data?
17. Will security and access codes be retained on archived data after the project? How?
18. If using a service other than your project group to archive research data, please describe the services that the archive provides in preserving and disseminating research data. Will there be a formal archiving agreement? 5

**3. Data Sharing**

Research to be shared 6

19. Of the data products generated during the project, which data will be shared? 7
20. Which data will be publicly-accessible, if at all? 8
21. When will you share those data? 9
22. How will the data be shared with other stakeholders? (e.g., made available for general access through a public website or database, or released only upon specific request from an interested party. Specify any 3<sup>rd</sup> party resources or services used.)
23. Who is expected to use the shared data?

Policies for access and sharing 10

24. Identify who will be allowed to use your data, and how data are to be used and disseminated. Explain any restrictions on re-use, production of derivatives and how you will communicate these restrictions, (e.g., requiring citation, or Creative Commons licensing.)
25. Are there any data with confidentiality issues (e.g., embargo period)? If so, what are the conditions of use, sharing, and dissemination?
26. Are there any data with specific security or regulatory concerns with sharing (e.g. classified information or FDA handling requirements), and how will they be addressed?
27. Are there any data with intellectual property (e.g., patent, copyright) concerns with sharing? If so, what are the conditions of use, sharing, and dissemination?
28. Are there any data with privacy concerns to sharing (e.g., human subjects)? If so, what policies need to be adhered to and how will policies be enforced? 11
29. Is any of the data owned by someone else? If so, what are the conditions of use, sharing, and dissemination?

This work is licensed under the Creative Commons Attribution--Noncommercial 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc/3.0/> .

**Additional Tips and Instructions** (See corresponding endnote number in the text)

1 **Data source** can include instruments, people, and data centers. **Data product** examples: transcripts, tables, 3D models, digital audio, geospatial data. Format examples: RTF text, MS Excel converted to CSV, MATLAB, WAV audio, shapefile. (Specify any instrument-specific formats or software packages). Estimated amount can include rate produced, e.g. 1 TB/year, 50GB/experiment. Include any sources and data products created by others that you are using. It may help to think through the steps of your research workflow to identify data types and sources requiring management.

2 Metadata is the information that captures the who, what, when, where, why and how of your data, providing the details necessary for another researcher to use your data sets. Some scientific communities have established metadata standards, such as Content Standard for Digital Geospatial Metadata (CSDGM), Data Documentation Initiative (DDI), Climate and Forecast (CF) metadata convention, and Dublin Core. Metadata may take the form of “readme files” that explain variables and file structures; however, it is preferable if metadata files are machine readable for better re-usability and processing.

3 Storing data is defined differently than archiving data. Storage is a necessary step towards archiving your data; however, storing data (e.g., on an external drive) does not safeguard against media degradation (e.g., CD file corruption), obsolescence of data formats (e.g., VisiCalc spreadsheets) or providing easy access in the future. Archiving encompasses both active preservation of the digital object and increased discoverability and access to those data. Your plan should discuss how you will store your research data during the project and your preservation strategy for after the project, particularly of research data that will be reused and shared. The next two sections help frame these different topics.

4 JHU requires retention of research data for a minimum period of 5 years after the date of any publication upon which it is based ([http://jhuresearch.jhu.edu/Data\\_Management\\_Policy.pdf](http://jhuresearch.jhu.edu/Data_Management_Policy.pdf)). The NSF Engineering Directorate requires retention for 3 years after conclusion of the award or 3 years after public release, whichever is later.

5 Different data archives provide different kinds of services, such as the creation of persistent, unique identifiers for citation, format migration, disaster recovery plans, and free, publicly-accessible downloading of data files. If you plan to use a data repository, we strongly recommend that you contact the repository to ensure that their archive can handle your data, and determine their archiving fees to include in your budget. Johns Hopkins University has built a research data archive. Please contact [datamanagement@jhu.edu](mailto:datamanagement@jhu.edu) to learn more about it.

6 Briefly address the following questions for each data product in Table 1. (You might refer to each by number).

7 NSF expects data sharing to follow the norms of your research community, but encourages efforts to broaden the range of data shared and of potential users beyond your field. Data can often be of unanticipated interest in the future if it can be located, understood, and cited.

8 “Accessible” generally means unmediated public access to your data distributed through a “cyber resource,” unless you specify conditions, such as embargo periods. “Sharing” can include direct release to interested parties upon request.

9 Specify a time period, e.g., “Data will be made available for sharing, in principle, two years after acquisition.”

10 This section will detail any reasons for sharing delays (e.g., embargo, publisher, patent, or political reasons) or restrictions (e.g. ecological endangerment concerns, IRB restrictions of sensitive data). You should also address granular methods for control and access (e.g., maintaining formal consent agreements, anonymous data, and restricted access to secured networks.)

11 State if there are IRB restrictions on data and steps to prepare accessible datasets, such as deidentifying transcripts. NSF requires fewer details than IRB forms, and respects when IRB restrictions put sharing beyond a reasonable effort, but they do sometimes ask for some attempt to create sharable datasets.

*Note: Your DMP for NSF grants should not exceed two pages. Contact Research Services in the Libraries for consultation ([ljohnsto@umn.edu](mailto:ljohnsto@umn.edu)).*

# Data Management Plan

V1 last updated **MM-DD-YYYY**

Name of student/researcher(s)	<b>Your Name</b>
Name of group/project	<b>Project Name or Research Lab (for group plan)</b>
Funding body(ies)	
Partner organisations	
Project Duration	Start: <b>MM-DD-YYYY</b> End: <b>MM-DD-YYYY</b>
Date Written	<b>MM-DD-YYYY</b>

## Table of Contents

[1. Introduction](#)

[2. Data Types](#)

*Section 2 Checklist*

[3. Data Organization, Documentation and Metadata](#)

*Section 3 Checklist*

[4. Data Access and Intellectual Property](#)

*Section 4 Checklist*

[5. Data Sharing and Reuse](#)

*Section 5 Checklist*

[6. Data Preservation and Archiving](#)

*Section 6 Checklist*

## 1. Introduction

The research project described in this data management plan (DMP) ....

## 2. Data Types

This types of data generated and/or used in this project include ...

*Note: Your DMP for NSF grants should not exceed two pages. Contact Research Services in the Libraries for consultation (ljohnsto@umn.edu).*

### *Section 2 Checklist*

- o What type of data will be produced?
- o How will data be collected? In what formats?
- o How to document data collection?
- o Will it be reproducible? What would happen if it got lost or became unusable later?
- o How much data will it be, and at what growth rate? How often will it change?
- o Are there tools or software needed to create/process/visualize the data?
- o Will you use pre-existing data? From where?
- o Storage and backup strategy?

## **3. Data Organization, Documentation and Metadata**

The plan for organizing, documenting, and using descriptive metadata to assure quality control and reproducibility of these data include ...

### *Section 3 Checklist*

- o What standards will be used for documentation and metadata?
- o Is there good project and data documentation format/standard?
- o What directory and file naming convention will be used?
- o What project and data identifiers will be assigned?
- o Is there a community standard for metadata sharing/integration?

## **4. Data Access and Intellectual Property**

The data have the following access and ownership concerns ...

### *Section 4 Checklist*

- o What steps will be taken to protect privacy, security, confidentiality, intellectual property or other rights?
- o Does your data have any access concerns? Describe the process someone would take to access your data.
- o Who controls it (e.g., PI, student, lab, University, funder) ?
- o Any special privacy or security requirements (e.g., personal data, high-security data) ?
- o Any embargo periods to uphold?



*Note: Your DMP for NSF grants should not exceed two pages. Contact Research Services in the Libraries for consultation ([ljohnsto@umn.edu](mailto:ljohnsto@umn.edu)).*

## 5. Data Sharing and Reuse

The data will be released for sharing in the following way ...

### *Section 5 Checklist*

- o If you allow others to reuse your data, how will the data be discovered and shared?
- o Any sharing requirements (e.g., funder data sharing policy) ?
- o Audience for reuse? Who will use it now? Who will use it later?
- o When will I publish it and where?
- o Tools/software needed to work with data?

## 6. Data Preservation and Archiving

The data will be preserved and archived in the following ways ...

### *Section 6 Checklist*

- o How will the data be archived for preservation and long-term access?
  - o How long should it be retained (e.g., 3-5 years, 10-20 years, permanently) ?
  - o What file formats? Are they long-lived?
  - o Are there data archives that my data is appropriate for (subject-based? Or institutional)?
  - o Who will maintain my data for the long-term?
-



## Research Data Toolkit

[Home](#)
[Agency Rqs/Guidance](#)
[Templates/Examples](#)
[Services](#)
[Training](#)
[Assess Data Needs](#)
[Tools](#)
[Contact Us](#)
[About](#)

### Templates

**Please Note:** These examples are not officially sanctioned by any UNC office. They are only intended to serve as examples for what you *might* do. Likewise, the sample plans linked below are very context-specific and are intended only to give a general idea of what others have done.

**DMPTool** - service of the University of California Curation Center (UC3) and the California Digital Library but customized for UNC at Chapel Hill. Select UNC from the list of institutions and login with your Onyen to see resources specific to our campus. Walks you through requirements for specific funding agencies. Allows you to work in sequence or jump around, save drafts, and export text files.

#### Sample Plans

Odum Institute's [sample plans](#)

ICPSR's [sample plan](#) (for deposit with ICPSR)

[Natural Science examples](#), from a wide range of projects and agencies (links collected on the ICPSR website)

#### Guides for Formulating Data Management Plans

[Guidelines for Effective Data Management Plans](#) (ICPSR)

[Managing and Sharing Data: Best Practice for Researchers](#) (UK Data Archive)

IRB application with sections that relate to data management

Applicable sections (on pages 9 and 10) include:

- A.4.11 Confidentiality of Data;
- A.4.12 Data sharing;
- A.4.13 Data security for storage and dissemination; and
- A.4.14 Post-study disposition of identifiable data or human biological materials

Odum Institute's [data management plan checklist](#)

[Data Management & Frequently Asked Questions \(FAQs\)](#) (NSF)

#### Other Resources

IRB-required consent form templates for research on human subjects (scroll down to "Consent")

### Example Language

For each of the five clauses presented in NSF's Grant Proposal Guide, Chapter II - Proposal Preparation Instructions, [Section J, Special Information and Supplementary Documentation](#) (second bulleted point), we have outlined below various points to consider in writing your plan. Where possible, we have adapted text from actual data management plans (although not necessarily plans for NSF) under the heading *Sample Text*. Such examples are not available for every section.

**Please Note:** These examples are not officially sanctioned by any UNC office at this time. They are only intended to serve as examples for what you *might* do.

If you are willing to share text from your own plan, please [contact us](#).

Here is a [Word version](#) in which to begin drafting your own plan.

From the NSF's Grant Proposal Guide: "Plans for data management and sharing of the products of research. Proposals must include a supplementary document of no more than two pages labeled "Data Management Plan." This supplement should describe how the proposal will conform to NSF policy on the dissemination and sharing of research results (see [AAG Chapter VI.D.4](#) ), and may include:

[Show All](#)

#### 1. Types of Data

[Data Description](#)  
[Existing Data](#)

#### 2. Standards

[Format](#)  
[Metadata](#)  
[Data Organization](#)  
[Quality Assurance](#)  
[Responsibility](#)

#### 3. Access and Sharing (Including Protected Data)

[Storage and Backup](#)  
[Data Access](#)  
[Ethics and Privacy](#)  
[Proprietary Data](#)  
[Intellectual Property](#)  
[Legal Requirements](#)

#### 4. Re-use

[Access and Sharing](#)  
[Re-Use](#)

#### 5. Archiving Data

[Archiving and Preservation](#)  
[Disaster Preparedness](#)  
[Budget](#)  
[Selection and Retention](#)

Last Updated: Apr 22, 2013 | URL: <http://guides.lib.unc.edu/researchdatatoolkit> | [Print Guide](#) | [SHARE](#) [f](#) [t](#) [e](#)

Powered by Springshare; All rights reserved. Report a tech support issue.  
View this page in a format suitable for printers and screen-readers or mobile devices.

[Admin Sign In](#)

## Job Descriptions

---

**Classified Title:** Data Management Specialist

**Working Title:** Data Management Consultant

**Role** : ATP

**Level** : 4

**Range:** PD

**Status:** Full Time

**Hours Worked:** 37.5

**Work Week:** Monday-Friday

**Contact:** Homewood HR 410-516-7196

**Personnel Area:** Libraries

**Org Unit:** Entrepreneurial Library Program

**Location:** 3400 N Charles Street

**Approximate  
Starting Salary:**

**General Description:** The Data Management Consultant provides consultative data management planning support to JHU Principal Investigators. The primary duties and responsibilities of the job include:

- Manage inquiries from Principal Investigators for data management planning support.
- Provide consultative support to PIs including
- Evaluate data planning needs, assess short and long term options and benefits, cater planning to specific granting agency requirements, and provide guidance on editing data management plans.
- Track specific scientific domain areas building knowledge and expertise in data types, formats, and needs within domain.
- Identify data standards, metadata standards, best practices for data management, etc. to continuously build expertise to support the JHU data archiving service.
- Maintain knowledge on a broad range of data repositories including their submission, Intellectual Property and use arrangements, and provide guidance on repository selection for deposit.
- Proactively collaborate and coordinate with team to implement data management plans for data being deposited into the JHU Data Archive.
- Collaborate with others in the library to effectively communicate services to faculty, researchers, and departments.
- Manage short and long-term communications and relationships with PIs, including outreach and training in data management best practices.
- Liaise with the Data Conservancy leadership.

**Additional information:**

The Sheridan Libraries and University Museums encompass the Milton S. Eisenhower

Library and its collections at the George Peabody Library, the Albert D. Hutzler Reading Room, the DC Centers, the Evergreen Museum and John Work Garrett Library, and Homewood Museum. Staff from the libraries and museums teach classes, curate exhibitions, produce scholarship and serve as principle investigators for research initiatives. Rich in resources and expertise, the libraries and museums focus on the needs of faculty and students but also serve as ambassadors to communities well beyond the borders of the Hopkins' campuses. A key partner in the academic enterprise, the library is a leader in the innovative application of information technology and has implemented notable diversity and organizational development programs. The Sheridan Libraries and University Museums are strongly committed to diversity. A strategic goal of the Libraries and Museums is to 'work toward achieving diversity when recruiting new and promoting existing staff.' The Libraries and Museums prize initiative, creativity, professionalism, and teamwork. For information on the Sheridan Libraries, visit [www.library.jhu.edu](http://www.library.jhu.edu). For information on Evergreen Museum and Library and Homewood Museum, visit [www.museums.jhu.edu](http://www.museums.jhu.edu).

***Qualifications:***

- Masters of Science, Engineering or Library Science.
- A minimum of three (3) years combined of library, information technology, informatics, and/or scientific research experience.
- Experience working with scientific data management and/or curation preferred.
- Experience with one or more components of the research data life cycle: creation, processing, analyzing, preserving, providing access to, and re-using.
- Must be self motivated, pro-active, willing to take on new challenges and solve problems with minimal supervision.
- Good listener with a high degree of customer orientation.
- Superb people skills, strong team-orientation, and professional attitude.
- Clear and consistent communicator.
- Strong writing skills.
- Strong project planning, management, and execution skills.
- Demonstrated ability to work with and easily adapt to new technology.

UNIVERSITY OF MARYLAND LIBRARIES  
POSITION DESCRIPTION FORM

Check one: Faculty  Exempt  Non-Exempt  Other

Date Prepared: January 19, 2012 Division: ITD

Prepared by: WITHHELD

Reports to: Manager, Digital Stewardship

Department: Digital Stewardship Unit, Information Technology Division

Position Title: Research Data Librarian [Post-Master's Program at the University of Maryland Libraries]

**NATURE OF WORK:**

The Post-Master's Program, a hiring initiative of the University of Maryland Libraries, matches recent post-master's professionals with short-term positions aligned with the Libraries' strategic priorities. Both sides win. The post-graduate professional develops their skills in a professional workplace, and the University Libraries gain the expertise of recent graduates to respond to a rapidly changing environment. Post-Master's Program professionals and the University Libraries each make a 2 year commitment to the position. Relocation costs are not available for Post-Master's Program professionals.

The University of Maryland Libraries at College Park is engaged in the exciting work of defining the future work of academic libraries. We are seeking employees who want to push the frontiers, to anticipate, model and lead in the provision of new services, and revise our definitions of collections, the library, and librarians themselves. Risk takers and highly flexible, creative problem solvers are most welcome!

The Research Data Librarian position provides an opportunity for a new librarian to get exposed to an academic library environment and exercise leadership in the development and implementation of policies and practices relating to e-Research, e-Science, and data management, a new area of engagement for the University of Maryland Libraries. The incumbent will help the University of Maryland Libraries define a completely new role for librarians - a role that will allow them to become more closely integrated in the whole educational and research process at the University of Maryland.

Reporting to the Manager, Digital Stewardship, the Research Data Librarian: actively participates in university-wide initiatives to develop and design policies, sustainable services, and infrastructure to enable faculty and students to preserve and make available their research data; partners with internal units (such as GIS and DRUM – Digital Repository of University of Maryland) and external units (such as Vice President for Research, Office of Information Technology (OIT) Enterprise Technical Infrastructure and Learning Technologies and Environments, and the Maryland Institute for Technology in the Humanities) to develop a data-publishing model that leverages library

services in support of data management and preservation; assists faculty with development of data management plans for grant applications; serves as an active member of the Information Technology Division, contributing to divisional initiatives and leading specific projects; incorporates support for data management and preservation into library services; and maintains close engagement with issues relating to scholarly communications such as copyright, open access, and data management and preservation.

**DUTIES AND RESPONSIBILITIES:**

- Develops an understanding of e-Research, e-Science and data services in selected fields; develops models for characterizing and interrelating datasets
- Performs research, evaluates approaches and implements best practices for gathering information on the developments in e-Research, data curation, metadata creation, and data preservation
- Performs and analyzes surveys to find out what practices and approaches researchers are using to collect, store, and re-use large data sets and how librarians can help them in this activity
- Provides support for researchers in implementing data storage and data management plans as required by funding agencies
- Investigates and implements new technologies and research tools that would support data services initiatives
- Develops and communicates a set of guidelines for best practices in data management for research
- Participates in preparation of grant proposals for development and advancement of the e-Research and data services program at the UMD Libraries
- Maintains a research guide and writes reports, articles and reviews related to data services; for example presents seminars/workshops data management and data curation
- Participates in library and campus committees as appropriate

**PHYSICAL DEMANDS:** Extensive use of the computer.

**SUPERVISORY RESPONSIBILITIES:** None

**EDUCATION:**

**Required:** Master's degree in Library or Information Science from an ALA-accredited institution of higher education by the start of employment.

**Preferred:** Advanced or undergraduate degree in science or engineering discipline.

**EXPERIENCE:**

**Required:**

Demonstrated knowledge of issues and technical challenges related to use and archiving of digital data. Experience with XML technologies and relational databases. Familiarity with academic, research, or special libraries. Excellent oral and written communication

skills. Excellent interpersonal skills with the ability to function independently and in groups, and to build and maintain relationships with partners.

**Preferred:**

Demonstrated subject knowledge and experience in sciences, social sciences, or engineering including understanding of issues related to scientific research and scholarly publishing. Familiarity/experience with data preservation, curation, management, content description and representation, metadata standards, and relevant workflows; experience with institutional or subject repository systems. Experience with DSpace, Fedora, or other repository software. Experience with HTML, CSS, JavaScript, PHP, Perl, or Java. Familiarity with linked open data.

**Employee's Signature** \_\_\_\_\_ **Date** \_\_\_\_\_

**Print Employee's Name** \_\_\_\_\_

**Supervisor's Signature** \_\_\_\_\_ **Date** \_\_\_\_\_



## **JOB DESCRIPTION**

---

<b>POSITION TITLE:</b>	Social Sciences Research Liaison Librarian
<b>INVENTORY NUMBER:</b>	201659
<b>LIBRARY:</b>	Satellite Social Sciences
<b>EFFECTIVE DATE:</b>	July 1, 2012
<b>POSITION REPORTS TO:</b>	Head, Teaching and Research Support – 201663

---

Titre français: Bibliothécaire de liaison en recherche pour les sciences sociales

### **SUMMARY OF FUNCTIONS:**

The incumbent is responsible for the delivery of an effective research liaison program to graduate students, post-doctoral fellows, faculty members and research teams in the social sciences domains.

The main functions of this position include the provision of in-depth reference services for individuals as well as research project liaison and support for research teams; the planning and delivery of customized instructional programs and workshops; and collection development in all formats including data resources in social sciences. Through collaborative outreach and liaison, the incumbent will gain an understanding of research teams' information resource and service needs that will be applied to developing, identifying and evaluating new services and information resources.

The main objective of this position is the enhancement of research output by creating efficiencies in the researcher information discovery process in support of the research mission of the University of Ottawa.

### **MAIN ACTIVITIES:**

#### *A. Outreach and liaison activities*

1. Provide reference assistance and in-depth research assistance to meet the information needs of researchers in the social sciences domains.
2. Maintain outreach to the social sciences community through engagement in departmental activities, awareness of current research and regular communication with faculty and students.
3. Liaise with academic units and researchers to promote library resources and services, reference and teaching activities and to identify ongoing needs.
4. Collaborate with faculty to create subject guides and use technologies such as social media to achieve seamless and integrated information and knowledge services for the assigned disciplines.
5. Develop and implement effective subject-based instructional and information literacy programs for assigned disciplines; collaborate with faculty in the design of innovative library and classroom instruction.
6. Working with the social sciences librarians team, incorporate support for e-science, research data management and curation into library services and assist researchers and faculty with development of data management plans.

May 2012

1

*B. Collections development*

1. Evaluate and develop scholarly information resources in assigned disciplines in accordance with current policies and practices and in cooperation with faculty and the social sciences librarians' team.
2. Elaborate, write and revise collections development policies for assigned disciplines and manage collections and gifts in kind in assigned disciplines including evaluation and transfer to the Library Annex;

*C. Other duties*

1. Contributes to librarianship by carrying out professional research and/or scholarly work.
2. Perform other duties as assigned by the Head, Teaching and Research Support Services.

**AUTONOMY:**

1. Carry out her functions under the responsibility of the Head, Teaching and Research Support Services.
2. Exercise full autonomy in the development of the collections in the assigned disciplines.

**RELATIONS:**

1. Frequent communication with the directors of academic units, library representatives, professors, students and other library clients.
2. Frequent communication with librarians and library technicians from other network libraries.
3. Occasional communication with the heads of specialized libraries and collections, and other library services.
4. Occasional relations with colleagues from external libraries.

**ESSENTIAL QUALIFICATIONS:**

1. A Master's degree in Library and Information Studies (M.L.I.S.) from an ALA accredited institution or equivalent;
2. Four (4) years of professional experience, or fewer, depending on relevance of experience to the position;
3. University degree in social sciences or experience working in a social sciences library;

4. Knowledge of scholarly information and research methods in social sciences acquired through studies or professional experience;
5. Familiarity with research data curation and metadata standards;
6. High level of technological literacy including knowledge of or experience with instructional technologies;
7. Pertinent knowledge of and experience in the areas of reference, teaching and collection development;
8. Excellent interpersonal and communication skills;
9. Bilingual (English and French), written and spoken, including the ability to teach in both official languages.

**INVENTORY NO:** 201XXX

**SIGNATURES:**

\_\_\_\_\_ Date:  
(Incumbent)

\_\_\_\_\_ Date:  
(Immediate Superior)

\_\_\_\_\_ Date:  
(Director)

\_\_\_\_\_ Date:  
(University Librarian)

Human Resources PD  
January 2012

**PURDUE UNIVERSITY-POSITION DESCRIPTION**

Date: 1/18/2012 Reason: [Create New Position](#)

**Libraries** **333** **STD: 50096371 & 50096372**  
**Org Unit Name** **Org Unit #** **Position ID #**

Supervisor Name: Michael Witt  
Supervisor Title: Interdisciplinary Research Librarian; Assistant Professor of Library Science  
Supervisor Position ID: 50031466 Phone: 4-8703 E-mail: mwitt@purdue.edu

**Position title: Digital Library Software Developer**  
*(Final determination rests with HRS)*

**Employee Group** *(Final determination rests with HRS)*

Non-exempt: **SELECT ONE** Exempt: **Administrative/Professional**

**Time Reporting:**  Full time  Part time (< 1.00) % = \_\_\_\_\_ **Shift: Day**

**Employee Subgroup:** **Non-exempt position SELECT ONE Exempt position FY 12**

**Education:** Indicate the **minimum** education required. **BA/BS degree**

**List the required and/or preferred course work or degree field(s):**

Bachelor's degree in computer technology, computer science, library and information science, or a related field. Master's of Library Science (MLS or MLIS) preferred.

**Experience:** Indicate the **minimum** years of experience required. **1 yr**

**Describe the type of experience required and/or preferred:**

Experience gathering requirements, evaluating tools, and designing, developing, and implementing software. Complementary experience in which documentation, analysis, problem-solving, and communication have been demonstrated both independently and as a member of a team. Mastery of two or more current programming languages. Preferred: experience with PHP, java, MySQL, XML, or RDF; Experience working in a library or with systems that manage digital information in a library context.

**Knowledge, Skills, Abilities:** List any knowledge, skills, or abilities, special training, certificates or licenses.

Functional competencies include: knowledge of databases, computer programming and scripting languages, software and web development, and information systems. Ability to learn and evaluate current and relevant technologies, standards, and practices in digital preservation such as OAIS, PREMIS, TRAC, LOCKSS, DuraSpace, Archivematica, and curation micro services

Leadership competencies include: teamwork, initiative, and innovation; skill to analyze data and use sound judgment to make decisions; skill to develop specific goals and plans to prioritize, organize, and accomplish work; skill to communicate effectively and clearly in both written and oral forms;

Interpersonal competencies include: effective teamwork and collaboration; efficient management of time, utilizing excellent time and project management; analysis of data and use of sound judgment to make decisions; completion of work with a high level of accuracy and attention to detail; effective and clear communication; development of specific goals and plans to prioritize, organize, and accomplish work; ability to build productive and respectful relationships with others and maintain them over time; ability to pay attention to detail and concentrate on a task over a period of time without being distracted;

Customer Service competencies include: ability to build and respond courteously and effectively to customer needs; ability to address shortfalls and problems in service delivery; ability to share solutions and improvements with others.

Change Management competencies include: ability to adjust productively to and communicate about change; ability to explore and try new ideas, methods, and approaches; ability to suggest changes that fit with unit/Libraries strategic plans.

Performance Development competencies include: ability to take initiative to learn new knowledge and skills; receptive to feedback and takes appropriate action in response; productively applies new knowledge and skills.

Teamwork and Collaboration competencies include: skill to communicate productively about group decisions; ability to treat coworkers with respect; ability to contribute productively to group/unit results.

<b>AAP 710</b>	<b>Census 101</b>	<b>EEO 08</b>	<b>JIC</b>	<b>Supervision No</b>	<b>FLSA Exemption</b>	<b>Exempt-Administrative</b>
<b>Donna Dye</b>		<b>2/17/2012</b>		<b>Job Long Text: App. Developer II</b>		
<b>Compensation Specialist</b>		<b>Validity Date</b>		<b>Job ID: 1260</b>		

Human Resources PD  
January 2012

**POSITION SUMMARY: What is the main purpose of this position? Why does it exist?**

The Digital Library Software Developer will be responsible for implementing and developing software to build out a long-term preservation environment for research data allowing the Purdue University Research Repository (PURR) to sustain published materials. This position will also collaborate with Purdue colleagues, and with both national and international partners, to develop and implement software in support of policies and practices that enable long-term digital data management and preservation. This is an 18 month position with the possibility of continued funding.

**ESSENTIAL DUTIES AND RESPONSIBILITIES** include but are not limited to the following:

*Describe the essential responsibilities of the position in order of importance. Essential responsibilities are those functions, if removed, would fundamentally alter the purpose of the position. It is not necessary to list each individual task. Percentages should be listed in no less than 5% increments and must total 100%.*

USE THE TAB KEY TO MOVE FROM FIELD TO FIELD

Essential	Percent
DEVELOP AND IMPLEMENT DIGITAL LIBRARY SOFTWARE	90%
<ul style="list-style-type: none"> <li>• Learn and evaluate current and relevant technologies, standards, and practices in digital preservation such as OAIS, PREMIS, TRAC, LOCKSS, Dura Space, Archivematica, and curation micro services (focus: preservation) AND/OR Learn and evaluate current and relevant technologies, standards, and practices in systems that are used to manage digital information in a library context, in particular those that relate metadata management, persistent identifiers, data interoperability, and discovery tools (focus: systems).</li> <li>• Gather requirements, consult and collaborate with constituents including project team members, project partners—the Libraries, Information Technology at Purdue (ITaP), and the Office of the Vice President for Research (OVPR), archivists, and users to identify needs and design software solutions to support user and archival workflows, policies, and best practices for digital preservation</li> <li>• Design, develop, and implement a standards-based preservation environment for digital research data as a component of the Purdue University Research Repository (PURR)</li> <li>• Provide documentation, support, and continuous improvement of preservation software and systems</li> <li>• Contribute code to the HUB zero and other, related open source projects</li> <li>• Assist in the development and certification of PURR as a Trustworthy Digital Repository (ISO 16363)</li> </ul>	
ADMINISTRATIVE	10%
<ul style="list-style-type: none"> <li>• Regularly meet, communicate, and collaborate with project partners and library units</li> <li>• Prepare reports or correspondence concerning project specifications, activities, or status</li> <li>• Other duties and projects as assigned</li> </ul>	

Human Resources PD  
January 2012

**SUPERVISION ROSTER**

**Libraries**  
**Org Unit Name**

**333**  
**Org Unit #**

**STD: 50096371 & 50096372**  
**Position ID #**

**Supervision exercised:** Must be listed as an essential responsibility of the position and described along with the percentage of time under the "Responsibilities Section" on the previous page.

**Functional:** limited to assigning, instructing and reviewing work of others. Also includes hiring, terminating and pay decisions for both undergraduate and graduate student employees.

- Indicate the total number of **Monthly exempt** staff this position functionally supervises: \_\_\_\_\_
- Indicate the total number of **Bi-weekly non-exempt** staff this position functionally supervises: \_\_\_\_\_

**Temporary/Student(s) supervision:** *List the total number of positions supervised below.*

- **Monthly temporary/Grad student(s):** \_\_\_\_\_ **Hourly temporary/Student(s):** 1

**Administrative:** responsible for making decisions/recommendations for hiring, terminations, pay adjustments, promotions and training of direct reports as well as performing other supervisory duties. **(If revising existing position, only list changes to reporting below.)**

- Indicate the total number of **Monthly exempt** staff this position administratively supervises: \_\_\_\_\_
- Indicate the total number of **Bi-weekly non-exempt** staff this position administratively supervises: \_\_\_\_\_

**List IDs of the position(s) below:**

**(Required)** List the **IDs of the Position(s)** (not the person) for each direct report this position **administratively supervises**. Must match the total number listed above. *Do not include graduate student, temporary, or grouped positions.*

**Monthly exempt:**


**(Required)** List the **IDs of the Position(s)** (not the person) for each direct report this position **administratively supervises**. Must match the total number listed above. *Do not include undergraduate student, temporary, or grouped positions.*

**Bi-weekly non-exempt:**


Human Resources PD  
January 2012

**PHYSICAL, ENVIRONMENTAL, AND HAZARDOUS SPECIFICATIONS**

Identify below the physical, environmental, and hazardous conditions under which the essential responsibilities of the position are performed.

**Physical Requirements**

From the list of physical requirement descriptions below, check the box that best describes the physical requirements of the position.

- |  |   |
|--|---|
| 1. <input type="checkbox"/> SEDENTARY ACTIVITY: Lift and carry up to 10 lbs. occasionally; sedentary work involves sitting most of the time. | 4. <input type="checkbox"/> MODERATE PHYSICAL ACTIVITY: lift and carry 25 to 50 lbs. frequently, and up to 60 lbs occasionally. |
| 2. <input checked="" type="checkbox"/> LIMITED PHYSICAL ACTIVITY: Lift and carry up to 10 lbs. frequently, and up to 20 lbs. occasionally.   | 5. <input type="checkbox"/> HEAVY PHYSICAL ACTIVITY: lift and carry 50 to 80 lbs. frequently, and up to 100+ lbs. occasionally. |
| 3. <input type="checkbox"/> LIGHT PHYSICAL ACTIVITY: Lift and carry 10 to 25 lbs. frequently, and up to 40 lbs occasionally.                 |   |

\*Occasional is defined as <50 percent of the time.  
\*\*Frequent is defined as >50 percent of the time.

**Machines, Tools, Electronic Devices & Office Equipment**

List the machines, tools, electronic devices, office equipment or other equipment necessary to perform the job.

- |                     |                   |                      |
|---------------------|-------------------|----------------------|
| 1. <u>Computers</u> | 2. <u>Servers</u> | 3. <u>Copier/Fax</u> |
| 4. <u>Printer</u>   | 5. _____          | 6. _____             |

**Environmental and Hazardous Conditions**

Check the boxes that best describe the environmental and hazardous conditions of the job.

- |  |   |
|--|---|
| 1. Work indoors (% of time: <b>100</b> )                                       | Work outdoors (% of time: _____)  |
| 2. Respiratory Conditions: Involving exposure to:                              | <input type="checkbox"/> Fumes/vapors <input type="checkbox"/> Dust <input type="checkbox"/> Odors        |
| <input type="checkbox"/> Gases <input type="checkbox"/> Inadequate ventilation | <input type="checkbox"/> Other conditions (list) _____  |
| 3. Skin Conditions: Involving exposure to:                                     | <input type="checkbox"/> Toxic chemicals <input type="checkbox"/> Radiation <input type="checkbox"/> Burn |
| <input type="checkbox"/> Electrical shock                                      | <input type="checkbox"/> Other conditions (list) _____  |
| 4. Working Conditions: Including use of, or exposure to:                       |   |
| <input type="checkbox"/> Heavy machinery                                       | <input type="checkbox"/> Machinery with moving parts <input type="checkbox"/> Vibration                   |
| <input type="checkbox"/> Working on scaffolding and high places                | <input type="checkbox"/> High voltage electricity <input type="checkbox"/> Lasers                         |
| <input type="checkbox"/> Steam pipes and/or tunnels                            | <input type="checkbox"/> Grease and oils <input type="checkbox"/> Cramped working quarters                |
| <input type="checkbox"/> Biological and/or chemical reagents                   | <input type="checkbox"/> Infectious diseases <input type="checkbox"/> Use of sharp objects                |
| <input type="checkbox"/> Extreme cold (temperatures below 32°)                 | <input type="checkbox"/> Noise (work requires employee to shout to be heard)                              |
| <input type="checkbox"/> Extreme heat (temperatures above 90°)                 | <input type="checkbox"/> Handling or maintaining animals  |
| <input type="checkbox"/> Other conditions (list) _____                         |   |

**DEPARTMENTAL/SCHOOL APPROVALS**

**Approval to Establish/Modify Position:** As supervisor of this position, I am certifying that this description is an accurate reflection of the primary purpose of the position and that the essential duties and responsibilities listed are those that the employee in this position is expected to perform. It does not limit or modify my responsibility or authority to assign and direct the work of the employee.

Supervisor Signature – **REQUIRED** \_\_\_\_\_ Date \_\_\_\_\_

Department Head Signature – **REQUIRED** \_\_\_\_\_ Date \_\_\_\_\_

Fiscal Authorization Signature – **REQUIRED** \_\_\_\_\_ Date \_\_\_\_\_  
(e.g., Business Office/Director/VP)

Human Resources PD  
January 2012

**PURDUE UNIVERSITY-POSITION DESCRIPTION**

Date: 2/3/2012 Reason: [Create New Position](#)

<b>Libraries</b>	<b>333</b>	<b>50096367</b>
<b>Org Unit Name</b>	<b>Org Unit #</b>	<b>Position ID #</b>

Supervisor Name: Michael Witt

Supervisor Title: Interdisciplinary Research Librarian; Assistant Professor of Library Science

Supervisor Position ID: 50031466

Phone: 4-8703

E-mail: mwitt@purdue.edu

**Position title: Digital Data Repository Specialist**

*(Final determination rests with HRS)*

**Employee Group** *(Final determination rests with HRS)*

Non-exempt: [SELECT ONE](#)

Exempt: [Administrative/Professional](#)

**Time Reporting:**  Full time  Part time (< 1.00) % = \_\_\_\_\_ **Shift:** [Day](#)

**Employee Subgroup:** [Non-exempt position SELECT ONE](#) [Exempt position FY 12](#)

**Education:** Indicate the **minimum** education required. [MS degree](#)

**List the required and/or preferred course work or degree field(s):**

MLS or MIS from an ALA-accredited institution or equivalent combination of education and experience.

**Experience:** Indicate the **minimum** years of experience required. [1 yr](#)

**Describe the type of experience required and/or preferred:**

Experience managing and/or developing repositories and digital collections. Experience in supporting and participating in scholarly communications and sponsored research. Experience as a successful collaborator in a collegial research library environment.

**Knowledge, Skills, Abilities:** List any knowledge, skills, or abilities, special training, certificates or licenses.

Functional competencies include: management or development of digital repositories, digital collections, and/or content management systems; one or more major descriptive metadata standards; standards and practices related to digital preservation such as ISO 16363 or TRAC; current digital preservation environment and practices and the research process, data life cycle, and trends in the organization and management of digital information; scholarly communication and intellectual property issues.

Leadership competencies include: teamwork, initiative, and innovation; ability to perform outreach and promotion for data services; skill to analyze data and use sound judgment to make decisions; skill to develop specific goals and plans to prioritize, organize, and accomplish work; skill to communicate effectively and clearly in both written and oral forms.

Interpersonal competencies include: effective teamwork and collaboration; efficient management of time, utilizing excellent time and project management; analysis of data and use of sound judgment to make decisions; completion of work with a high level of accuracy and attention to detail; effective and clear communication; development of specific goals and plans to prioritize, organize, and accomplish work; ability to build productive and respectful relationships with others and maintain them over time; ability to pay attention to detail and concentrate on a task over a period of time without being distracted; ability to teach something to others.

Customer Service competencies include: ability to build and respond courteously and effectively to customer needs; ability to address shortfalls and problems in service delivery; ability to share solutions and improvements with others.

Change Management competencies include: ability to adjust productively to and communicate about change; ability to explore and try new ideas, methods, and approaches; ability to suggest changes that fit with unit/Libraries strategic plans.

Performance Development competencies include: ability to take initiative to learn new knowledge and skills; receptive to feedback and takes appropriate action in response; productively applies new knowledge and skills.

Teamwork and Collaboration competencies include: skill to communicate productively about group decisions; ability to treat coworkers with respect; ability to contribute productively to group/unit results.

*For HRS use only:*

**AAP 710** **Census 101** **EEO 08** **JIC 26652** **Supervision No** **FLSA Exemption** **Exempt-Administrative**

**Donna Dye**

**2/17/2012**

**Job Long Text: Software Quality Splst. II**

**Compensation Specialist**

**Validity Date**

**Job ID: 1297**



Human Resources PD  
January 2012

**POSITION SUMMARY: What is the main purpose of this position? Why does it exist?**  
The Digital Data Repository Specialist will oversee and provide support for the launch and subsequent day-to-day operation of the Purdue University Research Repository (PURR) service. The position will partner with colleagues to support the adoption and improvement of PURR, as well as lead the ISO 16363 certification process for PURR as a Trustworthy Digital Repository. This is an 18 month position with the possibility of continued funding.

**ESSENTIAL DUTIES AND RESPONSIBILITIES** include but are not limited to the following:

*Describe the essential responsibilities of the position in order of importance. Essential responsibilities are those functions, if removed, would fundamentally alter the purpose of the position. It is not necessary to list each individual task. Percentages should be listed in no less than 5% increments and must total 100%.*

<b>Essential</b>	<b>Percent</b>
<p>REPOSITORY SERVICE MANAGEMENT</p> <ul style="list-style-type: none"> <li>• Lead and complete a successful ISO 16363 audit to establish and maintain PURR as a Trustworthy Digital Repository</li> <li>• Review, update and develop PURR policies (e.g., preservation) and procedures as the repository grows, and as technology and community practice evolves</li> <li>• Maintain a documented history of changes to repository's operations, procedures, software and hardware, and keeping records of actions and administrative processes relevant to storage and preservation.</li> <li>• Document change management and critical processes that potentially affect the repository's ability to comply with its mandatory responsibilities.</li> <li>• Review logs of access management failures and anomalies and respond accordingly.</li> <li>• Develop, maintain and revise the business plan to reflect both short- and long-term goals.</li> <li>• Analyze repository for security risk factors associated with personnel and physical plant.</li> <li>• Report on financial risk, benefit, investment, and expenditure (including assets, licenses, and liabilities).</li> <li>• Coordinate regularly scheduled self-assessment and external certification processes.</li> <li>• Coordinate staff roles, responsibilities, and authorizations related to implementing changes within the system and service.</li> <li>• Ensure that PURR meets its defined service level and policies.</li> <li>• Actively monitor the integrity of all digital archival objects, managing the number and location of copies of all digital objects, and maintaining information integrity measurements.</li> <li>• Coordinate and test the understandability of the Content Information and respond to the appropriate Designated Communities when Representation Information is inadequate for understanding the data holdings.</li> <li>• Review all reported incidents of data corruption and loss, and assess necessary revisions to software/hardware systems, operational procedures and management policies as needed.</li> <li>• Change preservation plans accordingly as a result of repository monitoring. Update and maintain Designated Community definitions and their accessibility, the delivery and access options available to the Designated Community, and address feedback from users.</li> </ul>	70%
<p>OUTREACH AND SUPPORT</p> <ul style="list-style-type: none"> <li>• Provide support to users in the context of their use of PURR (e.g., collaborators on a research project, dataset production and publication and archiving, end-users of datasets).</li> <li>• Work with subject specialist librarians to engage researchers on projects and review submitted datasets as needed.</li> <li>• Perform outreach and promotion for data services</li> <li>• Train staff and ensure PURR has adequate staff and skills to fulfill its duties and responsibilities.</li> </ul>	25%
<p>ADMINISTRATIVE</p> <ul style="list-style-type: none"> <li>• Partner with colleagues in the Office of the Vice President for Research, Information Technology at Purdue (ITaP), and the Purdue Libraries as a member of the PURR project team in the continuous improvement of PURR</li> <li>• Regularly meet, communicate, and collaborate with project partners, the project team, and library units</li> <li>• Other duties and projects as assigned</li> </ul>	5%



Human Resources PD  
January 2012

**PHYSICAL, ENVIRONMENTAL, AND HAZARDOUS SPECIFICATIONS**

Identify below the physical, environmental, and hazardous conditions under which the essential responsibilities of the position are performed.

**Physical Requirements**

From the list of physical requirement descriptions below, check the box that best describes the physical requirements of the position.

- |  |   |
|--|---|
| 1. <input type="checkbox"/> SEDENTARY ACTIVITY: Lift and carry up to 10 lbs. occasionally; sedentary work involves sitting most of the time. | 4. <input type="checkbox"/> MODERATE PHYSICAL ACTIVITY: lift and carry 25 to 50 lbs. frequently, and up to 60 lbs occasionally. |
| 2. <input checked="" type="checkbox"/> LIMITED PHYSICAL ACTIVITY: Lift and carry up to 10 lbs. frequently, and up to 20 lbs. occasionally.   | 5. <input type="checkbox"/> HEAVY PHYSICAL ACTIVITY: lift and carry 50 to 80 lbs. frequently, and up to 100+ lbs. occasionally. |
| 3. <input type="checkbox"/> LIGHT PHYSICAL ACTIVITY: Lift and carry 10 to 25 lbs. frequently, and up to 40 lbs occasionally.                 |   |
- \*Occasional is defined as <50 percent of the time.  
\*\*Frequent is defined as >50 percent of the time.

**Machines, Tools, Electronic Devices & Office Equipment**

List the machines, tools, electronic devices, office equipment or other equipment necessary to perform the job.

- |                     |                   |                      |
|---------------------|-------------------|----------------------|
| 1. <u>Computers</u> | 2. <u>Servers</u> | 3. <u>Copier/Fax</u> |
| 4. <u>Printer</u>   | 5. _____          | 6. _____             |

**Environmental and Hazardous Conditions**

Check the boxes that best describe the environmental and hazardous conditions of the job.

- |  |  |  |
|--|--|--|
| 1. Work indoors (% of time: <u>100</u> )                                       | Work outdoors (% of time: _____)   |  |
| 2. Respiratory Conditions: Involving exposure to:                              | <input type="checkbox"/> Fumes/vapors  | <input type="checkbox"/> Dust <input type="checkbox"/> Odors     |
| <input type="checkbox"/> Gases <input type="checkbox"/> Inadequate ventilation | <input type="checkbox"/> Other conditions (list) _____                       |  |
| 3. Skin Conditions: Involving exposure to:                                     | <input type="checkbox"/> Toxic chemicals                                     | <input type="checkbox"/> Radiation <input type="checkbox"/> Burn |
| <input type="checkbox"/> Electrical shock                                      | <input type="checkbox"/> Other conditions (list) _____                       |  |
| 4. Working Conditions: Including use of, or exposure to:                       | <input type="checkbox"/> Machinery with moving parts                         | <input type="checkbox"/> Vibration                               |
| <input type="checkbox"/> Heavy machinery                                       | <input type="checkbox"/> High voltage electricity                            | <input type="checkbox"/> Lasers                                  |
| <input type="checkbox"/> Working on scaffolding and high places                | <input type="checkbox"/> Grease and oils                                     | <input type="checkbox"/> Cramped working quarters                |
| <input type="checkbox"/> Steam pipes and/or tunnels                            | <input type="checkbox"/> Infectious diseases                                 | <input type="checkbox"/> Use of sharp objects                    |
| <input type="checkbox"/> Biological and/or chemical reagents                   | <input type="checkbox"/> Noise (work requires employee to shout to be heard) |  |
| <input type="checkbox"/> Extreme cold (temperatures below 32°)                 | <input type="checkbox"/> Handling or maintaining animals                     |  |
| <input type="checkbox"/> Extreme heat (temperatures above 90°)                 |  |  |
| <input type="checkbox"/> Other conditions (list) _____                         |  |  |

**DEPARTMENTAL/SCHOOL APPROVALS**

**Approval to Establish/Modify Position:** As supervisor of this position, I am certifying that this description is an accurate reflection of the primary purpose of the position and that the essential duties and responsibilities listed are those that the employee in this position is expected to perform. It does not limit or modify my responsibility or authority to assign and direct the work of the employee.

Supervisor Signature – **REQUIRED** \_\_\_\_\_ Date \_\_\_\_\_

Department Head Signature – **REQUIRED** \_\_\_\_\_ Date \_\_\_\_\_

Fiscal Authorization Signature – **REQUIRED** \_\_\_\_\_ Date \_\_\_\_\_  
(e.g., Business Office/Director/VP)



THE UNIVERSITY of  
**TENNESSEE** **UT**  
KNOXVILLE



Libraries

University Libraries Home
MyUTK | Online@UT | Tmail | Volmail | A-Z Index

Find Resources

---

Databases

---

Course Reserves

---

Request Materials

---

Services

---

Hours

---

Branch Libraries & Collections

---

About the Libraries

---

Libraries A to Z

---

My Account

---

## University Libraries Faculty Vacancy

[Libraries Employment](#)

<b>Position:</b>	<b>Data Curation Librarian</b>
<b>Appointment Rank:</b>	Assistant Professor
<b>Salary:</b>	\$48,000.00
<b>Available:</b>	June 1, 2013

The Data Curation Librarian will build on e-Science training initiatives and support new emphases and directions in liaison librarian assignments at the University of Tennessee, Knoxville. This new position will lead new initiatives in data curation and work collaboratively on new research initiatives and campus technology innovation.

Reporting to the Associate Dean for Research and Scholarly Communication, the data curation librarian will: strengthen the University's capacity to secure highly competitive grant funding; contribute to the development of long-term data management infrastructure; assist faculty in the discovery of relevant existing data sets and other information; serve as a PI, co-PI or grant team member on externally funded projects; and engage in research and professional activity at the national and international level. The librarian performs data management planning with PIs and researchers, serves as a consultant with researchers on research data issues, and trains researchers on the use of digital research and publishing tools, including UT's Trace digital repository.

The successful candidate will perform outreach and facilitate communication between the Libraries and research groups at UT. The librarian is a member of the Research and Scholarly Services department and a Learning, Research, and Collections liaison. As such, the incumbent is responsible for learning and engagement, research and scholarly communication, and stewardship and collections activities in assigned liaison areas. The Data Curation Librarian is responsible for building strong relationships with administrators, faculty, students, and staff on campus, within the Libraries, and beyond the university. Depending on qualifications and experience, the incumbent may be responsible for supervising library faculty and/or staff.

**Responsibilities:**

- Assist faculty with development of data management plans for grant applications and general data stewardship
- Working closely with other liaison librarians, incorporate support for data management, citation, and preservation into library services
- Maintain an awareness of emerging trends and best practices in e-science, data curation, and e-scholarship in all disciplines.
- Develop services to enhance access to data.
- Maintain awareness of subject or disciplinary repositories of potential interest to the UT research community
- Maintain awareness of tools and algorithms for computationally centered, data-driven science (data mining, visualization, text mining, etc.)
- Actively participate in university-wide initiatives to develop and design policies, services, and infrastructure to enable faculty and students to preserve and make available their research data
- Partner with internal units (such as Digital Initiatives, Learning and Outreach, and Agriculture & Veterinary Medicine Library) and external units (such as Office of the Vice Chancellor for Research and Engagement, Office of Information Technology, and Center for Information & Communication Studies) to implement data management and publishing services and workshops

**Required Qualifications:**

- ALA-accredited Master's degree in Library and/or Information Science, or doctorate in a relevant field.
- Demonstrated knowledge of issues and technical challenges related to the life cycle of research data
- Familiarity with two or more commonly used repository platforms (Fedora, DSpace, Dataverse, iRODS, etc.)
- Strong commitment to public service and ability to work well with diverse population of faculty, students, and academic colleagues
- Strong communication (oral and writing), interpersonal, and presentation skills



- Help from a Librarian -

Walk-Ins

Chat

E-Mail

Phone

Appointments

152 · Representative Documents: Job Descriptions

- Ability to initiate and manage collaborative projects and develop policies
- Ability to think creatively in developing and promoting the use of library services and collections through a variety of outreach efforts
- Familiarity with funding agency requirements for data management plans
- Familiarity with ISO 14721
- Must be able to meet the requirements of a tenure-track librarian position

**Preferred Qualifications:**

- Experience working with research data and researchers (e.g. a combination of academic work done in labs with research data, outreach work done with researchers and faculty, digital repository work, etc.)
- Second advanced degree in STEM (science, technology, engineering, mathematics) field or quantitative social science discipline
- Experience with DSpace, Fedora, Dataverse, or iRODS
- Experience with one or more of the following web technologies: HTML, CSS, JavaScript, PHP, Perl, Python, Java
- Experience with XML, XSLT, and relational databases
- Instruction or teaching experience
- Familiarity with at least one of the following metadata standards: Ecological Metadata Language (EML), Data Documentation Initiative (DDI), FGDC/ISO 19115, METS, PREMIS
- Ability to use various tools for metadata manipulation and scripting
- Successful track record of collaboration with other campus units around scholarly issues and/or technologies
- Experience working on an externally funded project
- Responsible conduct of research/research ethics training or certification
- Experience with a statistical software package (e.g., SPSS, SAS, R)
- Supervisory experience

**Environment:**

The University of Tennessee Libraries serves the flagship campus of the state university system. The UT Libraries supports the teaching, research, and service mission of the university and enhances the academic experience of each student at the Knoxville campus — through outstanding print and electronic collections, reference and instructional services, and top-notch facilities and technological resources.

The UT Libraries serves as an intellectual, cultural, and social center for the university and community. We are a national leader in the creation of regionally significant digital collections; in support of open access through our digital repository Trace; and through a rich history of designing innovative spaces and building key partnerships that enhance the teaching/learning enterprise. The University of Tennessee Libraries is a member of the Association of Research Libraries, the Association of Southeastern Research Libraries, the Digital Library Federation, Lyrisis, and the Center for Research Libraries. The UT Libraries collaborates actively at the state level with the other UT System Libraries as well as the libraries in the Tennessee Board of Regents system.

**Benefits:**

Excellent benefits include 24 annual leave days; choice of state retirement plan or ORP (AIG Retirement, ING, TIAA-CREF) with nonrefundable contributions paid for the employee by the University; optional group health and life insurance plans. Tuition remission is available for all university employees and partial undergraduate tuition remission is available to dependent children and spouses of UT employees. Faculty rank and status; twelve-month, tenure-track appointment.

**Application Procedures:**

A background check and official transcripts are required prior to hiring. Send cover letter addressing the above qualifications, a current resume, and the names, addresses, e-mail addresses, and telephone numbers of three recent references to: Elizabeth Greene, Library Human Resources, 1015 Volunteer Blvd., Knoxville, TN 37996-1000. Application materials may be sent via email attachment to [ejgreene@utk.edu](mailto:ejgreene@utk.edu). Review of applications will begin April 15, 2013 and will continue until the position is filled. Qualified spring graduates are encouraged to apply.

---

*All qualified applicants will receive equal consideration for employment and admissions without regard to race, color, national origin, religion, sex, pregnancy, marital status, sexual orientation, gender identity, age, physical or mental disability, or covered veteran status.*

*Eligibility and other terms and conditions of employment benefits at The University of Tennessee are governed by laws and regulations of the State of Tennessee, and this non-discrimination statement is intended to be consistent with those laws and regulations.*

*In accordance with the requirements of Title VI of the Civil Rights Act of 1964, Title IX of the Education Amendments of 1972, Section 504 of the Rehabilitation Act of 1973, and the Americans with Disabilities Act*

#### Data Consultant Description

##### Description:

Over the past 2 1/2 years, the University Library's Scientific Data Consulting Group has focused on assessment of the data management landscape at UVA, developing and testing service workflows and processes, building team capacity, and developing a strategy for addressing long-term research data management challenges at UVA. We are now positioned to offer significant value to the research process through a more complete set of lifecycle services and to become long-term data management consulting partners with researchers. We are now looking for an experienced researcher/scientist to join our group in a part-time capacity, for a one year period, to help advance specific objectives.

In addition to participation in day-to-day consulting and training activities, this individual will contribute in the following ways:

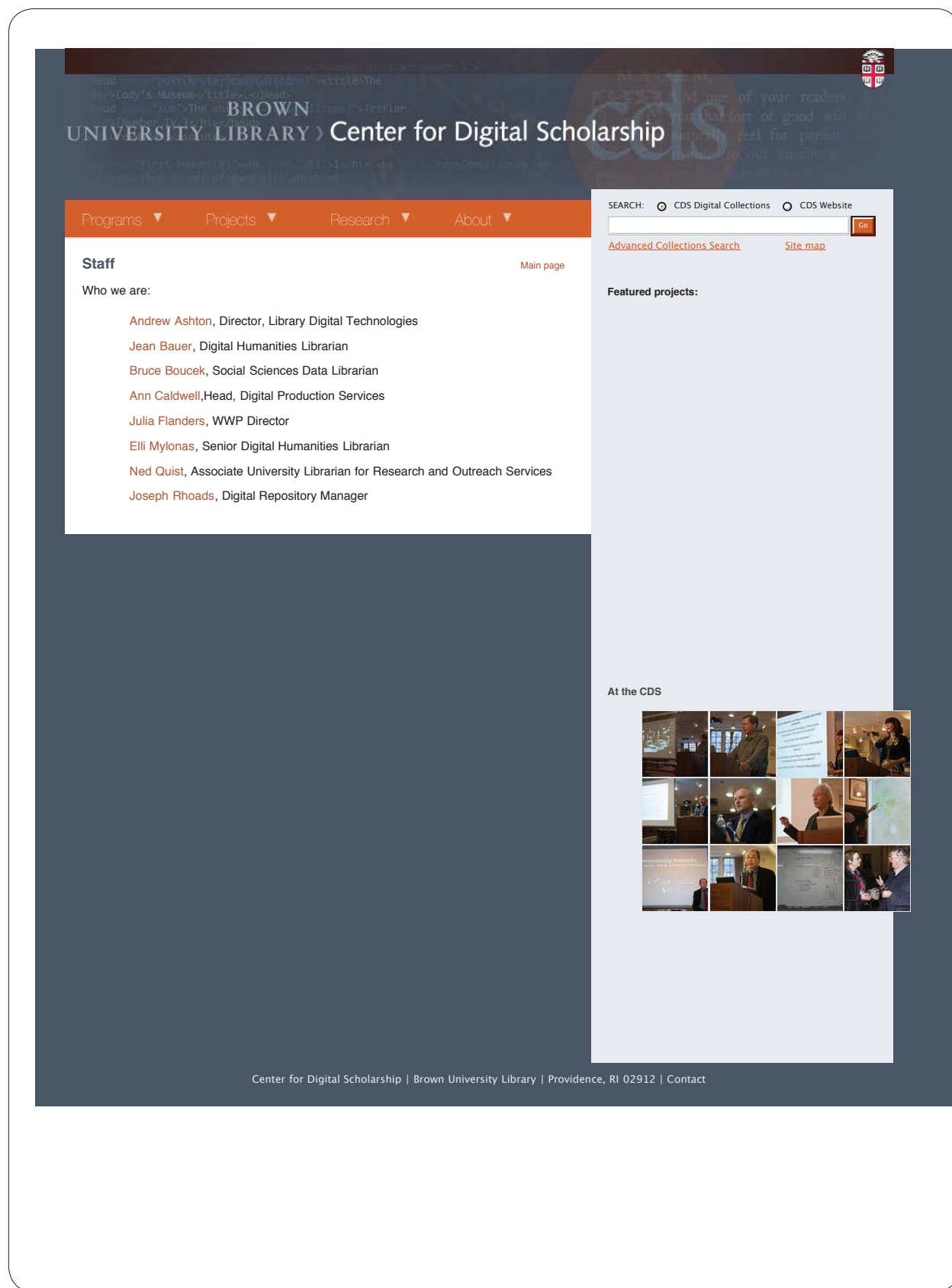
- 40% - Add capacity for implementation services – Apply knowledge of the University, infrastructure, services, and belief in our approach, to add team capacity for consulting and training services, especially in consulting with researchers on implementation of our data management process improvement recommendations.
- 40% - Evaluation of value/impact – Within context of responsibility centered management (RCM), look at how SciDaC generates value for researchers and the institution from the funding allocated to the unit, and how this ecosystem can be measured and reported in a clear manner to various stakeholders (primary focus is on researchers).
- 10% - Analyze from a different perspective – Conduct ongoing testing of consulting and training processes with target user groups.
- 10% - Reach deeper- Help SciDaC to better frame services in language which resonates with researchers, and to market more actively to those audiences.

##### Required experience/skills:

- Knowledge of research data policy, sponsored research trends, research software, open access trends, data sharing, security, preservation, IT environments, and academic research processes generally.
- A demonstrated commitment to improvement of research data management practices in the academic environment.
- Deep knowledge of how research is conducted, the incentive and motivation models involved in academic research, and how to marry the ideal vision for management and preservation of research data with the realities of day-to-day processes and cultural beliefs.
- A PhD in an academic discipline, as well as demonstrated experience applying for, obtaining, and completing sponsored research grants.
- More than 10 years experience as an academic researcher.

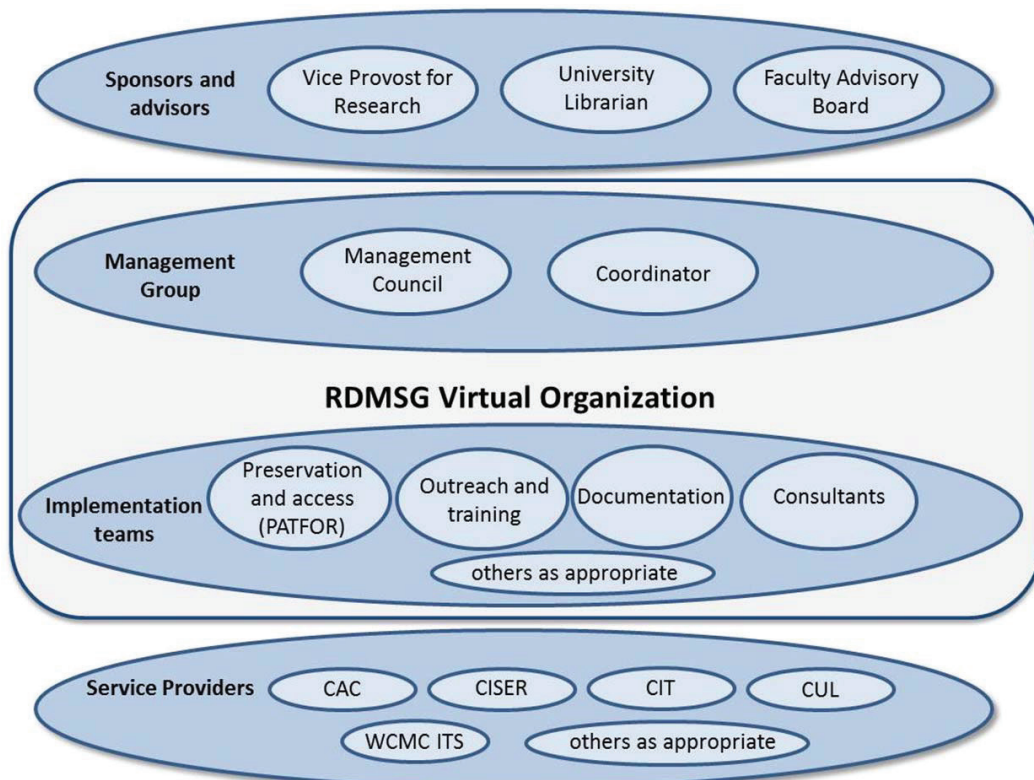
## Organization Charts

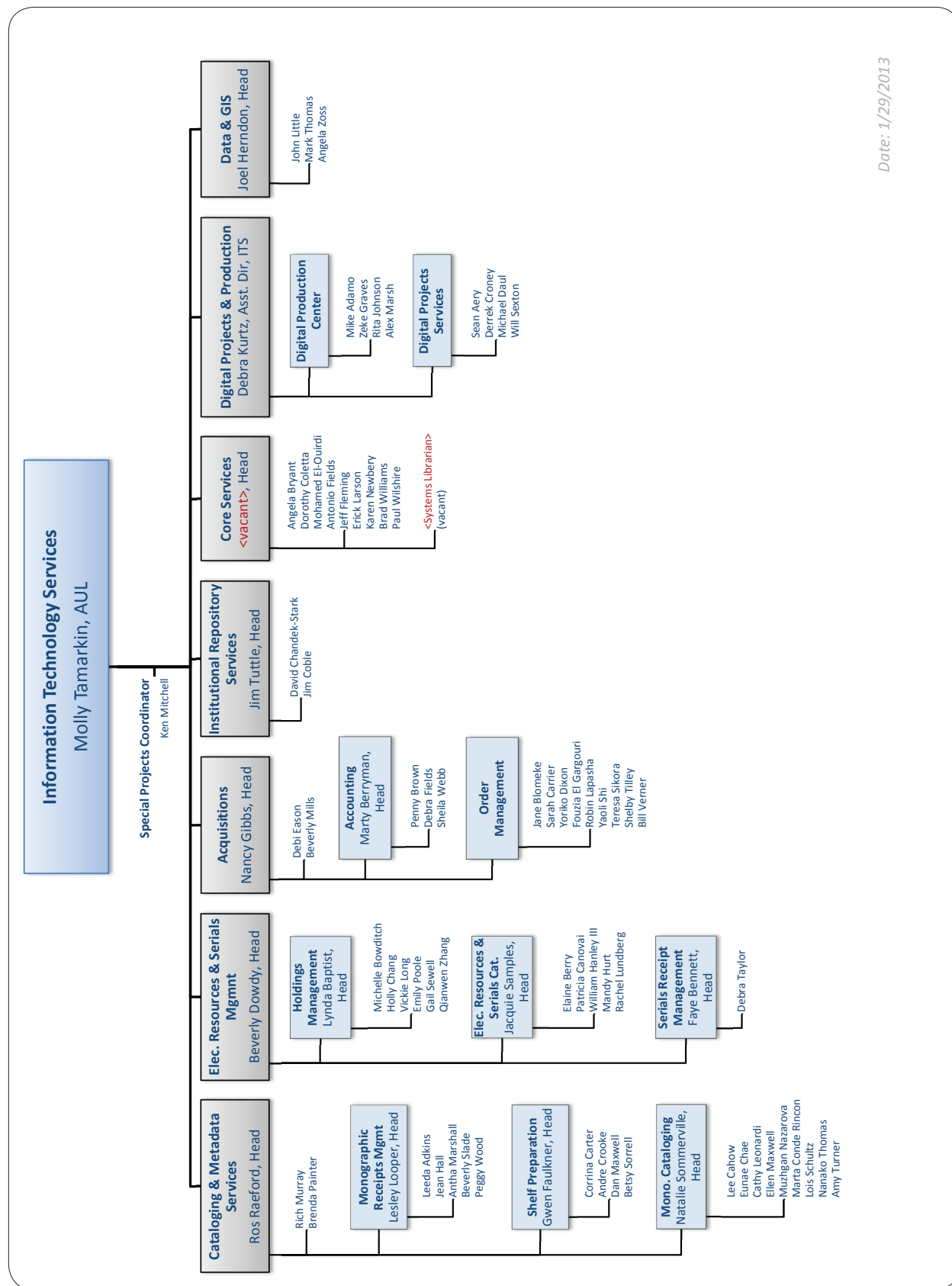
---



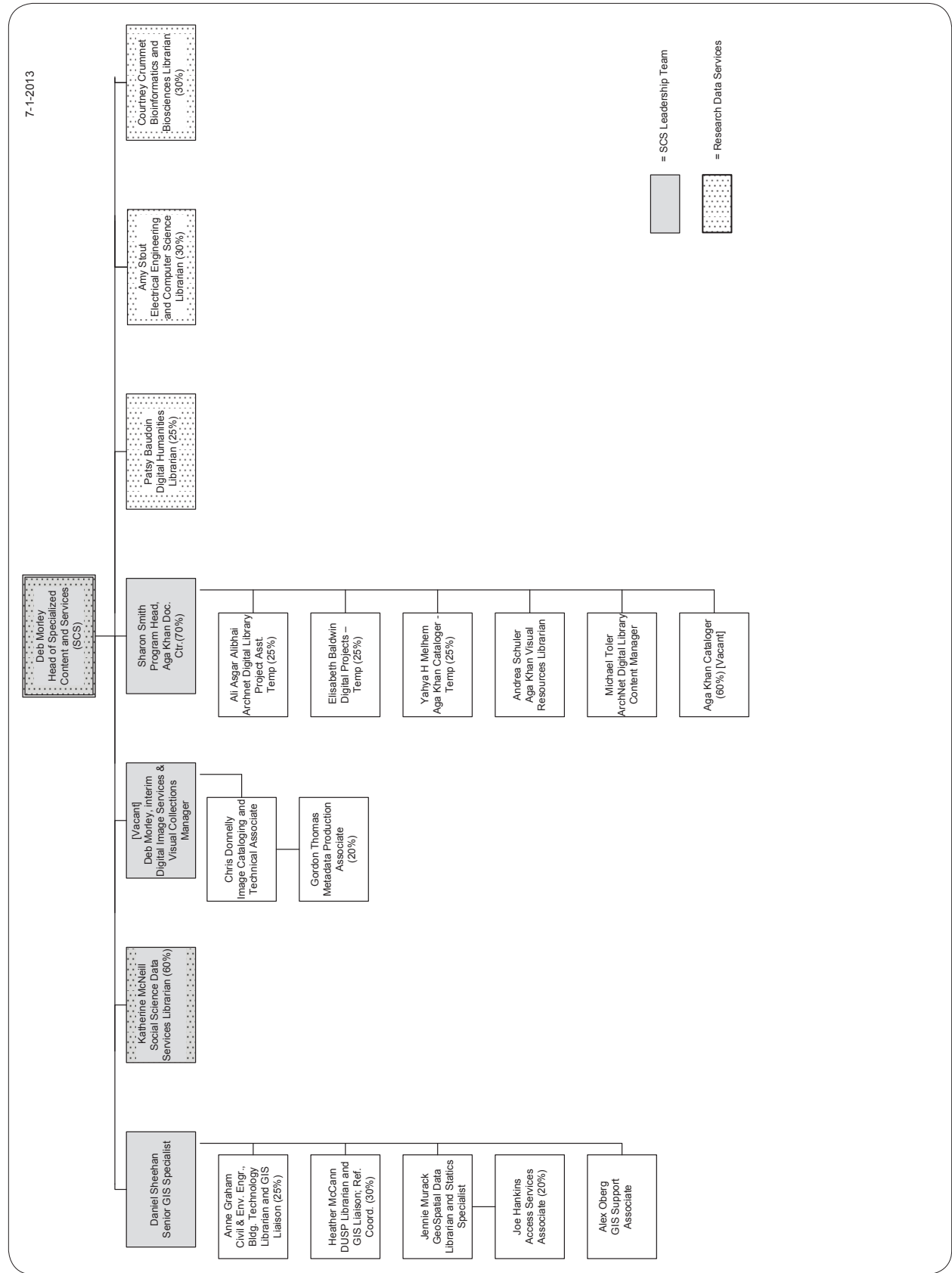


## Appendix A: RDMSG Organizational Structure





Date: 1/29/2013





## Research Data Policies

---

# Researcher Handbook

Home » 6. Conducting Research Responsibly

## 6g. Data Management: Research Records

1. [What are Research Records?](#)
2. [What points should I consider when managing my data?](#)
3. [What am I responsible for?](#)
4. [Who can help?](#)
5. [What are the relevant policies and procedures?](#)

### 1. What are Research Records?

Typically, research records refer to any type of records or materials that document your research effort. These can be electronic or hard copy as in various forms of logs, notebooks, correspondence, videos, computer databases, audio or digital records, or even the actual products of experiments.

In addition to maintaining accurate and complete research records for data analysis, all records relating to the conduct of the project are important including those that document the management of the research funds and the intellectual property.

Although not an inclusive list, research records typically include:

- Laboratory research: lab notes, notebooks, computer databases, microscopic slides, gels, images, photos, videos, laboratory equipment printouts, and records of statistical and other data analysis.
- Animal research: protocol binders with IACUC- approved protocols with all approved modifications, animal health records, surgical or treatment records, breeding records, drug records, research data files.
- Clinical trials: regulatory binders which include CHR approvals, protocols, informed consent documents, monitoring reports, adverse event reports, and other documents pertaining to sponsors, drugs and devices. Other clinical records can include records for research data, data analysis, audio and video tapes of subjects, images of subjects and any other type of record that can identify persons that data were collected from.
- Funding: records and correspondence relating to the grant financial records, purchasing records, scope of work, budgets, and service records.
- All correspondence with granting agencies, institutions, and collaborators.

The University of Iowa Operations Manual includes description of UI's [Records Management Program](#) which provides definitions for different types of UI documents and records.

## 2. What points should I consider when managing my data?

The integrity of your data is dependent upon having and using a system of data management. When determining how data will be collected, recorded, and stored, you should consider the following:

- Are the research records legible, accurate, and complete? Are they in sequence and dated? Is the researcher identified in the records?
- Are there reasonable plans for retention, retrieval and storage of the data?
- Have you managed the data so it can be shared if required by funding agencies?
- Would an audit of the research records support your claims in your publications?
- Could co-investigators confirm the accuracy of the manuscript from the laboratory or research notebooks?

Your research records are the source documents for verification of your research by governmental or University investigations and audits. Clear, permanent records of research are crucial for clarifying any challenges to your data authenticity, authorship and intellectual property.

## 3. What am I responsible for?

As the PI, you should observe sound practices for the maintenance, oversight, and storage of data as you have the final responsibility for the following:

- Validity and quality of the data and manuscripts.
- Fulfilling all departmental and University research standards, policies, and procedures.
- Training and monitoring the performance of your students, research fellows, residents, and staff to assure that each has the knowledge, information, and skills necessary to meet these standards.

At Iowa, researchers are encouraged to retain research data and records for a period of at least five years following publication to provide verification of the validity of the reported results, according to [27.6 c of the University of Iowa Operations Manual](#).

In addition to institutional responsibilities, a growing number of U.S. funding agencies such as the National Science Foundation, the National Institutes of Health, and National Endowment for the Humanities-Office of Digital Humanities require researchers to supply detailed, cost-effective plans for managing research data, called Data Management Plans. These plans typically detail:

- What data will be kept and for how long
- How data will be formatted and described for reuse and interpretation
- Policies around data access, use, and attribution/copyright, and preservation

## 4. Who can help?

UI researchers can seek assistance in developing data management plans from various sources. The [DMP Tool](#) helps researchers create and manage data management plans. The [University of Iowa Libraries subject liaisons](#) and its [Digital Research & Publishing unit](#) also provide advice on developing data management plans and long-term archiving and preservation for small sets through [Iowa Research Online](#).

## 5. What are the relevant policies and procedures?

University of Iowa Policy, Procedures, and Resources:

- [The University of Iowa Operations Manual 17.3 Records Management Program](#)
- [The University of Iowa Guidebook on Records Management](#)
- [Guide for Human Subjects Research at the University of Iowa](#) (guidance on records management can be found in Section F. Record Keeping)
- [UI Information Technology Services – Research Services](#)
- [The Iowa Social Science Research Center](#) offers data access and management services to UI social science researchers.

Partial list of Federal Policies, Procedures, and Resources:

**JOHNS HOPKINS UNIVERSITY  
POLICY ON ACCESS AND RETENTION  
OF RESEARCH DATA AND MATERIALS  
January 2, 2008**

INTRODUCTION

The following policy paper contains parameters for Research Data and Materials Management (hereafter to be referred to as Research Data). In recent years, the amount of scrutiny and inquiry into Research Data has increased from a variety of sources, which has prompted efforts at Johns Hopkins and elsewhere to evaluate and update their Research Data Management practices.

The purpose of this policy is to protect researchers and the university. These measures are designed to address compliance requirements for researchers while diffusing some of the burden associated with Research Data Management. At Johns Hopkins, the department, research administration, divisional and university administration and the researcher are partners in managing and protecting the Research Data produced at the university.

This policy provides an umbrella approach to Research Data Management across the university. Divisional and other policies may also apply but are not to conflict with the overarching policy. This policy has been carefully designed to serve the best interests of our researchers and the university in management of Research Data. This policy is designed to complement, not supersede, other policies of the Johns Hopkins University including (but not limited to) protection of human subjects, HIPAA, intellectual property, financial management, etc. This policy does not apply to academic issues.

1. DEFINITIONS

**RESEARCH DATA AND MATERIALS:** Research Data is defined as information recorded in physical form, regardless of form or the media on which it may be recorded. For the purposes of this policy, Research Data is further defined as including any records that would be used for the reconstruction and evaluation of reported or otherwise published results. Research Data also includes materials such as unmodified biological specimens, environmental samples, and equipment. Examples of Research Data and Materials include laboratory notebooks, notes of any type, photographs, films, digital images, original biological and environmental samples, protocols, numbers, graphs, charts, numerical raw experimental results, instrumental outputs from which Research Data can be derived and other deliverables under sponsored agreements.



**PRIMARY RESPONSIBLE INVESTIGATOR:** The individual who bears primary responsibility for technical, programmatic, fiscal, and administrative requirements of the project.

2. **APPLICABILITY OF POLICY:** This Policy on Access and Retention of Research Data and Materials shall apply to all Johns Hopkins University faculty, staff, postdoctoral fellows, students and any other persons, including consultants, involved in the design, conduct or reporting of research performed at or under the auspices of the University.
3. **OWNERSHIP OF RESEARCH DATA:** The University owns all Research Data generated by research projects conducted at or under the auspices of the Johns Hopkins University regardless of funding source, unless specific terms of sponsorship, other agreements or University policy supersede these rights.

This policy does not attempt to determine relative rights of researchers and issues surrounding collaborative efforts such as authorship.

4. **RETENTION AND ARCHIVING:** The Primary Responsible Investigator of a research project is responsible for selection of an appropriate method of storing and archiving Research Data, and for determining what needs to be retained in sufficient detail and for an adequate period of time to enable appropriate responses to questions about accuracy, authenticity, primacy, and compliance with laws and regulations governing the conduct of research. The Primary Responsible Investigator is responsible for educating all participants in the research project of their obligations regarding Research Data, and for protection of the University's rights and ability to meet obligations related to the Research Data. The Primary Responsible Investigator should also consult with University officials regarding the development of any contingency plans.
5. **RIGHTS TO ACCESS:** The Primary Responsible Investigator will have access to the Research Data generated by the project. Any other faculty, staff, student or person involved in the creation of Research Data may have the right to review that portion of the Research Data that he or she created. The University will have access to the Research Data as necessary for technology transfer, compliance and other purposes. The University also has the option to take custody of the Research Data as determined by the appropriate University official. Such option will not be invoked without cause and subsequent notification of the Primary Responsible Investigator. In some instances, a research sponsor has a legal right of access or access may be requested through the sponsoring agency under the federal Freedom of Information Act (FOIA). Such requests will be coordinated through the Office of the General Counsel and/or the appropriate Research Administration Office.

6. **DESTRUCTION OR REMOVAL:** Research Data must be maintained for the periods required by law, University policy and sponsored agreement terms (See Appendix V). Thereafter, Research Data must not be destroyed without prior approval of the appropriate University official. With respect to removal of the Research Data, the University recognizes the importance of Research Data to the future research and career of its faculty. Therefore, should removal of Research Data be approved, for example, because of the transfer of the investigator to another institution, the following requirements apply:

- I. Researchers may receive approval to remove original Research Data. The University may retain copies.
- II. Research Data generated during the Researcher's employment at the University will be maintained in accordance with Johns Hopkins policy
- III. Research Data that are integral to the ongoing research of another Johns Hopkins employee or student will continue to be made available for that purpose
- IV. The researcher bears full responsibility for making original Research Data available to Johns Hopkins or federal and legal entities upon request.

Others involved in the project may remove copies (but not originals) of the Research Data with permission of the Primary Responsible Investigator.

7. **MAINTENANCE AND REVISION OF THE RESEARCH DATA:** The Primary Responsible Investigator of the research project is the person directly responsible for maintenance of Research Data created on that project. In order to support the project's credibility and the University's rights and ability to meet obligations related to the Research Data, should any revisions to final Research Data be contemplated, the Primary Responsible Investigator must notify the appropriate offices in the University and the originator of the information. The Primary Responsible Investigator must retain the original Research Data. See also Appendix IV.

APPENDICES, WEB LINKS, AND/OR FORMS:

- I. [RESPONDING TO REQUESTS FOR ACCESS BY NON-HOPKINS ENTITIES UNDER FOIA](#) (Policy and Cost Reimbursement Form)
- II. [TRANSFER OF RESEARCH DATA FROM JHU CUSTODIANSHIP](#) (Optional Approval Form)
- III. [LINK TO UNIVERSITY POLICIES](#) (<http://jhuresearch.jhu.edu/policies.htm>)
- IV. [APPROVED METHODS OF ARCHIVAL](#)
- V. [TIME MINIMUMS FOR ARCHIVAL](#)



## REGULATION ON THE CONDUCT OF RESEARCH

### 4. RESEARCH DATA

- 4.1 A Researcher shall collect Data concerning human and animal subjects in accordance with the Regulatory Framework governing the use of such subjects.
- 4.2 A Researcher shall respect the laws governing access to personal information and privacy in his or her collection and use of Data.
- 4.3 A Student may engage in Research in which use of certain kinds of Data, in the custody of a government or Person, is restricted provided that:
- (i) the eventual publication of Research based on the Data is permitted; and
  - (ii) subject to section 4.3.1, any delay in publication does not exceed one (1) year.
- 4.3.1 A request by a third party for a delay in publication of Research undertaken by a Student for his or her thesis that exceeds one (1) year may be agreed to only in exceptional cases and shall require:
- (i) the written consent of the Student; and
  - (ii) the written approval of:
    - (a) the Vice-Principal (Research and International Relations); and
    - (c) the Dean of Graduate and Postdoctoral Studies.
- 4.4 A Researcher shall not use or publish Data which he or she knows to be, or has reasonable grounds to believe are, false or of unknown provenance unless it is so identified.
- 4.5 A Researcher shall organize his or her Data in a manner that allows for its verification by third parties.

### 4.6 *Retention of Research Data*

- 4.6.1 A Researcher shall retain Data in conformity with best practice in his or her discipline and for:
- (i) the period specified by the Agency supporting the Research; or
  - (ii) in the absence of an Agency specification, a period of seven (7) years from publication of the Data.
- 4.6.2 Each department or research unit shall establish procedures appropriate to its needs for the retention and recording of Data.
- 4.6.2.1 Data shall be retained by a Principal Investigator or the department or research unit in which they were generated as agreed to by the Principal Investigator and his or her Chair.
- 4.6.2.2 A Researcher who ceases to be a member of the University shall deposit his or her Data with the department or research unit where the Data were generated unless alternative written arrangements are made with his or her Chair.
- 4.6.3 In the event that Data obtained from a limited access database or under a Research Related Agreement cannot be retained by a Principal Investigator, the Principal Investigator must provide the Chair in writing with the location of the Data or the limited-access database.

### 4.7 *Access to Research Data*

- 4.7.1 Subject to exceptions based on a duty of confidentiality and the laws respecting intellectual property and access to information, a Researcher shall make his or her Data available after publication to an Agency or established scientific or scholarly journal presenting a reasonable and legitimate written request to examine the Data.

- 4.7.2** Where there is a disagreement between the Researcher and the Agency or journal requesting the Data, the disagreement shall be referred for resolution:
- (i) first to the Chair;
  - (ii) then, if necessary, to the Dean; and
  - (iii) finally, if necessary, to the Office of the Vice-Principal (Research and International Relations).

**4.8 Collaborative Data**

- 4.8.1** Research collaborators, at the commencement of their collaboration, shall make all reasonable efforts to reach agreement, preferably in writing, that is consistent with the law and the Regulatory Framework relating to intellectual property, on their rights to, and future use of, Data.
- 4.8.2** In the absence of an agreement between Research collaborators, their rights to and future use of the Data shall be governed by the law and the Regulatory Framework relating to intellectual property.
- 4.8.3** In the event that a dispute should arise between Research collaborators concerning rights to and future use of the Data, the University shall assist in facilitating the resolution of dispute in accordance with section 6.5.

## Research Data Policy

### 1. Objectives

Research Data are a valuable asset to The University of Tennessee (the University). This policy protects the faculty's and University's property rights by addressing definition, responsibility, control, and distribution of Research Data produced during activities supported by the University; supported by external sponsors; or produced with University facilities, resources, or other personnel.

This policy is applicable to Research Data developed by University employees in performing the duties of their employment by the University or through substantial use of funds and facilities provided by the University. This policy assures that Research Data are adequately recorded, archived, retained, and accessible for sufficient time to support the associated research that produced the data and any intellectual property developed by that research. This policy supports the academic freedom for free and broad dissemination of Research Data, consistent with University policy and needs.

### 2. Definition of Research Data

For purposes of this policy, Research Data includes all records necessary for the reconstruction and evaluation of reported results of research and the events and processes leading to those results, regardless of form or media. Research Data may include laboratory notebooks, databases documenting research, and other compilations of information developed during research.

Research Data are distinct and separate from, but may be associated with, other intellectual property such as patentable or copyrightable works, and trademarks. Intellectual property is subject to a separate policy (see The University of Tennessee Statement of Policy on Patents, Copyrights, and Other Intellectual Property), as is Tangible Research Property (see Tangible Research Property Policy).

### 3. Responsibility for Research Data

The University is ultimately responsible for the accuracy and sufficiency of research records, the cornerstone of rigorous research. Therefore, the University is responsible for Research Data developed by University personnel in performing the duties of their employment by the University or through substantial use of facilities or funds provided by the University. Such responsibility applies to research funded by external sources and managed by the University, unless the University agrees to another arrangement in a grant, contract, or other agreement.

The University's responsibility for the scientific record for projects conducted at the University, under University auspices, or with University resources is based upon (a) United States Office of Management and Budget Circular A-110, Sec. 53, (b) the University's need to assess and defend charges of intellectual dishonesty, (c) the University's need to support and commercialize the management of intellectual property, and (d) the University's mission to develop and disseminate new knowledge.

#### **4. Control of Research Data**

The University supports the principle of openness in research. Free dissemination of data, processes, and results of research and other sponsored activity is crucial to a vibrant and healthy academic environment. The University promotes the prompt and open exchange of Research Data with scientific colleagues outside the investigator's immediate laboratory or department, subject to relevant grants, contracts, other agreements, or applicable law.

In the case of externally sponsored research involving a grant, contract, or other agreement, the Principal Investigator (PI) is responsible for controlling storage, use, and distribution of Research Data arising from the research activity, subject to provisions of the applicable grant, contract, or other agreement, or University policy, or applicable law. The PI, or laboratory/department head is responsible in situations where the research is performed without a grant, contract, or other agreement, such as institutionally sponsored research. The PI or laboratory/department head is responsible for the following:

- a) Collection of Research Data, including production of defensible laboratory notebooks;
- b) Management of Research Data ensuring efficient and effective retrieval by the PI, other personnel within the research group, or appropriate administrative personnel or research sponsors;
- c) Development of a formal Research Data plan and procedures where appropriate;
- d) Consideration of a system for preserving Research Data in the event of a natural disaster or other emergency;
- e) Retention of Research Data for the requisite period of time (see below); and
- f) Documented communication of the management system and description of the data managed to members of a research group and to the Chief Research Officer.

Control of Research Data, however, remains at all times subject to the other provisions of this policy.

#### **5. Retention of Research Data**

The PI or laboratory/department head must preserve Research Data for a minimum of three (3) years after the final project close-out, with original data retained where feasible. The following circumstances may require longer retention:

- a) Where data supports a patent, such data must be retained as long as the patent and any derivative patents are valid;
- b) If allegations of scientific misconduct, conflict of interest, or other charges arise, data must be retained until such charges are fully resolved;
- c) If a student is involved, data must be retained at least until the degree is awarded or the student has unambiguously abandoned the work; and
- d) Data must be retained if required by the terms of a grant, contract, or other agreement, or applicable law.

Beyond these periods, destruction of the research record is at the discretion of the PI or the laboratory/department head. Research Data will normally be retained in the administrative unit where generated. Research Data must be retained on a University facility unless specific permission to do otherwise is granted by the Chief Research Officer.

## **6. University Responsibilities**

University responsibilities with respect to Research Data include the following:

- a) Ensuring the academic freedom of the faculty in pursuit of the University's mission of developing and disseminating new knowledge;
- b) Securing and protecting intellectual property rights for Research Data and commercialization of such data where appropriate and feasible;
- c) Protecting the rights, including those of access to data, of faculty, postdoctoral scholars, students, and staff;
- d) Avoiding undue interference with appropriate dissemination of Research Data in an academic community;
- e) Complying with the terms of a sponsored grant, contract, or other agreement;
- f) Facilitating the investigation of charges of scientific misconduct, conflict of interest, and similar charges or disputes; and
- g) Ensuring the appropriate care of animals, human subjects, recombinant DNA, radioactive materials, controlled substances and the like.

## **7. Research Data Transfer When a PI Leaves the University or a Grant is Transferred**

If a PI leaves the University and a research project is to accompany the PI to a new institution, ownership of the data may be transferred with the approval of the Chief Research Officer and with written agreement from the PI's new institution that ensures: (1) its acceptance of custodial and other responsibilities for the data; (2) the University

and any sponsors have access to the data when necessary and upon reasonable notice; and (3) protection of the rights of human subjects.

**8. Resolving Disputes Concerning Research Data Ownership or Policy**

Questions of Research Data ownership or other matters pertaining to the Research Data policy will be resolved by the Chief Research Officer in conformance with applicable University policies.

**9. University Access**

When necessary to assure access to Research Data, the University has the option to take custody of the data in a manner specified by the Chief Research Officer.



## Data Retention Policies

---

## **Policy on Research Records: Sharing, Retention and Ownership**

*As Approved by the Academic Council May 5, 1994*

*Revised by Research Policy Committee January 2007*

The preparation, sharing and retention of appropriate records are essential components of any research endeavor at the University. The University, its faculty and its trainees have a common interest and a shared responsibility to assure that research is appropriately recorded, shared and retained. Original records may be required to protect the University's intellectual property rights, to answer ongoing questions regarding management of a research program, to address possible questions that may arise regarding the propriety of research conduct and to comply with the data sharing requirements of many sponsors. Most importantly, it is essential that original research records be mutually available to all the collaborators on a research project.

### **Definition of Research Records**

Research records include, by way of example but not limitation, material contained in research notes, laboratory notebooks and in other media such as computer disks and instrument printouts. Significant research materials or products generated by any research are also part of the record and should be retained and available.

### **Sharing of Research Records**

Research records must always be available to collaborators (co-investigators, supervisors and their trainees). In collaborative projects, all investigators should know the status of all contributing research records and have access to them consistent with confidentiality restrictions. Investigators also should be aware if their research records are subject to specific data sharing requirements of a sponsor.

### **Retention of Research Records**

Faculty, or the responsible investigators, have the obligation to ensure that, for all aspects of their research program, sufficient records are kept to document the experimental methods and accuracy of data collection as well as the methods and accuracy of data interpretation. This policy does not create an obligation to retain the research records of an unfunded project unless it results in publication or involves the use of animals or human subjects. Research records should be archived for a minimum of five years after final reporting or publication of a project (or longer if required by an external sponsor, law, rule or regulation). The archived records should be the originals. In addition, the records should be kept for as long as may be required to protect any patents resulting from the work. If any questions regarding the research are raised during the required retention period, the records should be kept until such questions are fully resolved. In the event an investigator leaves the University for any reason, the original research records must be retained at the University and the investigator's department and collaborators notified as to their location.

### **Ownership of Research Records**

The primary owner of research records is the University. The University has the right of access to the supporting records for all research carried out through the University with the understanding that information or data that would violate the confidentiality of sources or subjects involved in the research should not be disclosed. In addition, extramural sponsors providing support for research at Duke University may have the right to review any data and records resulting from that extramural support.



## Policy 7.9

### Guidelines for Responsible Conduct of Scholarship and Research

**Responsible Official:** VP for Research Administration

**Administering Division/Department:** Research Compliance

**Effective Date:** April 30, 2007

**Last Revision:** April 30, 2007

#### Policy Sections:

- I. Overview
- II. Applicability
- III. Policy Details
- IV. Related Links
- V. Contact Information
- VI. Revision History

#### Overview

These guidelines describe a standard of practice for the conduct of scholarship and research at Emory University. The University complies with all applicable laws and regulations (see Appendix). The guidelines are intended as a statement of desirable practices. They are based on three important principles:

- I. The University is obligated to protect and foster the academic freedom and intellectual integrity of all members of the University community in their pursuit of knowledge;
- II. The University is accountable to outside funding sources that support the research and scholarship of its faculty; and
- III. Every scholar has ultimate responsibility for the accuracy and validity of his/her own work and that of junior co-investigators, fellows, and students. Each scholar shared this responsibility with colleagues with whom she/he establishes collaborative relationships.

#### Applicability

This document applies to research in all areas of intellectual inquiry. A separate section addresses issues specific to scientific research. These guidelines are intended to heighten awareness of potential ethical problems and to instruct individuals regarding appropriate procedures for resolving and documenting ethics-related matters. The focus is on the individual scholar; the purpose is to emphasize that his/her responsibility includes a duty to maintain high scholarly and ethical standards, and a commitment to instill those standards in co-investigators, students and trainees.

Scientific inquiry, scholarly contributions, creativity, and academic accomplishment can take many forms and may vary among disciplines. The issues addressed by these guidelines are essential to all scholarly activity within the University community. Scholarly responsibility, quality of scholarly activity, security of scholarly contributions and their sources, responsible authorship, and provision for training in ethics of each discipline are issues inherent to all areas. The implications of these guidelines apply as fully to the scholar who co-authors a history textbook as to the laboratory scientist who reports a biological discovery, or the clinician who publishes a case report.

The guidelines address the following concerns:

- the scholar's authority and responsibility for research activities;
- the establishment of the quality of research;
- authorship of publications, including multiple publications and requisites for authorship;
- the supervision of students and other trainees;
- the education of trainees in research ethics and integrity;
- access to and retention of scientific research protocols and data; and
- the social responsibility of the scholar.

**C. Access to and Retention of Scientific Research Protocols and Data**

1. Both the research director and the University have responsibilities and, hence, rights concerning access to, use of, and maintenance of original research data. ("Ownership of Research Data". Estelle A. Fishbein, *Academic Medicine*, 66:129, 1992 and "Workshop Summary". L.J. Rhoades, *Data Management in Biomedical Research: Report of a Workshop*, USPHS, pp. 2-9, 1990.) Consistent with the precepts of academic freedom and intellectual integrity, the investigator/scholar has the primary authority to make judgments involving the use and dissemination of the data.
2. Each faculty member/preceptor is ultimately responsible for the maintenance and proper retention of research records. These records should include sufficient detail to permit examination for the purposes of replicating the research, responding to questions that may result from unintentional error or misinterpretation, establishing their authenticity, and confirming the validity of the conclusions.
3. Each preceptor should maintain a laboratory manual that describes all major procedures. Correspondence with institutional review committees and records of the use of controlled substances and radioactive materials should be maintained as part of the research record in accordance with governmental, regulatory, and University policies.
4. A standardized system of data organization should be adopted and should be communicated to all members of a research group and to the appropriate administrative person. The appropriate administrative person should be determined by the sub-unit.
5. Where feasible, all original primary data are to be retained by the faculty member or by his or her designee. Accepted practices for retaining data vary among disciplines and depends on the perishability nature and logistics of retaining each type of data. Each investigator should treat data properly to ensure authenticity, reproducibility and validity and to meet the requirements of relevant grants and other agreements concerning the retention of data. Primary data should be reserved for a reasonable duration to ensure that any questions raised by the researcher, colleagues, or readers of any published results can be answered. It is recommended that, where feasible, data be retained for seven years; in circumstances where there are no federal or other requirements such as those referred to in the Appendix, sub-units of the University may wish to establish uniform standards and procedures for retention and destruction of data. Data should not be destroyed without proper notification of and approval by an appropriate administrative person. In unusual cases (e.g., data used for a patent application filed by the University), it may be necessary for original data to be kept at the University. Potentially patentable data should be signed and dated by the preceptor at the time they are entered into notebooks or maintained by other methods of retention in the event the results are questioned.
6. In the event the scholar leaves the University, an Agreement of Disposition of Research Data may be negotiated by the scholar and the department chair or dean to allow the scholar's data, notebooks, and other data retention materials (other than clinical research records) to be transferred to the new institution. Consistent with the same precepts, and to fulfill its obligations to funding sources and others, the University will ensure in such agreements access to the transferred data for purposes of review. In unusual cases (e.g., data used for a patent application filed by the University) it may be necessary for original data to be kept at the University. In such cases an individual written agreement shall be signed which preserves the scholar's right to access and copy (where practical) such data. In cases of multi-institutional studies, the institution of the primary study director is ultimately responsible for guaranteeing appropriate access to, use of, and retention of original data.

**7.9.03 References**

["Policies and Procedures for Investigation of Misconduct in Research"](#), Emory University, 2 March 1989.

Modified from "Ethical Guidelines for Publications of Research", *Endocrine Reviews*, 10:1, 1989.

"Authorship and Other Credits", N. Fotion and C. C. Conrad, *Annals of Internal Medicine*, 100:592, 1984.

Adapted from S. J. bird, President's Remarks, "Professional Responsibility", *AWIS Magazine*, 20:2, 1991.

"Ownership of Research Data", Estelle A. Fishbein, *Academic Medicine*, 66:129. 1992.

**APPENDIX V**

**Time Minimums for Research Data Archival**

<b>Research Data</b>	<b>Laws, Policies and Regulations</b>	<b>Time Periods</b>
Proposals not funded	Not defined, but may contain proprietary information	Not defined
Expired Grants and Contracts	- Office of Management and Budget (OMB) Circular A-110* - Grants Policy of Funding Agency	OMB - Three years after completion of the entire research project Federal - follows OMB Private – Varies--see specific policy
Clinical Trials (All relevant records)	Food and Drug Administration (FDA) Notice: “Good Clinical Practices: Consolidated Guidelines”	At least two years after the last approval of a marketing application or at least two years after formal discontinuation of clinical development of the investigational product or longer if required by contract, but in no instance less than three years after the completion of the Clinical Trial
- Patent files - Data in support of patent	U.S. Patent Law	17 years from the date of the patent application
Research Data which supported enactment of a federal, state or local law	Not defined	Indefinite

\* = OMB Circular A110 Uniform Administrative Requirements for Grants and Agreements with Institutions of Higher Education, Hospitals, and Other Non-Profit Organizations”

NOTE: If a sponsored agreement exists, see specific archival requirements contained therein.



NORTHWESTERN  
UNIVERSITY

**Responsible University Official:**

Vice President for Research

**Responsible Offices:** Office for Sponsored Research; Office for Research Integrity

**Origination Date:** December 18, 2012

## RESEARCH DATA: OWNERSHIP, RETENTION AND ACCESS

### Policy Statement

Individual researchers and the University have rights and responsibilities with respect to research data. This Policy describes the basis of data ownership and the standards for the collection and retention of data, in addition to requirements for data access. This Policy also provides guidance with respect to transfer of research data in the event a researcher leaves Northwestern University.

### Reason for Policy/Purpose

This Policy assures that research data are appropriately recorded and archived, and available for review under appropriate circumstances.

### Table of Contents

### Page #

Policy Statement.....	1
Reason for Policy/Purpose.....	1
Table of Contents.....	1
Who Approved This Policy.....	2
Who Needs to Know This Policy.....	2
Website Address for This Policy.....	2
Contacts.....	2
Definitions.....	2
Policy/Procedures.....	3
1.0 Ownership and Responsibilities.....;	3
2.0 Data Retention.....	4
3.0 Transfer in the Event a Researcher Leaves Northwestern.....	5
Forms/Instructions.....	5
Appendices.....	5
Related Information.....	5
History/Revision Dates.....	5

---

## Who Approved This Policy

---

Vice President for Research

---

## Who Needs to Know This Policy

---

Faculty, students, other trainees, staff, and all other members of Northwestern University's research community.

---

## Website Address for this Policy

---

[http://www.research.northwestern.edu/policies/documents/research\\_data.pdf](http://www.research.northwestern.edu/policies/documents/research_data.pdf)

---

## Contacts

---

If you have any questions about this policy, you may contact:

1. Executive Director, Office for Sponsored Research – (312) 503-7955, [osr-chicago@northwestern.edu](mailto:osr-chicago@northwestern.edu) (Chicago) or (847) 491-3003, [osr-evanston@northwestern.edu](mailto:osr-evanston@northwestern.edu) (Evanston)
2. Director, Office for Research Integrity – (312) 503-0054 or [nu-ori@northwestern.edu](mailto:nu-ori@northwestern.edu)

---

## Definitions

---

For the purpose of this Policy, “research data” means all information in whatever form (*e.g.*, both physical and electronic) collected and/or generated in the course of a sponsored research project conducted at the University, under the auspices of the University or with University resources. This includes original and derivatives of research data, including recordings of such data. Examples of research data include, but are not limited to:

- Records necessary for the reconstruction and evaluation of the results of research;
- Data contained in laboratory notebooks;
- Data collected using instrumentation or systems and stored in an electronic format;
- Case report forms and source documentation for human subject research studies.

Pursuant to Northwestern's [Policy on Retention of University Records](#), “Records” include “recorded information of any kind and in any form including writings, drawings, graphs, charts, images, prints, photographs, microfilms, audio and video recordings, data and data compilations, and electronic media, including e-mail.”

This Policy does not address ownership of intellectual property, which is governed by Northwestern's [Patent and Invention Policy](#) and [Copyright Policy](#).

---

## Policy/Procedures

---

### 1.0 Ownership and Responsibilities

As a federally funded research institution, the University, in order to meet the requirements of research sponsors, asserts ownership over the research data for projects conducted at the

University, under the auspices of the University or with University resources. Although the University as owner of the research data must meet the requirements of sponsors, good management practice and practical considerations necessitate that the University and researchers act in partnership to fulfill these obligations.

As custodians of research data, Principal Investigators (PIs) and other researchers are stewards of research data. At the same time, no matter how such responsibilities are delegated, the flow of accountability runs from the PI being responsible to the institution for the stewardship of research data, just as the institution is ultimately responsible to the research sponsor.

Research data are to be accessible to members of the University community, external collaborators and others as appropriate (*e.g.*, for patent applications or journal submissions). Where necessary to assure needed and appropriate access (*e.g.*, for research misconduct investigations), the University may take custody of research data in a manner specified by the Vice President for Research.

Northwestern's responsibilities with respect to research data include, but are not limited to:

1. complying with the terms of sponsored project agreements;
2. ensuring the appropriate use of project resources, *e.g.*, animals, human subjects, recombinant DNA, biological agents, radioactive materials, etc.;
3. protecting the rights of researchers, including, but not limited to, their rights to access to data from research in which they participated;
4. securing intellectual property rights;
5. facilitating the investigation of charges, such as research misconduct or conflict of interest;
6. maintaining confidentiality of research data, where appropriate; and
7. complying with applicable state and federal laws and regulations.

The PI's responsibilities with respect to research data include, but are not limited to:

1. ensuring proper management and retention of research data in accordance with this Policy;
2. establishing and maintaining appropriate procedures for the protection of research data and other essential records, particularly for long-term research projects;
3. ensuring compliance with program requirements;
4. maintaining confidentiality of research data, where appropriate; and
5. complying with applicable state and federal laws and regulations.

## **2.0 Data Retention**

Research data must be retained for a minimum of three years after the financial report for the project period has been submitted. In addition, any of the following circumstances may justify longer periods of retention:

1. research data must be kept for as long as may be necessary to protect any intellectual property resulting from the work;
2. if litigation or other dispute resolution, claim, financial management review or audit related to the research project is started before the expiration of the three year period, or commenced after the three year period but the relevant data and records have not been destroyed, the research data and other project records must be retained until all



- litigation/dispute resolution, claims, financial management review or audit findings involving the records have been resolved and final action taken;
3. if any charges regarding the research arise, such as allegations of research misconduct, research data must be retained consistent with the Northwestern University *Policy on Retention of University Records*, or as otherwise instructed by Northwestern's Office for Research Integrity or Office of General Counsel;
  4. if a student is involved, research data must be retained at least until the student's degree is awarded (or the student otherwise leaves Northwestern University) and any resulting papers are published;
  5. when research is funded by an award to or contract with Northwestern that includes specific provision(s) regarding ownership, retention of and access to technical data, the provision(s) of that agreement will supersede this Policy;
  6. research data from human subject research studies must be maintained consistent with the [Human Subjects Protection Program Policy Manual](#) and the *Policy on Retention of University Records*;
  7. if other regulations, federal oversight, sponsor policies or guidelines, journal publication guidelines or other University policies require longer retention, all applicable sources must be reviewed and the research data must be kept for the longest period of time applicable.

Beyond the period of retention specified here, the destruction of research data is at the discretion of the PI. Destruction of research data must follow applicable federal regulations, Northwestern policies on record retention and data disposal, sponsor requirements and other applicable guidelines.

Research data will normally be retained in the unit where they are produced. Please refer to the *Policy on Retention of University Records* for additional guidance on responsibilities related to the retention of research data and records.

### **3.0 Transfer in the Event a Researcher Leaves Northwestern**

When individuals other than the PI involved in research projects at Northwestern leave the University, they may take copies of research data for projects on which they have worked, subject to relevant confidentiality restrictions. Original data, however, must be retained at Northwestern by the PI.

If the PI leaves Northwestern, and a project is to be moved to another institution, ownership of the original data may be transferred from Northwestern to the PI's new institution upon request from the PI subject to: (a) the prior written approval of the Vice President for Research; (b) written agreement from the PI's new institution that guarantees (1) its acceptance of ongoing custodial responsibilities for the data and (2) Northwestern having access to the original data, should such access become necessary for any reason; and (c) relevant confidentiality restrictions, where appropriate.

---

## **Forms / Instructions**

---

N/A



## Policy: Laboratory Notebook and Recordkeeping

---

**Date:** 06/23/04

**Policy ID:** RES-002

**Status:** Final

---

**Policy Type:** University

**Contact Office:** [Vice President for Research \(Office of the\)](#)

**Oversight Executive:** Executive Vice President and Provost

**Applies To:** University-wide.

**Table of Contents:** [Policy Statement](#)  
[Recording and Storage of Laboratory Data](#)

**Reason for Policy:** This policy describes the University's position of the importance of recordkeeping in research to ensure that complete data is maintained in an accessible format to support verification of research processes undertaken and of the data obtained as an outcome of such processes.

**Policy Summary:**

**Definition of Terms in Statement:** **Laboratory Notebook:** The logbook of all processes and procedures performed in the course of research which shall be kept in such a manner as to enable an investigator to reproduce the steps taken.

**Data:** The results of research procedures.

**Policy Statement:** The investigators and all research fellows, assistants, technicians and students involved in research activities, shall maintain complete and verifiable records of the procedures they have followed in pursuing all research, and the subsequent data they have thereby obtained.

[BACK TO TOP](#)

**Recording and Storage of Laboratory Data:**

The retention of accurately recorded and retrievable results is of the utmost importance in the conduct of research, and it is the responsibility of each investigator to maintain such records in a secure location.

Data and notebooks resulting from sponsored research are the property of the University of Virginia. It is the responsibility of the principal investigator to retain all raw data in laboratory notebooks (or other appropriate format) for at least five years after completion of the research project (i.e., publication of a paper describing the work, or termination of the supporting research grant, whichever comes first) unless required to be retained longer by contract, law, regulation, or by some reasonable continuing need to refer to them.

If the principal investigator leaves the University of Virginia, he or she may transfer such data to another institution, provided that the Vice President for Research and Graduate Studies is informed of this transfer and approves of it. This shall be subject to the proviso that the University is given written assurance that the data will be retained for the required five-year minimum retention period.

The notebook or logbook shall be kept in a secure location where it cannot be removed by an unauthorized person.

[BACK TO TOP](#)

## Policy on Data Stewardship, Access, and Retention

Adopted By: Research Policy Advisory Committee

Adoption Date: December 16, 2010

Approved By: Vice Chancellor for Research

Approval Date: February 21, 2011

- 1.0 Purpose:** Establishes University policy to assure that research data are appropriately maintained, archived for a reasonable period of time, and available for review and use under the appropriate circumstances.
- 2.0 Scope:** This policy shall apply to all University of Wisconsin-Madison faculty, academic staff, visiting scholars, postdoctoral fellows or other trainees, research technicians, and graduate or undergraduate students and any other persons at UW-Madison involved in the design, conduct or reporting of research at or under the auspices of UW-Madison, and it shall apply to all research projects on which those individuals work, regardless of the source of funding for the project.

### 3.0 Definitions:

**Data** means recorded factual material, regardless of the form or media on which it may be recorded, that is commonly accepted in the research community as necessary to validate research findings. For example, data may include writings, films, sound recordings, pictorial reproductions, drawings, designs, or other graphic representations, procedural manuals, forms, diagrams, work flow charts, equipment descriptions, data files, statistical records, and other research data.

This definition pertains to both primary and secondary data. **Primary data** means data generated by research conducted at the University, under the auspices of the University, or with University resources. **Secondary data** means data owned and or generated by another party, data collected from administrative records, or data designated for public use, but used in whole or in part for research conducted at the University, under the auspices of the University, or with University resources.

This definition of **data** excludes research results based on data such as preliminary analyses, drafts of research papers, published papers, plans for future research, peer reviews, or communications with colleagues.

This definition does not supersede any campus policy pertaining to intellectual property.

**Principal investigator (PI)**, for purposes of this policy, means a researcher with primary responsibility for a research project, a definition that applies whether or not the research is sponsored by an external funding source. A PI's responsibility includes both leadership of the scientific/technical aspects and compliance with administrative aspects of the research.

Others on campus, including certain academic staff titles, visiting scholars, postdoctoral fellows or other trainees, and graduate or undergraduate students, who would initiate a research project and are not themselves eligible to be a PI, must identify a faculty member, academic staff person with permanent PI status, or other authorized person to serve as principal investigator.

**Other research contributors** mean any persons other than the PI who have made a substantial contribution to the conception and design of research, acquisition of data, or analysis and interpretation of data. Contributors may include faculty collaborators, academic staff, visiting scholars, postdoctoral fellows or other trainees, research technicians, and graduate or undergraduate students. In general, persons performing narrow technical or clerical tasks would not qualify as contributors.

**4.0 Policy:** UW-Madison must retain research *Data* in sufficient detail and for an adequate period of time to enable appropriate responses to questions about accuracy, authenticity, primacy and compliance with laws and regulations governing the conduct of the research. It is the responsibility of the *Principal Investigator* to determine what needs to be retained under this policy.

**4.1 Scope:** The University's requirements for stewardship of the research record for projects conducted at the University, under the auspices of the University, or with University resources are based on regulation (OMB Circular A-110, Sec. 53), UW System policy, and sound management principles. UW-Madison's responsibilities in this regard include, but are not limited to:

1. Complying with the terms of sponsored project agreements;
2. Ensuring the appropriate use of animals, human subjects, recombinant DNA, disease-causing agents, radioactive materials, and the like;
3. Protecting the rights of students, postdoctoral scholars, and staff, including, but not limited to, their rights to access *Data* from research in which they have participated;
4. Facilitating the investigation of charges, such as scientific misconduct or conflict of interest; and
5. Support university personnel in securing and protecting intellectual property rights.

Where research is subject to an agreement with UW-Madison that includes specific provision(s) regarding retention of and access to *Data* and other records of research conducted under the auspices of the University of Wisconsin-Madison, the provision(s) of that agreement will supersede this policy. However, University of Wisconsin System Financial & Administrative Policies on Extramural Support Administration (G2) Section V.B.(9) "Data" provides that "No agreement shall be entered into with any extramural sponsor which allows for the transfer of the ownership of data."

In the case where an outside party has provided a University of Wisconsin-Madison investigator with secondary *Data* for the purposes of research, requirements to retain research *Data* in sufficient detail and for an adequate period of time will apply to that portion of secondary *Data* used in the research.

**4.2 Stewardship and Retention:** *Principal Investigators* should adopt an orderly system of *Data* organization, access, and retention and should communicate the chosen system to all members of a research group and to the appropriate administrative personnel, where applicable. Particularly for long-term research projects, PIs should establish and maintain procedures for the protection of essential records in the event of a natural disaster or other emergency.

Research *Data* must be archived for a minimum of seven years after the final project close-out, with original *Data* retained wherever possible. Principles of good stewardship would justify longer periods of retention in the following cases:

1. *Data* must be kept for as long as may be necessary to protect any intellectual property resulting from the work;
2. If any charges regarding the research arise, such as allegations of scientific misconduct or conflict of interest, *Data* must be retained until such charges are fully resolved; and;
3. If a postdoctoral scholar or other trainee, graduate student, or undergraduate student is a *Research Contributor*, *Data* must be retained at least until the degree is awarded, training is completed, or it is clear that the individual has abandoned the work.

Beyond the period of retention specified here, the disposal of the research record is at the discretion of the PI and his or her department or work unit (e.g., laboratory). As a practical matter, *Data* may be translated to more efficient storage media as long as the essential nature of the *Data* is not lost. For example, lab notebooks may be scanned, audio recordings transcribed, questionnaires coded and digitized, and the like.

Records will normally be retained in the unit where they are produced. Research records must be retained on the UW-Madison campus, or in facilities under the auspices of University of Wisconsin-Madison, unless specific permission to do otherwise is granted by the Vice Chancellor for Research.

**4.3 Access:** As part of the stewardship of research *Data*, the *Principal Investigator* shall create explicit understandings with *Other Research Contributors* regarding access to and use of *Data*. These understandings ought to reflect access appropriate to one's role and contribution to the conception and design of research, acquisition of *Data*, or analysis, and interpretation of *Data*.

It will also be the responsibility of the *Principal Investigator* to follow the requirements of any sponsored agreements with regard to access to *Data*.

Where necessary to assure needed and appropriate access, the *Principal Investigator*, upon request of the university, must provide the university with research *Data*. Under extraordinary circumstances, such as research misconduct, the university will take all necessary steps to ensure integrity of the *Data* in a manner specified by the UW Policy for Misconduct in Scholarly Research (FP&P II-314).

None of these provisions is intended to supersede the *Principal Investigator's* right to keep *Data* proprietary until the results of the research have been published and the aims of the research have been fulfilled.

**4.4 Transfer in the Event a Researcher Leaves UW-Madison:** When individuals involved in research projects at UW-Madison leave the University or move to a different research group or position at UW-Madison, they may, with PI approval, take copies of research *Data* that they have generated or to which they have made a substantial contribution for projects on which they have worked. Original *Data*, however, must be retained at UW-Madison by the *Principal Investigator*.

If a *Principal Investigator* leaves UW-Madison, and a project is to be moved to another institution, the *Data* may be transferred with the approval of the Vice Chancellor for Research, and with written agreement from the PI's new institution that guarantees: 1) its acceptance of custodial responsibilities for the *Data*, and 2) UW-Madison access to the *Data*, should that become necessary.

**5.0 Roles and Responsibilities:** The *Principal Investigator* is responsible for the stewardship and retention of research *Data* as well as for determinations concerning access to and appropriate use of *Data*.

*Other Research Contributors* are responsible to cooperate with the PI in carrying out the requirements of this policy.

The dean(s) of the school(s)/college(s) in which the PI is appointed may hear appeals concerning issues of access to *Data* and determine who shall have access.

The Vice Chancellor for Research may hear appeals to a dean's determination concerning access to *Data* and make a final determination. The Vice Chancellor for Research may determine, consistent with campus policy, who is eligible to serve as a *Principal Investigator*.

**6.0 Related Documents/Resources:**

University of Wisconsin-Madison Research Data Services  
<http://researchdata.wisc.edu/>

University of Wisconsin-Madison Intellectual Property Policy and Procedures  
([www.grad.wisc.edu/research/ip/policies.html](http://www.grad.wisc.edu/research/ip/policies.html))

## Data Needs Assessment

---



**Research Data Management Group Survey of NSF Principal Investigators at Cornell University**

Developed and administered by: Dianne Dietrich, Eric Chen, Florio Arguillas, Stefan Kramer, and Gail Steinhart

This document contains the survey instrument used in a 2011 survey of NSF PIs at Cornell University.

**Block 0**

As you may be aware, on October 1, 2010, the National Science Foundation announced a new policy requiring a supplementary document for all grant proposals outlining the proposal's data management plan (see <http://www.nsf.gov/bfa/dias/policy/dmp.jsp>). This requirement will take effect on January 18, 2011. Individual programs and directorates within NSF may have additional guidelines. Other major research funders can be expected to implement similar policies, if they have not already done so.

The Research Data Management Service Group (RDMSG, <http://data.research.cornell.edu/>) is conducting this survey to estimate the demand on campus services for data management, and to identify potential gaps in existing services.

It should take you approximately 10 minutes to complete this survey and your participation is voluntary. You will not be required to provide any identifying information unless you choose to.

Your answers will provide valuable information for use in the RDMSG's planning efforts. Some results from this survey, such as general trends, may be used in external reports, but no identifying information or direct quotes will be used without your consent.

This survey will be closed and no further submissions will be accepted after February 1st, 2011.

**Information Sessions**

You are also invited to attend an informational session on the National Science Foundation's (NSF) new policy requiring a data management plan with all grant proposals. The new policy goes into effect January 18, 2011.

Staff from RDMSG will review the new requirement, describe how researchers can obtain assistance from the RDMSG to create data management plans, and answer questions.

Three sessions will be offered:

Thursday, January 13, 1:30-2:30pm, G01 Biotech  
Tuesday, January 18, 9:00-10:00am, 102 Mann Library  
Thursday, January 20, 12:30-1:30pm, 312 Hollister

---

By clicking the next button below, you voluntarily agree to participate in this online survey.

---

**Block 1**

Please answer the following questions with your most recent NSF award in mind.

---

Please specify the NSF directorate of your most recent award.

- Directorate for Biological Sciences
  - Directorate for Computer & Information Science & Engineering
  - Directorate for Education & Human Resources
  - Directorate for Engineering
  - Directorate for Geosciences
  - Directorate for Mathematical & Physical Sciences
  - Directorate for Social, Behavioral & Economic Sciences
  - Office of the Director (includes Office of Cyberinfrastructure, Polar Programs, and others)
-



Would you be interested in any sort of guidance, including consultation, for writing a data management plan in support of an NSF grant application?

- Yes
- No
- I'm not sure

Additional comments

**Block 2**

According to the NSF, a data management plan may include a description of "the types of data, samples, physical collections, software, curriculum materials, and other materials to be produced in the course of the project."

Please specify the types of data you have produced or anticipate producing for this project that you intend to share with others. Check all that apply.

- Text
- Image
- Audio
- Video
- Spreadsheets
- Databases
- Code
- Other
- I'm not sure

Please specify other data types

Please list the file extensions you have produced or anticipate producing for this project that you intend to share with others.

**Block 3**

According to the NSF, a data management plan may include a description of "the standards to be used for data and metadata format and content (where existing standards are absent or deemed inadequate, this should be documented along with any proposed solutions or remedies)."

Does the data you have produced or intend to produce conform to known standards in your discipline?

- Yes
- No
- I'm not sure

Please specify the standard(s) you are using.

"Metadata" refers to descriptive information or documentation about data.

Have you produced or do you anticipate producing metadata for this project?

- Yes
- No
- I'm not sure

Additional comments

Does the metadata you have produced or intend to produce conform to known standards in your discipline?

- Yes
- No
- I'm not sure

Please specify the standard(s) you are using.

Would you make use of a service to produce metadata for this project?

- Yes, and I would be willing to pay for this service
- Yes, but I would not be willing to pay for this service
- No, I would produce metadata myself

**Block 4**

According to the NSF, a data management plan may include a description of "policies for access and sharing including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements."

Do you anticipate needing to consult with an intellectual property specialist to create a license agreement or usage statement for the data you have produced or intend to produce?

- Yes
- No
- I'm not sure

When would you be able to share the data you have produced or intend to produce for this project?

- Immediately after collection
- Immediately after my team has analyzed the data
- Six months or more after my team has analyzed the data
- I would not be able to share this data

What might prevent you from sharing the data you have produced or intend to produce for this project? Check all that apply. (You may also check no boxes if none apply.)

- Little value to others
- Confidentiality or privacy issues
- Commercialization or patent issues
- Some or all of the data I work with has license or usage restrictions that prevent me from sharing
- Data requires secure access I am not capable of providing

Additional comments

**Block 5**

According to the NSF, a data management plan may include a description of "policies and provisions for re-use, re-distribution, and the production of derivatives." Furthermore, Investigators are expected to share with other researchers, at no more than incremental cost and within a reasonable time, the primary data, samples, physical collections and other supporting materials created or gathered in the course of work under NSF grants."

Given the NSF expectation to share data with other researchers, how much data would you intend to share?

- I do not plan on sharing data
- No more than 1 GB
- More than 1GB but less than 100 GB

- More than 100 GB but less than 1 TB
- More than 1 TB but less than 100 TB
- More than 100 TB

---

When you publish your findings from this research project, do you plan on submitting your supporting data to a journal publisher?

- Yes
- No
- I'm not sure

---

Additional comments

---

Do you plan on using a custom solution to share the data you have produced or intend to produce? (i.e., Sharing data on a personal or departmental website or FTP server...)

- Yes, and I plan to do this work in-house
- Yes, and I plan to contract all or part of this work
- No
- I'm not sure

---

Additional comments

---

**Block 6**

According to the NSF, a data management plan may include a description of "plans for archiving data, samples, and other research products, and for preservation of access to them."

---

Do you plan to deposit the data you have produced or intend to produce in Cornell's Institutional Repository, eCommons (<http://ecommons.cornell.edu/about.html>), or would you be interested in doing so to satisfy the NSF requirement?

- Yes
- No
- I'm not sure

---

Additional comments

---

---

Do you plan to deposit the data you have produced or intend to produce in CISER's Data Archive (<http://ciser.cornell.edu/info/about.shtml>), or would you be interested in doing so to satisfy the NSF requirement?

- Yes
  - No
  - I'm not sure
- 

Additional comments

---

Do you plan to utilize the Cornell Restricted Access Data Center [http://ciser.cornell.edu/CRADC/What\\_is\\_CRADC.shtml](http://ciser.cornell.edu/CRADC/What_is_CRADC.shtml)) to work with restricted access or limited use licensed data, or would you be interested in doing so to satisfy the NSF requirement?

- Yes
  - No
  - I'm not sure
- 

Additional comments

---

Do you plan to store the data you have produced in the Center for Advanced Computing Disk Farm (<http://www.cac.cornell.edu/services/storage.aspx>), or would you be interested in doing so to satisfy the NSF requirement?

- Yes
  - No
  - I'm not sure
- 

Additional comments

---

Do you plan to deposit the data you have produced or intend to produce in a data center or other non-Cornell repository, or would you be interested in doing so to satisfy the NSF requirement?

- Yes
- No
- I'm not sure

Please specify the repository (or repositories) you plan to deposit your data into.

**Block 7**

What is your current method of backing up the data you have produced or intend to produce for this project? Check all that apply.

- Own IT infrastructure (e.g., external hard drives)
- EZBackup or other campus-based solution
- Commercial solution (i.e., Google Docs, Amazon S3)
- No backup

Approximately how much data needs to be backed up?

- No more than 1 GB
- More than 1 GB but less than 100 GB
- More than 100 GB but less than 1 TB
- More than 1 TB but less than 100 TB
- More than 100 TB

**Block 8**

The NSF specifies that if "any PI or co-PI identified on the project has received NSF funding in the past five years, information on the award(s) is required." Specifically, applicants must indicate "evidence of research products and their availability, including, but not limited to: data, publications, samples, physical collections, software, and models, as described in any Data Management Plan."

Do you currently keep track of research outputs and their availability?

- Yes
- No

Additional comments

If there was a service offered where you could enter in basic information about your data (including the description, where it was available on the web) to demonstrate compliance with NSF's policy, would you make use of it?

- Yes
- No
- I'm not sure

Additional comments

Do you anticipate or would you be interested in any sort of guidance, including consultation or instruction, for any of the data management plan components mentioned above?

- Yes
- No
- I'm not sure

Additional comments

Which components are you interested in receiving consultation or instruction for?

A review of the datamangement components:

1. the types of data, samples, physical collections, software, curriculum materials, and other materials to be produced in the course of the project;
2. the standards to be used for data and metadata format and content (where existing standards are absent or deemed inadequate, this should be documented along with any proposed solutions or remedies);
3. policies for access and sharing including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements;
4. policies and provisions for re-use, re-distribution, and the production of derivatives; and
5. plans for archiving data, samples, and other research products, and for preservation of access to them.

**Block 9**

A review of the datamangement requirements:

1. the types of data, samples, physical collections, software, curriculum materials, and other materials to be produced in the course of the project;
2. the standards to be used for data and metadata format and content (where existing standards are absent or deemed inadequate, this should be documented along with any proposed solutions or remedies);
3. policies for access and sharing including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements;
4. policies and provisions for re-use, re-distribution, and the production of derivatives; and
5. plans for archiving data, samples, and other research products, and for preservation of access to them.


---

Please share any additional thoughts or concerns you have regarding campus support for complying with funders' data management policies.

---

Would you like to be contacted in the future about your response or participate in focus groups about Data Management Plans? (Please include your contact information if yes.)





**EMORY LIBRARIES** | Robert W. Woodruff Library

Main Library » Guides » Research Data Management Admin Sign In

Research Data Management Tags: data, e-research, electronic\_data\_center, research data management

Research Data Management at Emory University

Last Updated: Jun 18, 2013 URL: <http://guides.main.library.emory.edu/datamgmt> Print Guide RSS Updates SHARE

- Overview
  - Data Management Plans
  - Data Storage and Preservation
  - Data Sharing and Re-Use
  - Data Documentation
  - Data Publication
- Campus Resources
  - More Information
  - Faculty Survey Results
  - Researcher Interviews

**Faculty Survey Results** Comments(0) Print Page Search:  This Guide Search

**Faculty Practices and Perspectives on Research Data Management**


In fall of 2012, in collaboration with the Emory Office of Institutional Research, we invited all Emory University faculty members to complete an online survey of their research data management practices and perspectives.

Over 350 respondents from a wide range of schools and colleges stated that they generate some type of research data (e.g., spreadsheets, text, images, videos, audio files, instrument files, photographs, physical samples/specimens, etc.). Their responses are shown in preliminary form to the right.

To learn more about how Emory researchers manage their research data, we are currently conducting in-person interviews with faculty, research staff, postdocs, and graduate students. If you are interested in participating in these interviews, please see our [interview page](#) for more information.

Comments (0)

**Survey Results**



**EMORY LIBRARIES**

**Survey of Faculty Practices and Perspectives on Research Data Management**


Fall 2012

slideshow Share 1 / 14

[Research Data Management from RDMemory](#)

Comments (0)

**Data Management Specialist**




**Jennifer Doty**

**Contact Info**  
 Electronic Data Center  
 217 Woodruff Library  
 404-727-0498 (P)  
[jennifer.doty@emory.edu](mailto:jennifer.doty@emory.edu)  
[Send Email](#)

**Links:**  
[Website / Blog](#)  
[Profile & Guides](#)

**Subjects:**  
 Research Data Management

**Related Documents**

 RDM Survey Instrument

Comments (0)

**Presentations**



**RDAP13** <https://www.aisis.org/rdap/>  
Research Data Access & Preservation Summit 2013

**Faculty Practices and Perspectives on Research Data Management**

Jennifer Doty & Katherine G. Akers  
Emory University Libraries

slideshow Share 1 / 12

[Lightning Talk, Doty: Faculty Practices and Perspectives on Research Data](#)

Help Yourself : Subject Guides



## Data Management and Publishing

[Home](#)

[Why Manage Your Data?](#)

[Data Planning Checklist](#)

[What is Data?](#)

[Evaluate Your Data Needs](#)

[Funding Requirements](#)

[Data Management Plans](#)

[Writing an NSF Data Management Plan](#)

[Documentation and Metadata](#)

[File Formats](#)

[Organizing Your Files](#)

[Backups and Security](#)

[Sharing Your Data](#)

[Citing Data](#)

[Data Integration](#)

[Ethical and Legal Issues](#)

[Workshops](#)

[Guides to Data Management](#)

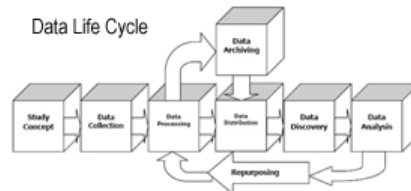
[Managing Research Data 101 \(PDF\)](#)

[Inquiries?](#)

**Related Guides:**

### A Data Planning Checklist

Managing your data before you begin your research and throughout its **life cycle** is essential to ensure its current usability and long-run preservation and access. To do so, begin with a planning process. See also our page on [data management plans](#).



Source: DOI Structural Reform Group "DOI Version 3.0 Conceptual Model" DOI Alliance 2004. Accessed on 11 August 2008. <<http://www.igsl.umich.edu/DOI/committee-info/Concept-Model-3.0.pdf>>

1. What **type of data** will be produced? Will it be reproducible? What would happen if it got lost or became unusable later?
2. **How much data** will it be, and at what growth rate? How often will it **change**?
3. **Who will use it** now, and later?
4. **Who controls it** (PI, student, lab, MIT, funder)?
5. How long should it be **retained**? e.g. 3-5 years, 10-20 years, permanently
6. Are there **tools or software** needed to create/process/visualize the data?
7. Any special **privacy** or security requirements? e.g., personal data, high-security data
8. Any sharing requirements? e.g., [funder data sharing policy](#)
9. Any other funder requirements? e.g., [data management plan](#) in proposal
10. Is there good project and [data documentation](#)?
11. What directory and file [naming convention](#) will be used?
12. What project and data [identifiers](#) will be assigned?
13. What [file formats](#)? Are they long-lived?
14. Storage and [backup strategy](#)?
15. When will I [publish it and where](#)?
16. Is there an ontology or other community standard for [data sharing/integration](#)?
17. Who in the research group will be responsible for data management?

#### Faculty Successes:

"I've had thousands of downloads of my published data--I am impressed that it's been so useful to others!"

Esther Duflo, Abdul Latif Jameel Professor of Poverty Alleviation and Development Economics, MIT

#### For advice on a data management project, contact:

[data-management@mit.edu](mailto:data-management@mit.edu)

**Courtney Crummett**  
Bioinformatics and Biosciences Librarian

**Anne Graham**  
Civil and Environmental Engineering, Building Technology Librarian

**Katherine McNeill**  
Social Science Data Services & Economics Librarian

**Daniel Sheehan**  
Senior GIS Specialist

**Amy Stout**  
Electrical Engineering and Computer Science Librarian

**Data Curation Profiles Toolkit**

About ▾ News Submit a Profile Completed Profiles Workshops ▾ Forums Resources

Welcome to the Data Curation Profiles Community!

A LOT IS GOING ON WITH DATA CURATION PROFILES: THREE (3) NEW TOOLS!

We are in the middle of renovating this space, but we felt we had to share!

Section Number	Section Title	Estimate of time needed for completion
Section 1	Introduction and Background	10 minutes
Section 2	Concepts and Definitions for the Data Curation Profiles	15 minutes
Section 3	The Need for Action	10 minutes
Section 4	Using the Profiles	20 minutes
Section 5	Uses of the Data Curation Profiles Toolkit	10 minutes
Section 6	Assessment	30 minutes
Section 7	Identifying Faculty and Preparing for the Data Interview	25 minutes
Section 8	Research and Data Description	20 minutes
Section 9	The Data Lifecycle	20 minutes
Section 10	Data Sharing	25 minutes
Section 11	Organization and Description	15 minutes
Section 12	Archival Planning	15 minutes
Section 13	Other DCP Modules	15 minutes
Section 14	Contributions to Data Curation Profiles	10 minutes
Section 15	Using the Data Curation Profiles to Identify Needs	15 minutes

The Data Curation Profiles Symposium was recorded and provides a video overview of work involving the Profiles and Toolkit. Additional presentations by experts in the field addressing curation.

<http://docs.lib.purdue.edu/dcpsymposium/>

A new tutorial on using the Data Curation Profiles is available to anyone who wants to learn more about the Profiles and the Toolkit.

Coming soon...

We have a new publication, the Data Curation Profiles Toolkit. You can publish Profiles you write, they can be found and studied, and they will be indexed to be easily found and cited.

<http://docs.lib.purdue.edu/dcp>

**PURDUE UNIVERSITY LIBRARIES**

INSTITUTE of Museum and Library SERVICES

D2C2 DISTRIBUTED DATA CURATION CENTER

This website is an environment where academic librarians of all kinds, special librarians at research facilities, archivists involved in the preservation of digital data and those who support digital repositories can find help, support and camaraderie in exploring avenues to learn more about working with research data and the use of the Data Curation Profiles Toolkit.

A Data Curation Profile is essentially an outline of the “story” of a data set or collection, describing its origin and lifecycle within a research project. The Profile and its associated Toolkit grew out of an inquiry into the changing environment of scholarly communication, especially the possibility of researchers providing access to data much further upstream than previously imagined. If researchers are interested in sharing or forced to provide access to data sets or collections, what does that mean for the data, for researchers, and for librarians?

Data Curation Profiles can:

- provide a guide for discussing data with researchers
- give insight into areas of attention in data management
- help assess information needs related to data collections
- give insight into differences between data in various disciplines
- help identify possible data services
- create a starting point for curating a data set for archiving and preservation

Look around and get to know the site. You will find everything from the history of Data Curation Profiles, the Toolkit for developing a profile of a research data set (registration is required), completed profiles from various disciplines, guidelines for submitting profiles, forums for discussion and resources to learn more about data curation. We hope you will register, download the Toolkit, submit a Profile of your own and join the conversation.

## Data Interview Protocol

**This document is the step-by-step set of instructions for the actual interview. This is to serve as the master copy, accompanied by a question template that is designed to be printed and used to ask focused questions along with check boxes to account for all of the protocol issues. The template will also allow for note taking during the interview.**

### Data Interview Constraints

---

- Interview will consist of:
  - Scientific Data Consultant Group members (two) and subject librarian (one)
  - Researcher (one) being interviewed.
  - Optionally, an additional technical expert invited by the researcher.
- An interview will last no more than sixty minutes.
- Interviews will be semi-structured to allow free-flowing discussion.
- Information to be gathered includes:
  - The state of current data management efforts.
  - Types of digital data created.
  - A prioritized needs assessment covering:
    - Current situation and future needs.
  - Functional specifications for services to meet those needs.

### Mission of the Data Interview

---

At the start of the interview, we will briefly review why we are doing these data interviews and why the library is suited to do it.

- Library goal of supporting researcher needs.
- Library focus on data management.
- Scientific Data Consultant Group Experience:
  - Research Computing Lab, Dataset Task Force, Metadata Steering Group, Institutional Repository Implementation Team
- Purpose of the Data Interview Initiative:
  - Identify common researcher data problems and needs.
  - Identify communities and individuals who are under the most pressure from upcoming grant regulations.
  - Provide data management recommendations and training.
  - Identify the types of digital “data” that are being created.
  - Identify potential partnerships for IR data deposit implementation.
- Remember – there are no “right” answers! We want an honest assessment of your practices. That includes your successes and your failures.
- Mention IRB and give them a copy

## Data Interview Protocol

### What Is Your Data All About?

---

To start the interview we'd like to get some background information on your research. If you'd like, you can discuss your lab's work as a whole, or focus on a specific project.

- 1.1 What question are you trying to answer?
- 1.2 What is the process/method to answer the questions?

### What Kind of Data Do You Have?

---

Now that we've heard about your research, let's talk specifically about what kind of data you produce. i.e. what they create and use, and their attitude towards digital material. Here we are looking for the data characteristics, types, sizes and transformations.

- 2.1 Describe the data you create in your research.  
*Here we are looking for the data characteristics, types, sizes and transformations.*
  - General Category (experimental, simulation/computational, observational, derived/compiled)
  - Creation (sensors, instruments, software)
  - Data Type (docs, emails, databases, images, videos, etc.)
  - Data Format (MS Word, Excel, spss, html, jpg, etc.)
  - Amount (#files, files sizes, growing?)

2.2 Another issue related to data is that of intellectual property. Who owns the Intellectual Property rights of the data you create? Are you familiar with the following UVa policies?

- Lab Notebook Policy
- UVa's Ownership Rights Policy

### How Do You Work With Your Data?

---

Now we'd like to talk about the practices you have in place to organize your data.

- 3.1 Who is responsible for managing the data? Are you using any filing or naming conventions for the files? How are the files organized? Is there any documentation on the files and/or data fields?

## Data Interview Protocol

*Here we are looking for information on managing the data. Are there set procedures? What role does each person play?*

- |   |  |
|---|--|
| <input type="checkbox"/> Management Plan    | <input type="checkbox"/> File Backup/loss/recovery |
| <input type="checkbox"/> Naming Conventions | <input type="checkbox"/> File storage              |
| <input type="checkbox"/> File Organization  | <input type="checkbox"/> Backups                   |
| <input type="checkbox"/> Documentation      |  |

3.2 Do you share data among lab group or other colleagues (e-mail, shared drive, removable devices, CD, web pages, other)? Do you typically have multiple people working on the same data files? If so, have you had issues regarding which version was “correct” or the latest? How are these issues controlled or resolved?

- File sharing
- Issues related to multiple file versions

### Preservation Concerns

---

We are looking for any digital preservation issues in this section of the interview. Continue discussion to ascertain whether any issues have been encountered when creating and using digital material to identify areas where practices could improve.

*Here we are looking for preservation issues on their own data in their own lab/computer.*

4.1 What challenges have you faced in terms of storage, formats, costs, and continued access to older data?

- |  |   |
|--|---|
| <input type="checkbox"/> Do they have older files? | <input type="checkbox"/> Lost or misplaced data |
| <input type="checkbox"/> Obsolete data formats     | <input type="checkbox"/> Storage space          |
| <input type="checkbox"/> Obsolete media            | <input type="checkbox"/> Costs                  |

## Data Interview Protocol

### Data Sharing and Long-term Accessibility

---

Get them thinking about the future of their data i.e. how can these files continue to be accessed and used (if appropriate), do they need to be preserved, if so, for how long?

5.1 Have you been asked to provide or share your data? Could or should your data be reused or repurposed by others, and if so, how and by whom?

- |  |  |
|--|--|
| <input type="checkbox"/> Publisher requirement | <input type="checkbox"/> Restrictions (Confidentiality, Sensitivity) |
| <input type="checkbox"/> Funder requirement    | <input type="checkbox"/> Documented for sharing                      |

### Long-term Preservation

5.2 Do your files need to be preserved? For how long? Does all of it need to be kept?

- Raw or processed data or both
- Who decides? Who is responsible?
- Where?
- Libra, the UVa IR
- How long?

### What Would Make Data Management Easier for You?

---

Ask where the interviewee currently gets advice and support and what else s/he would like to see provided by the University. Key thing is to gauge desire for preservation policy, suggested coverage and any supplementary support needed to implement it.

6.1 What would help you create and manage your data better?

6.2 Who should be responsible for digital preservation? Who should be responsible for funding it?

- Preservation responsibility
- Help, where?
- Library

6.3 What sort of impact might a University-wide policy on data preservation have upon you? What sort of policy do you think would be reasonable?

## Data Interview Protocol

### Follow-Up Plans

---

Review the steps that will come after the interview is complete (Script for interview is included below).

- 7.1 Team combines interview notes.
- 7.2 Send aggregated report to researcher for review/approval, corrections/additions on notes. To expediate things we need the approval/feedback back within one week.
- 7.3 Ask for feedback for interview process.
- 7.4 Provide a complete report that includes a summary of the conversation, responses to the interview questions, and recommendations on how to improve your data management.

### Script from the Data Interview Template (for this section):

Thank you for participating in our Data Interview. Here are our next steps:

- 7.1 Andrew, Sherry and I will combine our interview notes.
- 7.2 I will send you an aggregated report for your review/approval, corrections/additions on our notes. Please return the approval/feedback within one week.
- 7.3 When you send the report back to us, we would like to have your feedback on our interview process.
- 7.4 Once we have your comments on the report. I will provide you with a complete report that will include a summary of the conversation, responses to the interview questions, and recommendations on how to improve your data management.



## RDM Staff Resources

---

**UF** | George A Smathers Libraries Smathers Libraries Site

- Data Management Planning
- Storage, Backup, & Security
- Preservation & Data Sharing
- Copyright & Ethical Use of Data
- Frequently Asked Questions (FAQ)
- Workshops & Trainings
- Speaker Series
- Data Management Task Force
- Smathers Libraries Home

Laurie N. Taylor  
Digital Humanities Librarian  
533 Library West  
Gainesville FL 32611  
Phone: 352.273.2902  
Email: [laurien@ufl.edu](mailto:laurien@ufl.edu)

## Data Management/Curation Task Force

Please see the [charge for background information](#) on this collaborative Task Force with the UF Libraries and UF Research Computing.

- [Task Force Charge](#)
- [Email List](#)
- [Related Wiki](#)
- [UF Research Computing](#)
- [Materials for our work \(reports, templates for presentations and group discussions, etc.\)](#)

**Meetings:**  
Reminders are sent on the email list. Meetings are every other Wednesday from 1-2pm.  
Locations rotate: Health Science Center Library C2-41, Library West 429, and Marston Science Library L107.  
Next Meeting: July 24, 1-2pm, MSL

**Overall Activities**

- Description of Responsibilities from the Charge**  
This group is charged to assess needs, make recommendations, and develop support for the role of the Libraries in campus-wide data management and curation.
  - Specific advisory activities include:
    - Formally assess, through surveys, interviews, and focus groups, campus-wide data management needs and current support resources and activities
    - Review and consider the best practices and models of peer institutions
    - Develop recommendations for the Libraries' campus-level role in support of data management and curation
    - Propose a corresponding framework and resources for library support of the data life cycle
    - Recommend the role of the institutional repository and research computing in storing, finding, and accessing working and final data, and linking publications to supporting data
    - Recommend a framework for liaisons and subject specialists to incorporate data instruction and consultation into their workflows
  - Specific operational activities include:
    - Develop materials and sessions for training of liaisons, subject specialists, and other library staff to prepare them to support campus data management services
    - Develop training and outreach materials to be used by liaisons, subject specialists, and other library staff in their work with clients
    - Develop means to enhance and expand the librarian liaison model with the goal of making librarians partners in research activities
    - Develop and implement templates and support training and services for the DMPTool (Data Management Plan Tool) and other resources
  - Additional goals identified by the group (draft):**  
Ensure that the recommendations and plans resulting from the group both support immediate campus needs and support ongoing, long-term needs for full support including, but not limited to:
    - Develop and promote training and resources: for data management/curation related concerns, finding data, citing data, creating data management plans, and implementing data management
    - Develop and submit recommendations for a full approach for supporting the implementation of data management and curation across campus, for collaborative implementation and support with the libraries, Office of Research, Research Computing, and others as appropriate, along with resources and requirements for the recommendations to be fully operational

**Current Activities and Coordinator**

- [SURA, collaboration on pilot test for Dataverse Network](#) (Laurie, all)
- [DMP Tool, customization for UF resources](#) (Val)
- Survey (Hannah and Rolando)
- Focus Groups (all)
- Sharing and promoting activities (all)
- Coordinating existing training activities related to data management (all)
  - Including training by [Research Computing](#), Libraries (ICPSR, Census and Gov Data, [Best Practices in Research Data Management at HSCL](#), HSCL specific classes, etc.), and others
- Coordinating and promoting speakers and events related to data management (all)
- Integrating resources within a single consolidated portal for all users, with the [Research Data Management LibGuide](#)

**Draft timeline for current activities in the meeting agendas and reports:**  
<http://ufdc.ufl.edu/AA00014835/00001/allvolumes>

[Staff web](#) | [Conduct in the library](#) | [Contact us](#) | [Privacy policy](#) | [Staff login](#)

Send suggestions and comments to the [library web manager](#).  
© 2013 University of Florida George A. Smathers Libraries.  
All rights reserved.  
Terms of Use for Electronic Resources and Copyright Information  
This page uses Google Analytics - ([Google Privacy Policy](#))

**UF UNIVERSITY of FLORIDA**  
*The Foundation for The Gator Nation*

The screenshot shows the Purdue University Libraries website. At the top, the logo for Purdue University Libraries is displayed. Below the logo, the page title is "Supporting Information for Data Services" and there is an "Admin Sign In" link. The main content area includes a central location for handouts, talking points, and articles. A navigation menu contains links for Home, FAQ/Glossary, Resource Guides, Other Libraries' Data Services, Purdue-authored articles, Workshop Handouts, and Webliography. A search bar is located at the bottom of the navigation menu. The main content area is divided into three columns: "Mentors" (listing contact information for Megan Sapp Nelson, Jeremy Garritano, Jake Carlson, Michael Witt, and Lisa Zilinski), "Welcome" (a message from the Data Education Working Group), and "Purdue University Research Repository (PURR)" (describing the repository's purpose). At the bottom, there is a footer with a disclaimer about equal access and a small logo.

**PURDUE LIBRARIES**  
UNIVERSITY

Libraries » LibGuides » Supporting Information for Data Services Admin Sign In

### Supporting Information for Data Services

Central location for all handouts; talking points; Purdue authored articles; Data literature; Tools and any other useful files

Last Updated: Jun 12, 2013 | URL: <http://guides.lib.purdue.edu/dataservices> | [Print Guide](#) | [RSS Updates](#) | [Email Alerts](#) | [SHARE](#) [f](#) [t](#) [e](#) ...

[Home](#) [FAQ/Glossary](#) [Resource Guides](#) [Other Libraries' Data Services](#) [Purdue-authored articles](#) [Workshop Handouts](#) [Webliography](#)  
[Helpful Data Literature](#) [Tools](#) [DMP Requirements](#) [Sample DMPs and Templates](#) [Data Curation Profiles](#) [Conferences](#)

[Home](#) [Print Page](#) Search:  [This Guide](#) [Search](#)

#### Mentors

The following individuals are happy to consult with you for preparation or carrying out data interviews and other data work. If you would like to be added to this list, please contact Megan.

Megan Sapp Nelson - [msn@purdue.edu](mailto:msn@purdue.edu)

Jeremy Garritano - [jgarrita@purdue.edu](mailto:jgarrita@purdue.edu)

Jake Carlson - [jrcarlso@purdue.edu](mailto:jrcarlso@purdue.edu)

Michael Witt - [mwitt@purdue.edu](mailto:mwitt@purdue.edu)

Lisa Zilinski - [lzilins@purdue.edu](mailto:lzilins@purdue.edu)

#### Welcome


This page was created by the Data Education Working Group to assist library faculty to find relevant, useful resources to support developing roles in data curation. The page has been designed to house a variety of resources, which have been selected based upon previous utility and relevance. This includes locally produced documents relevant to Purdue's data services such as the Data Management Plan documents and the Data Curation Profiles. Relevant resources produced by other institutions are highlighted as well. To recommend resources for inclusion, please contact the individuals listed at the left. If you have feedback on this tool, please give that to the authors listed at the left as well.

#### Purdue University Research Repository (PURR)

The Purdue University Research Repository (PURR) provides an online, collaborative working space and data-sharing platform to support the data management needs of Purdue researchers and their collaborators.

Powered by Springshare; All rights reserved. Report a tech support issue.  
View this page in a format suitable for printers and screen-readers or mobile devices.

Purdue University is an equal access/equal opportunity university.  
If you have trouble accessing this page because of a disability, please contact [accessibility@lib.purdue.edu](mailto:accessibility@lib.purdue.edu).



A discussion group at Yale for issues relating to data, science, and research support across disciplines.

[home](#) [upcoming events](#) [meeting notes](#) [useful links](#) [bibliography](#)

[home](#) [Print Page](#)

Search:

#### What is DaEG?

DaEG is an informal discussion group interested in data curation, management, sharing, publication, citation and the services and infrastructures surrounding those activities at Yale and universities and data centers in the US and internationally.

Topics discussed are: the latest in data literature, what's going on at Yale, how we can learn more, implement best practices, and spread awareness about data at Yale.

Join the mailing list here to be informed of upcoming readings and meetings:  
<http://mailman.yale.edu/mailman/listinfo/daeg>

Questions? Email [michelle.hudson@yale.edu](mailto:michelle.hudson@yale.edu)



#### Related research guides

##### [Data and Statistics in the Social Sciences](#)

by [Kristin Bogdan](#), [Michelle Hudson](#) - Last Updated May 23, 2013

Sources for locating published statistics on topics broadly related to the social sciences, as well as numeric datasets for statistical analysis.

2,959 views this year

##### [eScience Institute 2012](#)

by [Michelle Hudson](#) - Last Updated Mar 7, 2013

Materials for those of us participating in the eScience Institute

82 views this year

##### [Research Data Management](#)

by [Kristin Bogdan](#), [Michelle Hudson](#), [Melanie Maksin](#), [Stacey Maples](#) - Last Updated Mar 8, 2013

Resources for learning about best practices in research data management across a variety of disciplines.

745 views this year

##### [Science Data Resources](#)

by [Kristin Bogdan](#), [Kayleigh Bohemier](#), [Michelle Hudson](#) - Last Updated Aug 17, 2012

Resources for data in the sciences.

86 views this year

Powered by [Springshare](#); All rights reserved. [Report a tech support issue](#).  
View this page in a format suitable for [printers](#) and [screen-readers](#) or [mobile devices](#).

## Outreach Materials

---

RESEARCH DATA SERVICES

Need help with your data?

Librarians can help you:

find data

share data

manage data

cite data



email: [data-help@colorado.edu](mailto:data-help@colorado.edu)  
<http://data.colorado.edu>

RESEARCH  
DATA  
SERVICES



NEED HELP?

- Writing a data management plan
- Locating a repository for sharing and archiving your data
- Identifying best practices for research data management
- Finding data sets for your research
- Citing data sets
- Tracking data impact

Contact us at [data-help@colorado.edu](mailto:data-help@colorado.edu)

UF
George A Smathers Libraries
Smathers Libraries Site

- Data Management Planning
- Storage, Backup, & Security
- Preservation & Data Sharing
- Copyright & Ethical Use of Data
- Frequently Asked Questions (FAQ)
- Workshops & Trainings
- Speaker Series**
- Data Management Task Force
- Smathers Libraries Home

**Speaker Series**

The Data Management/Curation Task Force at UF is a collaborative Task Force with representatives from the George A. Smathers Libraries, Research Computing, and the Office of Research. In order to promote awareness of data management and curation concerns on campus and broader impacts for research, teaching, and service, the group is highlighting existing and planning new speaker events related to of data management and curation. Highlighted speakers and events are being noted as being part of the Data Management/Curation Speaker Series.

**Events for 2013-2014:**

**Big Data Event, August 7, 1-2pm**

**UF Digital Humanities Day and THATCamp-UF**  
April 24-25, 2014

**Past Events:**



**Dense, Intense and Complex Data Workshop, June 19**

Andrea Matsunga/ Mauricio Tsugawa	<a href="#">Big Data Support for Scientific Disciplines through Information Technology Engineering</a>
S. Balachandar	<a href="#">Big Data from simulations of extreme and environmental problems</a>
Jim Jones	<a href="#">Data and Research for Sustainable Food Security</a>
Liang Mao	<a href="#">Big Geographic Data and GISciences</a>
Paul Gader	<a href="#">Big Data for Environmental Monitoring</a>
Herman Lam	<a href="#">Big Data Meets High-Performance Reconfigurable Computing</a>
Sanjay Ranka	<a href="#">Big Data: Research and Education</a>
Erik Deumens	<a href="#">HiPerGator and infrastructure for working with data</a>
Bill Farmerie	<a href="#">A role for glue people in big data research</a>
Pam Soltis	<a href="#">Big Data in Biodiversity Studies</a>
Eric Triplett	<a href="#">Big Data and the search for a microbial cause for disease</a>
Mike Conlon	<a href="#">Opportunities for Big Data Medical Records</a>
Betsy Shenkman	<a href="#">The Power of Multiple Data Streams and Big Data in Health and Health Care</a>
Brad Barbazuk	<a href="#">Biology and Big Data</a>
Sixue Chen	<a href="#">Proteomics Data Generation and Analysis Toward Systems Biology</a>
Art Edison	<a href="#">Big data prospective from metabolomics</a>
Thomas Mareci	<a href="#">Imaging Structure and function in Biology with Magnetic Resonance</a>
Joanna Long	<a href="#">Big data, big noise, and big simulations</a>
Kamran Mohseni	<a href="#">Novel methods for in situ measurement and simulation of hurricanes</a>
Daisy Wang	<a href="#">Big Data Systems for Knowledge Base Construction from Text, Images and Crowds</a>
George Lan	<a href="#">Empowering nonlinear and stochastic optimization for large-scale data analysis</a>
David Hale	<a href="#">Automated Analysis of Traffic Simulation</a>
Kevin Knudson	<a href="#">Topological Data Analysis</a>


**National Agricultural Library: A Vision for Preservation and Accessibility of Agricultural Data**  
Dr. Simon Liu, Director of the National Agricultural Library  
Wednesday April 10th, 3:00 p.m., Smathers Library, Room 1A

**Digital Humanities Day and Interface**  
Thursday April 25th, 9:00 a.m. to 4:15, Smathers Library, Room 1A

For more information or to request a workshop, contact the Data Management Task Force: [datamgmt@lists.ufl.edu](mailto:datamgmt@lists.ufl.edu)

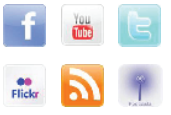

Faculty Graduate Students Undergraduates Visitors IUCAT
 
Ask a Librarian 
Libraries & Hours

[Libraries](#) » [Data Management](#) » [Workshops](#)



- [Data Management Home](#)
- [Data Management at IU](#)
- [Describing Data with Metadata](#)
- [Storage and Preservation](#)
- [Sensitive Data](#)
- [Funder Requirements and Data Management Plans](#)
- [Workshops](#)

Connect with us online



Contact Information

Contact: Stacy Konkiel

Location: 1320 E. 10th Street  
Wells Library E159  
Bloomington, IN 47405  
Phone: 812-856-5295  
Email [iuswdata](mailto:iuswdata)

» [View Map of IUB Libraries](#)

last updated: 6/12/2013

## Workshops

The IU Data Management Service brings together experts in research data preservation and management policy to bring you a series of workshops co-sponsored by the IU Libraries, the OVPR, and ORA.

### Past

**Meeting the NSF Data Management Plan Mandate at IUB**

Thu Jan 24, 2013 1:00 pm - 3:00 pm  
Persimmon Room, IMU

Learn the fundamentals for preparing a data management plan that conforms to the January 2011 NSF mandate. Find information on free, fully supported campus resources for data storage, access, and preservation; resources for DMP development; and key staff that can help you develop your proposal. Q&A will follow the presentation. Anyone interested in or planning to apply for NSF funding should attend.

**RCR Series: Data Management**

Mon Feb 18, 2013 10:00 am - 12:00 pm  
Redbud Room, IMU

In collaboration with other academic units both in Bloomington and Indianapolis, REEP has developed an on-going series of workshops for post-docs and graduate students covering the topics of responsible conduct of research. Sessions are offered on each campus twice per semester on various topics.

**Teaching Research Ethics (Poynter): Responsible Data Management**

May 14-17, 2013  
IMU

Each year the Poynter Center at Indiana University and additional sponsors offer the Teaching Research Ethics Workshop (TRE) to provide training for those involved in teaching research ethics or in administering research programs. The workshop emphasizes a variety of pedagogical approaches to teaching research ethics through sessions on ethical theory, research ethics, trainee and authorship issues, assessment and evaluation, responsible data management, integrity in research, conflict of interest, and international research.

**Managing Your Research Data at IUB**

Fri Sep 21, 2012 12:00 pm - 1:00 pm  
Wells Library E174

Learn about the research data storage, preservation and access resources that IUB has to offer researchers. We will also examine funder mandates for data management planning and how to meet them.

**Meeting the NSF Data Management Plan Mandate at IUB**

Tue Oct 9, 2012 10:00 am - 11:00 am  
Office of Research Administration

Learn the fundamentals for preparing a data management plan that conforms to the January 2011 NSF mandate. Find information on free, fully supported campus resources for data storage, access, and preservation; resources for DMP development; and key staff that can help you develop your proposal. Q&A will follow the presentation. Anyone interested in or planning to apply for NSF funding should attend.

Copyright © 2010 The Trustees of Indiana University | [Copyright Complaints](#) | [IUB](#) | [Libraries Privacy Policy](#) | [News & Events](#) | [Contact Us](#) | [About IUB Libraries](#) | [Outages and Downtime](#)  
Intranet (Library Staff Only)





## **SELECTED RESOURCES**



## Research Data Management Services (RDMS) and the Library

---

- Antell, Karen, Jody Bales Foote, Jaymie Turner, and Brian Shults. "Dealing with Data: Science Librarians' Participation in Data Management at Association of Research Libraries Institutions." *College & Research Libraries*. Accepted: March 7, 2013; Anticipated Publication Date: July, 2014. <http://crl.acrl.org/content/early/2013/04/05/crl13-464>.
- Association of Research Libraries. "E-Research Focus Area." 2013. <http://www.arl.org/focus-areas/e-research>.
- Canadian Association of Research Libraries. "Data Management Sub-Committee." 2013. <http://www.carl-abrc.ca/en/scholarly-communications/data-management-sub-committee.html>.
- Christensen-Dalsgaard, Birte, et al. "Ten Recommendations for Libraries to Get Started with Research Data Management." *Final report of the LIBER working group on E-Science/Research Data Management*. Association of European Research Libraries, 2012. <http://www.libereurope.eu/news/ten-recommendations-for-libraries-to-get-started-with-research-data-management>.
- Graham Pryor, editor. *Managing Research Data*. London: Facet Publishing, 2012.
- Heidorn, P. Bryan. "The Emerging Role of Libraries in Data Curation and E-science." *Journal of Library Administration* 51, no. 7/8 (2011): 662–72. <http://dx.doi.org/10.1080/01930826.2011.601269>.
- Jones, Sarah, Graham Pryor, and Angus Whyte. "How to Develop Research Data Management Services — a Guide for HEIs." DCC How-to-Guides. Edinburgh: Digital Curation Centre, 2013. <http://www.dcc.ac.uk/resources/how-guides/how-develop-rdm-services>.
- Jones, Sarah, Marieke Guy, and Miggie Pickton. "Research Data Management for Librarians." Edinburgh: Digital Curation Centre, 2013. [http://www.dcc.ac.uk/webfm\\_send/1245](http://www.dcc.ac.uk/webfm_send/1245).
- Lewis, Martin. "Libraries and the Management of Research Data." In *Envisioning Future Academic Library Services: Initiatives, Ideas and Challenges*, edited by Sue McKnight, 145–68. London: Facet Publishing, 2010. <http://eprints.whiterose.ac.uk/11171/>.
- Monastersky, Richard. "Publishing Frontiers: The Library Reboot." *Nature* 495, no. 7442 (2013): 430–32. <http://dx.doi.org/10.1038/495430a>.
- Reznik-Zellen, Rebecca, Jessica Adamick, and Stephen McGinty. "Tiers of Research Data Support Services." *Journal of eScience Librarianship* 1, no. 1: Article 5 (2012). <http://dx.doi.org/10.7191/jeslib.2012.1002>.

Tenopir, Carol, Ben Birch, and Suzie Allard. "Academic Libraries and Research Data Services: Current Practices and Plans for the Future." *Association of College and Research Libraries White Paper* (June 2012). [http://www.ala.org/acrl/sites/ala.org.acrl/files/content/publications/whitepapers/Tenopir\\_Birch\\_Allard.pdf](http://www.ala.org/acrl/sites/ala.org.acrl/files/content/publications/whitepapers/Tenopir_Birch_Allard.pdf).

University of Massachusetts Medical School. "e-Science Portal for New England Librarians: A Librarian's Link to e-Science Resources." <http://esciencelibrary.umassmed.edu/>.

Wakimoto, Jina Choi. "Developing Research Data Management Services." *EDUCAUSE Review Online* (2013). <http://www.educause.edu/ero/article/developing-research-data-management-services>.

## Supporting Data Management Planning

---

Association of Southeastern Research Libraries (ASERL) and Southeastern Universities Research Association (SURA) "ASERL/SURA Research Data Management Group." 2013. [http://www.lib.ua.edu/wiki/sura/index.php/Main\\_Page](http://www.lib.ua.edu/wiki/sura/index.php/Main_Page).

"DMPTool Bitbucket Wiki." 2013. <http://bitbucket.org/dmptool/main>.

Fary, Michael, and Kim Owen. "Developing an Institutional Research Data Management Plan Service." *White Paper. EDUCAUSE: ACTI-DM Working Group* (2013). <http://www.educause.edu/library/resources/developing-institutional-research-data-management-plan-service>.

Hswe, Patricia, and Ann Holt. "NSF Data Sharing Policy." *Association of Research Libraries, Guide for Research Libraries*, 2011. <http://www.arl.org/focus-areas/e-research/data-access-management-and-sharing/nsf-data-sharing-policy>.

## Data Archiving

---

Mayernik, Matthew S., et al. "The Data Conservancy Instance: Infrastructure and Organizational Services for Research Data Curation." *D-Lib Magazine* 18, no. 9/10 (2012). <http://www.dlib.org/dlib/september12/mayernik/09mayernik.html>.

University of California, San Diego (UCSD). *Blueprint for the Digital University: A Report of the UCSD Research Cyberinfrastructure Design Team*. 2009. <http://rci.ucsd.edu/files/Blueprint.pdf>.

Witt, Michael. "Co-designing, Co-developing, and Co-implementing an Institutional Data Repository Service." *Journal of Library Administration* 52, no. 2 (2012). <http://dx.doi.org/10.1080/01930826.2012.655607>.

## Staffing and Staff Training

---

- Auckland, Mary. *Re-skilling for Research*. London: Research Libraries UK, January 2012. <http://www.rluk.ac.uk/content/re-skilling-research>.
- Carlson, Jake. "Demystifying the Data Interview: Developing a Foundation for Reference Librarians to Talk with Researchers about Their Data." *Reference Services Review* 40, no. 1 (2012): 7–23. <http://dx.doi.org/10.1108/00907321211203603>.
- Cox, Andrew, Eddy Verbaan, and Barbara Sen. "Upskilling Liaison Librarians for Research Data Management." *Ariadne* 70 (2012). <http://www.ariadne.ac.uk/issue70/cox-et-al>.
- Creamer, Andrew, Myrna Morales, Donna Kafel, Javier Crespo, and Elaine Martin. "A Sample of Research Data Curation and Management Courses." *Journal of eScience Librarianship* 1, no. 2 (2012). <http://dx.doi.org/10.7191/jeslib.2012.1016>.
- Creamer, Andrew, Myrna Morales, Javier Crespo, Donna Kafel, and Elaine Martin. "An Assessment of Needed Competencies to Promote the Data Curation and Management Librarianship of Health Sciences and Science and Technology Librarians in New England." *Journal of eScience Librarianship* 1, no. 1 (2012). <http://dx.doi.org/10.7191/jeslib.2012.1006>.
- Digital Curation Centre. "RDM for Librarians." 2013. <http://www.dcc.ac.uk/training/rdm-librarians>.
- EDINA and Data Library, University of Edinburgh. *MANTRA: Do-It-Yourself Research Data Management Training Kit for Librarians*. Edinburgh: Digital Curation Centre, 2013. <http://datalib.edina.ac.uk/mantra/libtraining.html>.
- Gabridge, Tracy. "The Last Mile: Liaison Roles in Curating Science and Engineering Research Data." *Research Library Issues: A Bimonthly Report from ARL, CNI, and SPARC*, no. 265 (August 2009): 15–21. <http://publications.arl.org/rli265/16>.
- Garritano, Jeremy R., and Jake R. Carlson. "A Subject Librarian's Guide to Collaborating on e-Science Projects." *Issues in Science & Technology Librarianship* 57 (Spring 2009). <http://www.istl.org/09-spring/refereed2.html>.
- Guy, Marieke. "RDM Training for Librarians." *DCC RDM Services Case Studies*. Edinburgh: Digital Curation Centre, 2013. <http://www.dcc.ac.uk/resources/developing-rdm-services/rdm-training-librarians>.
- Lyon, Liz. "The Informatics Transform: Re-Engineering Libraries for the Data Decade." *International Journal of Digital Curation* 7, no. 1 (2012): 126–38. <http://dx.doi.org/10.2218/ijdc.v7i1.220>.
- Pryor, Graham, and Martin Donnelly. "Skilling Up to Do Data: Whose Role, Whose Responsibility, Whose Career?" *International Journal of Digital Curation* 4, no. 2 (2009): 158–70. <http://dx.doi.org/10.2218/ijdc.v4i2.105>.
- Salo, Dorothea. *Syllabus: LIS 855, Digital Curation*. University of Wisconsin-Madison: School of Library and Information Studies, Spring 2012. <http://dsalo.info/wp-content/uploads/2011/05/855syll2012.pdf>.

Tenopir, Carol, Robert Sandusky, Suzie Allard, and Ben Birch. *Academic Librarians and Data Research Services: Preparation and Attitudes*. Paper presented at IFLA World Library and Information Congress, Helsinki, Finland, August 2012. <http://conference.ifla.org/past/ifla78/116-tenopir-en.pdf>.

## RDMS Policies

---

Digital Curation Centre. "UK Institutional Data Policies." 2013. <http://www.dcc.ac.uk/resources/policy-and-legal/institutional-data-policies/uk-institutional-data-policies>.

Hall, Nathan, Bill Corey, Wendy Mann, and Tom Wilson. *Model Language for Research Data Management Policies*. ASERL / SURA Research Data Coordinating Committee, 2013. [http://www.aserl.org/wp-content/uploads/2013/01/ASERL-SURA\\_Model\\_Language\\_RDM\\_Policy\\_Language\\_FINAL.pdf](http://www.aserl.org/wp-content/uploads/2013/01/ASERL-SURA_Model_Language_RDM_Policy_Language_FINAL.pdf).

## Scanning and Assessment of the Landscape for RDMS

---

Beitz, Anthony, Kheeran Dharmawardena, and Sam Searle. 2012. *Research Data Management Strategy and Strategic Plan 2012–2015*. University of Monash, 2012. <https://confluence-vre.its.monash.edu.au/display/rdmstrategy/Research+Data+Management+Strategy+and+Strategic+Plan+2012-2015>.

Brandt, Scott. *Learning Hands-on and by Trial & Error with Data Curation Profiles*. Paper presented at Digital Curator Vocational Education Europe (DigCurV) Project International Conference, Florence, Italy, May 2013. <http://www.digcur-education.org/eng/International-Conference/Programme/Learning-Hands-on-and-by-Trial-Error-with-Data-Curation-Profiles>.

Jones, Sarah, Alexander Ball, and Çuna Ekmekcioglu. "The Data Audit Framework: A First Step in the Data Management Challenge." *International Journal of Digital Curation* 3, no. 2 (2008): 112–20. <http://dx.doi.org/10.2218/ijdc.v3i2.62>.

Peters, Christie, and Anita Riley Dryden. "Assessing the Academic Library's Role in Campus-Wide Research Data Management: A First Step at the University of Houston." *Science & Technology Libraries* 30, no. 4 (2011): 387–403. <http://dx.doi.org/10.1080/0194262X.2011.626340>.

Provost's Task Force on the Stewardship of Digital Research Data. *Research Data Stewardship at UNC: Recommendations for Scholarly Practice and Leadership*. Chapel Hill, NC: University of North Carolina Libraries, 2012. [http://sils.unc.edu/sites/default/files/general/research/UNC\\_Research\\_Data\\_Stewardship\\_Report.pdf](http://sils.unc.edu/sites/default/files/general/research/UNC_Research_Data_Stewardship_Report.pdf).

Soehner, Catherine, Catherine Steeves, and Jennifer Ward. *E-Science and Data Support Services: A Study of ARL Member Institutions*. Washington, DC: Association of Research Libraries, August 2010. <http://www.arl.org/storage/documents/publications/escience-report-2010.pdf>.

Steinhart, Gail, et al. *Digital Research Data Curation: Overview of Issues, Current Activities, and Opportunities for the Cornell University Library*. Ithaca, NY: Cornell University Library, 2008. <http://hdl.handle.net/1813/10903>.

## Key Papers in the Development of RDMS

---

Association of Research Libraries. "To Stand the Test of Time: Long-term Stewardship of Digital Data Sets in Science and Engineering." *A Report to the National Science Foundation from the ARL Workshop on New Collaborative Relationships: The Role of Academic Libraries in the Digital Data Universe* (September 2006). <http://www.arl.org/storage/documents/publications/digital-data-report-2006.pdf>.

Association of Research Libraries, Joint Task Force on Library Support for E-Science. "Agenda for Developing E-Science in Research Libraries." *Final Report and Recommendations to the Scholarly Communication Steering Committee, the Public Policies Affecting Research Libraries Steering Committee, and the Research, Teaching, and Learning Steering Committee* (November 2007). <http://www.arl.org/storage/documents/publications/escience-report-final-2007.pdf>.

Gold, Anna. "Cyberinfrastructure, Data, and Libraries, Part 1. A Cyberinfrastructure Primer for Librarians." *D-Lib Magazine* 13, no. 9/10 (2007). <http://www.dlib.org/dlib/september07/gold/09gold-pt1.html>.

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

———. "Data Curation and Libraries: Short-Term Developments, Long-Term Prospects." *Office of the Dean, California Polytechnic State University Library* (2010). <http://works.bepress.com/agold01/9>.

"Harnessing the Power of Digital Data for Science & Society." *Report of the Interagency Working Group on Digital Data to the Committee on Science of the National Science and Technology Council* (January 2009). [http://www.nitrd.gov/About/Harnessing\\_Power\\_Web.pdf](http://www.nitrd.gov/About/Harnessing_Power_Web.pdf).

Jahnke, Lori, Andrew Asher, and Spencer D. C Keralis. *The Problem of Data*. Council on Library and Information Resources (CLIR) Report 154 (2012). <http://www.clir.org/pubs/reports/pub154>.

Luce, Richard E. "A New Value Equation Challenge: The Emergence of eResearch and Roles for Research Libraries." *No Brief Candle: Reconceiving Research Libraries for the 21st Century*. Council on Library and Information Resources (CLIR) Report 142 (2008): 42–50. <http://www.clir.org/pubs/reports/pub142/luce.html>.

National Science Board. *Long-lived Digital Data Collections: Enabling Research and Education in the 21st Century*. Washington, DC: National Science Foundation, September 2005. <http://www.nsf.gov/pubs/2005/nsb0540/nsb0540.pdf>.

## ARL/DLF/Duraspace eScience Institute

---

Dobbs, Arta. "E-Science Institute Sponsored by ARL/DLF." 2012. <http://uchc.libguides.com/content.php?pid=228797&sid=1892384>.

Parchuck, Jill, Ed Kairiss, and Michelle Hudson. *Proposed Strategic Agenda for eScience Support at Yale: eScience Institute Capstone Project*. Yale University Library, 2013. <http://guides.library.yale.edu/loader.php?type=d&id=696861>.

Penn State University Libraries. "The ARL/DLF e-Science Institute: Purpose, Activities, Intended Outcomes." 2012. <http://www.libraries.psu.edu/psul/pubcur/esci.html>.

Rankin, Patricia, et al. *Research Data Management at the University of Colorado Boulder: Recommendations in Support of Fostering 21st Century Research Excellence*. University of Colorado Vice Chancellor for Research's Data Management Task Force, 2012. <http://hdl.handle.net/10971/1398>.

Sutton, Shan, David Barber, and Amanda Whitmire. *Oregon State University Libraries and Press Strategic Agenda for Research Data Services*. Oregon State University Libraries, 2013. <http://hdl.handle.net/1957/38794>.

Note: All URLs accessed July 23, 2013.