

Building Blocks:

Laying the Foundation for a Research Data Management Program

Building Blocks: Laying the Foundation for a Research Data Management Program

Ricky Erway

OCLC Research

Laurence Horton

The London School of Economics and Political Science

Amy Nurnberger

Columbia University

Reid Otsuji

University of California, San Diego

Amy Rushing

The University of Texas at San Antonio



© 2016 OCLC Online Computer Library Center, Inc.
This work is licensed under a Creative Commons Attribution 4.0 International License.
<http://creativecommons.org/licenses/by/4.0/>



April 2016

OCLC Research
Dublin, Ohio 43017 USA
www.oclc.org

ISBN: 1-55653-507-4 (978-1-55653-507-9)
OCLC Control Number: 945700114

Please direct correspondence to:

Jackie Dooley
Program Officer
dooleyj@oclc.org

Suggested citation:

Erway, Ricky, Laurence Horton, Amy Nurnberger, Reid Otsuji, and Amy Rushing. 2015. *Building Blocks: Laying the Foundation for a Research Data Management Program*. Dublin, Ohio: OCLC Research.
<http://www.oclc.org/content/dam/research/publications/2016/oclcresearch-data-management-building-blocks-2016.pdf>.

ACKNOWLEDGMENTS

The authors thank the working group members for helping to determine the direction of the report and for providing input at various stages.

- Cunera Buys, Northwestern University
- Ricc Ferrante, The Smithsonian Institution
- Donna Gibson, Memorial Sloan Kettering Cancer Center
- Ryan Johnson, University of California, San Diego
- Jenna Ryan, Boston University
- Stephanie Wright, University of Washington

Also thanks to Jackie Dooley, OCLC Research, who served as a subject-matter expert through the final stages of publication after Ricky Erway retired at the end of 2015.

CONTENTS

Introduction.....	5
Part 1: Laying the Foundation.....	5
Needs Assessment.....	5
Information Provision.....	6
Training.....	6
Data Management Planning.....	7
Active Data Management.....	7
Metadata.....	8
Rights and Privacy.....	9
Data Publication, Preservation and Archiving.....	9
Awareness, Promotion and Outreach.....	10
Part 2: Building Up and Out.....	11
Needs Assessment.....	11
Institution.....	11
Researcher.....	12
Defining Policies, Guidelines and Strategies.....	12
Service Provision.....	13
Training.....	13
Data Management Planning.....	14
Active Data Management.....	16
Metadata.....	17
Data Publication, Preservation and Access.....	18
Awareness, Promotion and Outreach.....	20
Community Engagement and Awareness.....	20
Online Presence.....	21
Partnership Development.....	21
Conclusion.....	22
Other Resources.....	22

INTRODUCTION

Many public funders require that grant recipients make data publically accessible, which makes good sense: research data is a substantial university asset. Often researchers and university administrators look to the library to support these needs.

Sometimes the library proactively seeks to fill the need. In many cases no additional funds or personnel are available for the development and provision of data support services, so the library has to start small, establishing a foundation on which to build.

This document is intended for those who are just beginning to offer data services to researchers at their universities. Part 1 assumes that very little, if anything, is in place, and that resources are limited. It seeks to guide the individual who has data management program responsibilities in directions that will lay a very basic foundation. Part 2 helps identify steps for building on that foundation as needs become evident and as resources allow.

Part 1: Laying the Foundation

This section outlines the first steps for building a research data management (RDM) program from the ground up. It is intended as a low-barrier entry point for library practitioners who have been tasked with the job of providing RDM support and services with limited staff and resources. It includes recommendations on how to get started, issues to consider and existing resources you can leverage.

NEEDS ASSESSMENT

It can be daunting to confront the amount of time and resources that are necessary to conduct a complete needs assessment. Fortunately, not everyone has to do so. A quick and easy starting point is to look at the needs identified by other institutions. Once you have read a few assessments that were done elsewhere, you'll develop a sense of what rings true for your institution. Take advantage of what others have learned until you can learn about your own campus's needs. Your needs assessment can then focus on a particular area, obviating the need to undertake an all-in-one study. You may decide you don't need to conduct a full needs assessment at all.

Examples

- Reznik-Zellen, Rebecca, Jessica Adamick, and Stephen McGinty. 2012. "Tiers of Research Data Support Services." *Journal of eScience Librarianship*. 1(1): e1002. doi:10.7191/jeslib.2012.1002.
- Rolando, Lizzy, Chris Doty, Wendy Hagenmaier, Alison Valk, and Susan Wells Parnham. 2013. *Institutional Readiness for Data Stewardship: Findings and Recommendations from the Research Data Assessment*. Atlanta, Georgia: Georgia Institute of Technology. <http://hdl.handle.net/1853/48188>.
- Weller, Travis, and Amalia Monroe-Gulick. 2015. "Differences in the Data Practices, Challenges, and Future Needs of Graduate Students and Faculty Members." *Journal of eScience Librarianship*. 4(1): e1070. doi:10.7191/jeslib.2015.1070.

- Williams, Sarah. 2013. "Gathering Feedback from Early-Career Faculty: Speaking with and Surveying Agricultural Faculty Members about Research Data." *Journal of eScience Librarianship*. 2 (2): e1048. doi:10.7191/jeslib.2013.1048.
- Wolf, Alan, Mike Simpson, Dorothea Salo, Doug Flee, and Bruce Barton. 2009. *Summary Report of the Research Data Management Study Group*. Madison, Wisconsin: University of Wisconsin–Madison. <https://minds.wisconsin.edu/handle/1793/34859>.

INFORMATION PROVISION

The cornerstone of a foundational service is to provide information and advice. A raft of resources exist that you can consult to build your knowledge and from which you can select the resources you'll recommend to researchers.

Create a web presence for data services, either by adding a data management page to the library's website, by creating library research guides (e.g., LibGuides.com or SubjectsPlus.com), or by using a low-barrier web publishing platform, such as WordPress.

Resources

- Library Success: A Best Practices Wiki. "Subject Guides." http://www.libsuccess.org/Subject_Guides.
- WordPress.com. "Getting Started." <https://learn.wordpress.com/get-started/>.

Training

To help researchers become familiar with RDM concepts and develop skills, include links to existing training resources on your Website.

Examples

- Coursera. "Research Data Management and Sharing." <https://www.coursera.org/learn/data-management>.
- e-Science Portal for New England Librarians. "DIL [Data Information Literacy] Course Materials." <http://esciencelibrary.umassmed.edu/data-literacy/course-materials>.
- Massachusetts Institute of Technology. "Data Management Workshops." <http://libraries.mit.edu/data-management/services/workshops/>.
- The University of Edinburgh. "MANTRA: Research Data Management Training." <http://datalib.edina.ac.uk/mantra/>.
- University of Minnesota. "Training for Researchers and Students: Online Data Management Course." <https://www.lib.umn.edu/datamanagement/workshops>.
- Zenodo. "Research Data Management (RDM) Training Materials." <https://zenodo.org/collection/user-dcc-rdm-training-materials>.

Use elements and ideas from some of these resources to design a short presentation on RDM essentials that can be presented to departments or at sessions open to interested researchers. Consider basing the session on the data requirements of your institution's primary funder. Although most funder requirements address the same topics, researchers are more likely to be interested in the most relevant specifics.

Data Management Planning

Consult your research or grant support office to build a profile of the institution's research funding so you can concentrate on funder requirements that are relevant to your researcher community.

Researchers may have questions about how to create a data management plan (DMP). They will first need to become familiar with their particular funding agency's requirements for a plan.

Example

- Digital Curation Centre (DCC). "Overview of Funders' Data Policies."
<http://www.dcc.ac.uk/resources/policy-and-legal/overview-funders-data-policies>.
- University of Minnesota. "Funding Agency Guidelines."
<https://www.lib.umn.edu/datamanagement/funding>.
- Whitmire, Amanda, Kristin Briney, Amy Nurnberger, Margaret Henderson, Thea Atwood, Margaret Janz, Wendy Kozlowski, Sherry Lake, Micah Vandegrift, and Lisa Zilinski. 2015. *A Table Summarizing the Federal Public Access Policies Resulting from the US Office of Science and Technology Policy Memorandum of February 2013*. Version 4. figshare: November 2015. doi:10.6084/m9.figshare.1372041.

There is no need to develop your own guidance for creating DMPs. Many DMP authoring tools exist and are widely used.

Resources

- California Digital Library, University of California Curation Center. "DMPTool: Data Management Planning Tool" <https://dmptool.org/> (for US researchers).
- DCC "DMPonline." <https://dmponline.dcc.ac.uk/> (for UK/Europe researchers).
- Interdisciplinary Earth Data Alliance. "Data Management Plan Tool v.2."
<http://www.iedadata.org/compliance/plan>.

Active Data Management

The data management process doesn't start when a research project has ended; it should begin when the work is being proposed and designed. Researchers will need to store data while they are actively using it during the course of the research, and they may need to share it with collaborators. Provide examples of best practices for managing data during the research process.

Examples

- Briney, Kristin. 2015. *Data Management for Researchers: Organize, Maintain and Share your Data for Research Success*. Exeter, UK: Pelagic Publishing.

- California Digital Library. "Data Management General Guidance." https://dmptool.org/dm_guidance.
- Corti, Louise, Veerle van den Eynden, Libby Bishop, and Matthew Woollard. 2014. *Managing and Sharing Research Data: A Guide to Good Practice*. London: Sage Publications, Ltd.
- DMPTool. "Data Management General Guidance." https://dmptool.org/dm_guidance.
- Stanford University. "Data Best Practices." <http://library.stanford.edu/research/data-management-services/data-best-practices>.
- University of California (UC) San Diego. "Data Management: Follow Best Practices." <http://libraries.ucsd.edu/services/data-curation/data-management/best-practices.html>.

During the active stages of the research project, researchers will need space to safely and securely store their data. Refer to campus/organizational resources such as institutional information technology services, cloud services or regional/national data centers.

Resource

- Boston University. "Information Security: How to Safely Store Your Data in the Cloud." <http://www.bu.edu/infosec/howtos/how-to-safely-store-your-data-in-the-cloud/>.

Metadata

Datasets are useful only when they can be understood. Encourage researchers to provide structured information about their data, providing context and meaning and allowing others to find, use and properly cite the data. At minimum, advise researchers to clearly tell the story of how they gathered and used the data and for what purpose. This information is best placed in a *readme.txt* file that includes project information and project-level metadata, as well as metadata about the data itself (e.g., file names, file formats and software used, title, author, date, funder, copyright holder, description, keywords, observation unit, kind of data, type of data and language).

Point researchers to existing metadata standards and guidance.

Resources

- Cornell University. "Guide to Writing 'readme' Style Metadata." <http://data.research.cornell.edu/content/readme>.
- DataCite. <https://www.datacite.org/>.
- DCC. "Disciplinary Metadata." <http://www.dcc.ac.uk/resources/metadata-standards>.
- DCC. "How to Cite Datasets and Link to Publications." <http://www.dcc.ac.uk/resources/how-guides/cite-datasets>.

Encourage researchers to obtain unique identifiers for both themselves as creators and for the digital objects they create. These identifiers can be obtained from several sources that have a demonstrated commitment to persistence. Including identifiers in the metadata will provide stable, long-term references to both creators' names and the data itself, thereby allowing both to be cited reliably and authoritatively.

Refer to best practices for persistent researcher identifiers and determine how target repositories assign persistent digital object identifiers for datasets.

Resources

- The DOI® (Digital Object Identifier) System. <http://www.doi.org/>.
- International Standard Name Identifier (ISNI). <http://www.isni.org/>.
- Open Researcher and Contributor ID (ORCID). <http://orcid.org/>.

Rights and Privacy

Make researchers aware of the privacy, intellectual property, copyright and licensing issues that pertain to sharing data; they are ultimately responsible for ensuring that the data are used legally and ethically. Refer to your institution's research compliance or research integrity office for Institutional Review Board (IRB) information, policies on data retention and ownership, and conflict of interest issues.

Example

- University of Virginia. "Research Data Services: Copyright & Privacy." <http://data.library.virginia.edu/data-management/plan/privacy/>.

Point to appropriate licenses for research data and software, and encourage researchers to archive their materials with licenses that permit reuse.

Resources

- DCC. "How to License Research Data." <http://www.dcc.ac.uk/resources/how-guides/license-research-data>.
- Open Knowledge. "Open Definition: Conformant Licenses." <http://opendefinition.org/licenses/>.
- Opensource.org. "Open Source Initiative: Licenses & Standards." <https://opensource.org/licenses>.

Data Publication, Preservation and Archiving

At the end of the active stage of research, researchers may want to archive their data for long-term preservation, access and reuse. Refer them to a discipline-specific repository, if one exists, or to your institutional repository, if it accepts data. Few external repositories commit to preservation, so be sure to examine their terms of service.

Resources

- Open Access Directory. "Disciplinary Repositories." http://oad.simmons.edu/oadwiki/Disciplinary_repositories.
- Registry of Research Data Repositories (re3data.org). "Browse by Subject." <http://www.re3data.org/browse/by-subject/>.

If no discipline-specific or institutional repository exists, investigate other open data repositories.

Resources

- The Dataverse Project. “An Open Source Research Data Repository Framework.” <http://dataverse.org/>.
- Dryad. <http://datadryad.org/>.
- figshare. <http://figshare.com/>.
- Zenodo. <https://zenodo.org/>.

AWARENESS, PROMOTION AND OUTREACH

It is important to make researchers aware of the importance of RDM and to promote the services and resources that you are offering.

- Find ways to promote your Web presence across campus.
- Work with your office of research to raise awareness of RDM support in the institution.
- Offer to become a point of contact and referral for RDM queries.
- Request invitations to departmental meetings to present on RDM essentials and to explain how the library can assist researchers in meeting their obligations.
- Make contact with new researchers when their employment begins.

Develop internal partnerships with likely stakeholders, such as high-performance computing centers, information technology resources, library information technology staff and the sponsored projects office.

Resources

- Erway, Ricky. 2014. *Starting the Conversation: University-wide Research Data Management Policy*. Dublin, Ohio: OCLC Research. <http://www.oclc.org/research/publications/library/2013/2013-08r.html>.

Allocate some time for keeping your knowledge about RDM up to date.

- Bailey, Charles W. 2015. *Research Data Curation Bibliography*. Houston: Digital Scholarship. <http://digital-scholarship.org/rdcb/rdcb.htm>.
- DataQ: A Collaborative Platform For Research Data Questions in Academic Libraries. <http://researchdataq.org>.
- Roy Rosenzweig Center for History and New Media. “Zotero: Research Data Management Services.” https://www.zotero.org/groups/research_data_management_services.

Journals

- *Journal of eScience Librarianship*. <http://escholarship.umassmed.edu/jeslib/>.

- *IASSIST (International Association for Social Science Information Services & Technology) Quarterly.* <http://www.iassistdata.org/iq>.
- *International Journal of Digital Curation.* <http://www.ijdc.net/index.php/ijdc/index>.

Conferences

- International Association for Social Science Information Services & Technology (IASSIST). <http://www.iassistdata.org>.
- International Digital Curation Conference (IDCC). <http://www.dcc.ac.uk/events/international-digital-curation-conference-idcc>.
- Research Data Access & Preservation Summit (RDAP). <https://www.asis.org/rdap/>.
- Research Data Alliance (RDA). “Plenary Meetings.” <https://rd-alliance.org/plenaries>.

Follow these Twitter hashtags: [#datalibs](#), [#OSTPResp](#), [#jisrdm](#).

Part 2: Building Up and Out

After following the steps suggested in Part 1, you will have established a firm, supportive foundation on which you can build. The following elements can then be assembled to create the structure of your RDM program. You may not find it necessary to implement every activity described, nor will you necessarily want to do them in the order listed. Given that you have already learned about your researchers and their needs, you can make good decisions about which steps will be of the most value.

NEEDS ASSESSMENT

You’ve acquired a general understanding of what needs exist at other institutions in terms of data curation and management. Now it’s time to focus on what sets your situation apart, and therefore which specific needs to address.

Institution

What are your institution’s needs relating to policy, infrastructure, data repositories and institutional strategic goals?

Level 1. Conduct an institutional survey on research data needs.

- What internal partnerships would be of benefit? Identify a variety of institutional stakeholders and explore their perspectives in the survey.
- Consider reviewing the particular disciplines emphasized at your institution and their distinct practices and needs regarding research data.

Resources

- DCC. “CARDIO: Collaborative Assessment of Research Data Infrastructure and Objectives.” <http://cardio.dcc.ac.uk/>.

- Jones, Sarah, and Ed Pinsent. 2009. "Using DAF and AIDA to Scope Data Management Needs." Presentation for JISC Data Management Call, Leeds, July 15, 2009. http://www.data-audit.eu/docs/DC101_DAF_AIDA_150709.pdf.

Level 2. Develop program, policy and service recommendations based on survey outcomes, taking into account how service recommendations fit into stakeholders' current workflows.

Researcher

What needs are individuals at your institution expressing? What are the pain points for researchers? These may include areas such as DMP support and guidance, long-term storage, metadata, training, recognition of impact and adhering to funder requirements.

Level 1. Read the data curation profiles for research projects at other universities.

Examples

- "Data Curation Profiles Directory." <http://docs.lib.purdue.edu/dcp/>.
- DCC. "[RDM] Case Studies." <http://www.dcc.ac.uk/resources/case-studies>.

Level 2. Create targeted requests for researcher interviews based on approved DMPs. Set up semi-formal interviews or consultations with individual researchers on a by-request basis using a less intensive form of the data curation profile.

Example

- Boston University. "Research Data Management: Building a Data Curation Profile." <http://www.bu.edu/datamanagement/outline/building-a-data-curation-profile/>.

Level 3. Create full data curation profiles and contribute the results back to the community. Discuss with your sponsored projects office whether profiles should be required as part of the grant proposal process.

Resources

- "Data Curation Profiles." <http://datacurationprofiles.org/>.
- "Data Curation Profiles Directory." <http://docs.lib.purdue.edu/dcp/>.

Level 4. Continue the conversation! Check back with researchers to understand how their research has progressed and whether needs and expectations have changed. Follow up with them after the project has ended to find out which aspects of their DMP worked and didn't work. Would they change their plan for future projects? If so, in what ways?

DEFINING POLICIES, GUIDELINES AND STRATEGIES

Guidelines and policies are key for establishing common commitments and expectations; they will inform all parties of the services offered and how to take advantage of them. Include metadata, selection for retention and access, long-term preservation and sensitive data.

Level 1. Develop an institution-wide strategy for research data curation services.

Resource

- Coates, Heather L. 2014. "Building Data Services From the Ground Up: Strategies and Resources." *Journal of eScience Librarianship*. 3(1): e1063. doi:10.7191/jeslib.2014.1063.

Level 2. Develop RDM policies differentiated by organizational level to align with department, school or center policies.

Resources

- DCC. "UK Institutional Data Policies." <http://www.dcc.ac.uk/resources/policy-and-legal/institutional-data-policies>.
- Erway, Ricky. 2014. *Starting the Conversation: University-wide Research Data Management Policy*. Dublin, Ohio: OCLC Research. <http://www.oclc.org/research/publications/library/2013/2013-08r.html>.

SERVICE PROVISION

Building an RDM program means going beyond provision of the information recommended in Part 1. Here, we urge you to step up to providing services.

Training

Make the library the center for data management training.

Librarians

Level 1. Create talking-point guidance for librarians. Include RDM training resources and support materials for liaison librarians.

Resources

- DataONE. "Librarian Outreach Kit." <https://www.dataone.org/for-librarians>.
- RDMRose. "RDMRose Learning Materials." http://rdmrose.group.shef.ac.uk/?page_id=10.
- The University of Edinburgh. "MANTRA: Do-It-Yourself Research Data Management Training Kit for Librarians." <http://datalib.edina.ac.uk/mantra/libtraining.html>.

Level 2. Build library staff expertise by providing active RDM training for liaison librarians.

Level 3. Maintain and sustain librarian interest and investment by creating an RDM interest group in the libraries. Work with liaison librarians to develop collaborative partnerships with other campus departments.

Researchers

Level 1. Develop an RDM learning support plan to identify target audiences and objectives for the training and implementation. Plan to develop various instructional resources, both online and offline. This aids in developing a comprehensive training plan for RDM.

Resource

- Otsuji, Reid, Dominique Turnbow, Amanda Roth Heath, and Mary Linn Bergstrom. 2015. "Learning to Plan, Planning to Learn!" Poster presented during RDAP Summit 2015, Minneapolis, Minnesota, April 23, 2015.
http://www.slideshare.net/asist_org/learning-toplan-rdapposter4236final.

Level 2. Provide general, in-person RDM training. Develop basic training workshops or materials for writing DMPs, following best practices and managing data. Advertise these sessions as being oriented to requirements of specific funders, even if the content is fairly generic, to promote interest in the training sessions.

Resource

- University of Massachusetts Medical School. "New England Collaborative Data Management Curriculum." <http://library.umassmed.edu/necdmc/modules>.

Level 3. Develop in-person training for discipline-specific RDM issues, working with liaison librarians and researchers in the field. Bring training to the researchers by hosting workshops that target faculty researchers and graduate students in their home departments or labs.

Students

Level 1. Develop a dedicated RDM services website and RDM learning support plan, with online video demos, tutorials and other training materials. Develop DMP self-assessment materials and post them online.

Level 2. Develop curricular materials or guides that can be used in classroom or lab settings to assist instructors in developing a well-defined teaching plan that includes RDM principles.

Resource

- Haverford College. "Project TIER: Teaching Integrity in Empirical Research." <http://www.haverford.edu/TIER/>.

Data Management Planning

Tools and Consultation

Get involved early in the proposal process by providing tools, training and consultation for drafting a DMP. Liaise with your sponsored projects office to ensure that appropriate resources, metadata and RDM practices are cited in proposals for funding.

Level 1. Sign up for institutional use of DMPTool/DMPonline so that researchers can use institutional credentials to log in.

Resource

- DMPTool. “About Partners: Becoming a DMPTool Partner.” <https://dmptool.org/partners#become-partner>.

Level 2. Customize DMPTool/DMPonline templates to give them your institution’s look and feel, to list resources available at your institution and to accommodate any institution-specific policies or preferences.

Level 3. Conduct DMP consultations based on referrals from, or in partnership with, subject librarians, referrals from the sponsored projects office, and researcher requests. Understand how funder policies, journal requirements and professional association expectations affect requirements for data management.

Resources:

- Briney, Kristin, Abigail Goben, and Lisa Zilinski. 2015. “Do You Have an Institutional Data Policy? A Review of the Current Landscape of Library Data Services and Institutional Data Policies.” *Journal of Librarianship and Scholarly Communication*. 3 (2). doi:10.7710/2162-3309.1232.
- Whitmire, Amanda, Kristin Briney, Amy Nurnberger, Margaret Henderson, Thea Atwood, Margaret Janz, Wendy Kozlowski, Sherry Lake, Micah Vandegrift and Lisa Zilinski. 2015. *A Table Summarizing the Federal Public Access Policies Resulting from the US Office of Science and Technology Policy Memorandum of February 2013*. Version 4. figshare: November 2015. doi:10.6084/m9.figshare.1372041.
- DCC. “Overview of Funders’ Data Policies.” <http://www.dcc.ac.uk/resources/policy-and-legal/overview-funders-data-policies>.

Level 4. Perform periodic reviews throughout a researcher’s project to assist in executing and adapting the DMP, culminating in a final review to ascertain the final disposition of the research data.

Data Security

Make sure researchers are aware of the types of security they may need and the resources available to them.

Level 1. Work with your IRB to provide researchers with information about data security policies and compliance regarding personally identifiable information (PII) or protected health information and to provide wording that permits broad data sharing, to the extent that it is responsible and practicable.

Example

- UC Santa Cruz. “Information Technology Services: Personal Identity Information.” <http://its.ucsc.edu/security/pii.html>.

Resources

- Center for Applied Internet Data Analysis (CAIDA). "Data Anonymization." <http://www.caida.org/data/anonymization/index.xml>.
- US Department of Health and Human Services. 2012. *Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule*. Washington, DC: US Department of Health and Human Services. <http://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html>.

Level 2. Obtain institutional policy on copyrights and patents from your sponsored projects office so you'll be able to provide information on intellectual property issues to your researchers.

Example

- Cornell University. "Introduction to Intellectual Property Rights In Data Management." <http://data.research.cornell.edu/content/intellectual-property>.

Level 3. Provide assistance with tools to monitor data validity. Provide pointers to checksum utilities and file format validation tools.

Resources

- "Checksum Verification Tools: Guest Post by Carol Kussmann." 2012. *Practical E-Records* (blog). Posted by Chris Prom (ed.), 6 June 2012. <http://e-records.chrisprom.com/checksum-verification-tools>.
- Prom, Chris. n.d. "Characterizing Files." *Practical E-Records* (blog). <http://e-records.chrisprom.com/resources/software/accessioningingest/identifying-and-characterizing-files/>.

Active Data Management

Assist researchers in managing data during the course of their research and in making it accessible to collaborators.

Level 1. Work closely with all Information Technology (IT) and other stakeholders to ensure that needs for working storage are met. If necessary, set a maximum storage threshold above which fees will be charged.

Examples

- Boston University. "Research Data Management: Storage of Ongoing Data." <http://www.bu.edu/datamanagement/outline/elements/storage/>.
- Texas A&M University. "Scholarly Communication: Data Sharing and Storage Options." <http://scholarlycommunication.library.tamu.edu/repository-getting-started/policies/texas-a-m-university-libraries-data-sharing-and-storage-options1.html>.

Level 2. When needs cannot be met locally, determine what external services might be employed.

Resources

- Anderson, Seth. 2014. *Feet On The Ground: A Practical Approach To The Cloud*. New York, New York: AudioVisual Preservation Solutions, Inc. <https://www.avpreserve.com/wp-content/uploads/2014/02/AssessingCloudStorage.pdf>.
- AudioVisual Preservation Solutions, Inc. 2014. "Cloud Storage Vendor Profiles." <https://www.avpreserve.com/papers-and-presentations/cloud-storage-vendor-profiles/>.
- The London School of Economics and Political Science. 2013. "Using Dropbox and Other Cloud Storage Services." <http://www.lse.ac.uk/intranet/LSEServices/IMT/guides/softwareGuides/other/usingDropboxCloudStorageServices.aspx>.

Level 3. Craft technological solutions to smooth workflow and data transitions through the research life cycle, from collection to analysis, curation, publication and finally preservation and access.

Resources

- Center for Open Science (COS). "Open Science Framework." <https://cos.io/osf/>.
- DataONE. "Best Practices." <https://www.dataone.org/best-practices>.
- DCC. "DCC Curation Lifecycle Model." <http://www.dcc.ac.uk/resources/curation-lifecycle-model>.

Metadata

Information about the dataset helps others find it, understand it, and use it. Help your researchers provide the needed metadata.

Level 1. Provide metadata consultation services that assist with basic metadata description and data documentation.

Resources

- DMPTool. "Data Management General Guidance: Metadata Data Documentation." https://dmptool.org/dm_guidance#metadata.
- DCC. "Disciplinary Metadata." <http://www.dcc.ac.uk/resources/metadata-standards>.
- Research Data Alliance (RDA). "Metadata Directory." <http://rd-alliance.github.io/metadata-directory/standards/>.

Level 2. Identify and disambiguate your researchers' names and datasets.

- a. Establish a relationship with a service that provides persistent identifiers. Help your researchers get permanent identifiers.

Resources

- California Digital Library, University of California Curation Center. “EZID.” <http://ezid.cdlib.org/>.
 - International DOI Foundation. “DOI Registration Agencies.” http://www.doi.org/registration_agencies.html.
- b. Help your researchers obtain persistent identifiers, such as ORCID or ISNI for their own names.

Resources

- ORCID, Inc. “Register for ORCID ID.” <https://orcid.org/register>.
 - ISNI International Agency. “Do you have an ISNI?” <http://www.isni.org/do-you-have-an-isni>.
- c. Require use of a data citation format standard.

Resources

- DataCite. <https://www.datacite.org/>.
- Data Citation Synthesis Group. 2014. *Joint Declaration of Data Citation Principles – Final*. Edited by M. Martone. San Diego, California: FORCE11. <https://www.force11.org/group/joint-declaration-data-citation-principles-final>.
- Economic & Social Research Council (ESRC). 2016. “*Data Citation: What You Need to Know*.” Swindon, UK. http://ukdataservice.ac.uk/media/104397/data_citation_online.pdf.

Level 3. Perform quality control and metadata cleanup on behalf of your researchers.

Data Publication, Preservation and Access

When the research project is complete and the outcomes have been published, the data outputs of the research should be placed in appropriate locations for preservation and access.

Level 1. With the help of the researchers, determine which datasets should be preserved and made accessible, taking in to account the value of the data for replicating and validating the research or for new research, including reuse for others. As necessary, address issues pertaining to sensitive information.

Resources

- DCC. “How to Appraise and Select Research Data for Curation.” <http://www.dcc.ac.uk/resources/how-guides/appraise-select-data>.
- UK Data Service. 2015. *Collections Development Selection and Appraisal Criteria*. Essex, UK: University of Essex. <http://ukdataservice.ac.uk/media/455175/cd234-collections-appraisal.pdf>.

Level 2. Help researchers evaluate external preservation repositories, and help them prepare their data for deposit. Consider establishing formal relationships with one or more external discipline-specific or non-discipline-specific service providers.

Examples

- figshare. <http://figshare.com/>.
- figshare for institutions. <http://figshare.com/services/institutions>.
- Inter-university Consortium for Political and Social Research (ICPSR). “Deposit Data.” <http://www.icpsr.umich.edu/icpsrweb/deposit/>.
- re3data.org. “Browse by Subject.” <http://www.re3data.org/browse/by-subject/>.
- Zenodo. <https://zenodo.org/>.

Level 3. Provide a dark archive service for researchers who have nowhere else to deposit their data. Consider partnering with a network for replication of the data.

Example

- UC San Diego. “Chronopolis.” <https://library.ucsd.edu/chronopolis/index.html>.

Resource

- The Digital Preservation Network (DPN). <http://www.dpn.org/>.

Level 4. Extend and promote your institutional repository to accommodate datasets for preservation and access.

Examples

- Oregon State University. “ScholarsArchive@OSU.” <https://ir.library.oregonstate.edu/xmlui/>.
- Indiana University. “Knowledge Base. What is IUScholarWorks?” <https://kb.iu.edu/d/aujq>.

Resource

- EPrints Repository Software. “ReCollect.” <http://wiki.eprints.org/w/ReCollect>.

Level 5. Provide a dedicated data repository for preservation of and access to your institution’s researchers’ datasets. Consider becoming part of a data network.

Examples

- Johns Hopkins University Data Management Services. “Johns Hopkins Data Archive Dataverse Network.” <https://archive.data.jhu.edu/dvn/>.

- University of Nebraska-Lincoln. “University of Nebraska-Lincoln Data Repository.” <https://dataregistry.unl.edu/>.
- University of North Texas. “UNT Digital Library: UNT Data Repository.” <http://digital.library.unt.edu/explore/collections/UNTDRD/>.

Resources

- Data Archiving and Networked Services (DANS). “Data Seal of Approval.” <http://datasealofapproval.org/en/>.
- DCC. “Data Management and Curation Education and Training.” <http://www.dcc.ac.uk/training/data-management-courses-and-training>.
- Harvard University. “The Dataverse Project.” <http://dataverse.org/>.

Level 6. Create a seamless data-stewardship experience throughout the life cycle, from inception through collection, analysis, description and curation, as well as preservation and access.

Resources

- Foster, Ian, Hakizumwami Birali Runesha, and Vas Vasiliadis. 2013. “*Campus Support For Research Data Management: A Perspective from the University of Chicago*.” Paper presented at Research Data Management Implementations Workshop, Arlington, Virginia, March 13–14, 2013. https://rdmi.uchicago.edu/sites/rdmi.uchicago.edu/files/uploads/Foster,%20I.,%20Runesha,%20H.B.,%20Vasiliadis,%20V_Campus%20support%20for%20reaserach%20data%20management.pdf.
- Mistry, Himanshu, Samantha Guss, and Andy Rutkowski. 2013. “Shared Vision for Data Life-Cycle: Targeting Graduate Students.” Poster presented at Research Data Symposium, Columbia University, New York, New York, February 27, 2013. <http://hdl.handle.net/10022/AC:P:19175>.

AWARENESS, PROMOTION AND OUTREACH

Increase the value of your program by making sure your services are well known and by opening avenues for addressing further needs.

Community Engagement and Awareness

- Level 1.** Create a community group and/or an institutional advisory committee on research data. Develop promotional materials (e.g., free items, postcards, fliers, buttons).
- Level 2.** Increase your presence in individual departments or schools by means such as hosting RDM events, presenting information sessions or getting on the agendas of regular meetings. Encourage library staff to promote services directly to researchers, departments or schools as part of research consultations or presentations. Work with presenters to deposit their data in advance of research events to enable sharing and citing during the event.

Level 3. Communicate the benefits of investing resources in the curation and long-term preservation of research data.

Resources

- Collaboration to Clarify the Costs of Curation (4C Project). 2015. "Investing in Curation: A Shared Path to Sustainability." <http://www.4cproject.eu/roadmap>.
- Charles Beagrie. 2013. "KRDS/I2S2 Digital Preservation Benefit Analysis Tools Project." <http://beagrie.com/krds-i2s2.php>.
- Whyte, Angus, and Jonathan Tedds. 2011. "Making the Case for Research Data Management." *DCC Briefing Papers*, September 2011. <http://www.dcc.ac.uk/resources/briefing-papers/making-case-rdm>.

Online Presence

Level 1. Create and manage a dedicated RDM services website. Keep the content brief, and be sure to link to internal and external policies and support tools.

Level 2. Establish profiles on social media. Use these to follow research activities in your institution, to communicate about your work, to stay up to date on developments at other institutions, and to learn from and contribute to the RDM community.

Partnership Development

Develop and maintain constructive relationships with other organizations to ensure that library systems and services interact with campus systems and services, as well as with external systems and services.

Level 1. Create an advisory committee for research data at the institutional level. Develop library and liaison librarian partnerships with major campus research initiatives.

Level 2. Join external organizations for awareness and coordination. Follow other institutions that are leading in this area, that have a situation similar to yours or are part of a regional group.

Resources

- DPN. <http://www.dpn.org/>.
- IASSIST. <http://iassistdata.org/>.
- RDA. <https://rd-alliance.org/>.

CONCLUSION

This report describes an extensive set of actions and resources that can be leveraged by libraries establishing basic research data management programs. Part 1, “Laying the Foundation,” is intended to help institutions establish the underpinning for the next steps suggested in Part 2, “Building Up and Out.” The order in which any of these steps are implemented should be determined by the needs of your researchers in combination with the capabilities and resources of your library.

Establishing a firm understanding of basic data management needs, providing introductory information and training, and implementing an outreach plan will provide the scaffolding to allow you to continue building your RDM infrastructure. In order for your program to be sustainable, it is important that you assess the needs of the researchers and the institution on an ongoing basis—and be ready to build to the next level, while also, being open to signs that a remodel is called for.

Above all, RDM program development should be guided by strategies appropriate to the needs of the library and the institution. Whether you are responding to the need for initial development of a program or for expansion of an established program, we hope you find these suggested actions and resources helpful in improving your support for your institution’s research mission.

OTHER RESOURCES

- Australian National Data Service (ANDS). “Working with Data: Data Management.” <http://ands.org.au/working-with-data/data-management>.
- DCC. “How to Develop RDM Services – A Guide for HEIs.” <http://www.dcc.ac.uk/resources/how-guides/how-develop-rdm-services>.
- Prior, Graham, Sarah Jones, and Angus Whyte, eds. 2013. *Delivering Research Data Management Services: Fundamentals of Good Practice*. London: Facet Publishing.
- Raboin, Regina, Rebecca C. Reznik-Zellen, and Dorothea Salo. 2013. “Forging New Service Paths: Institutional Approaches to Providing Research Data Management Services.” *Journal of eScience Librarianship*. 1(3): e1021. doi:10.7191/jeslib.2012.1021.
- Strasser, Carly, Robert Cook, William Michener, and Amber Budden. 2012. *Primer on Data Management: What You Always Wanted to Know*. Albuquerque, New Mexico: DataOne. doi:10.5060/D2251G48.

For more information about our work related to research collections and support, please visit:
www.oclc.org/research/themes/research-collections.html



6565 Kilgour Place
Dublin, Ohio 43017-3395

T: 1-800-848-5878

T: +1-614-764-6000

F: +1-614-764-6096

www.oclc.org/research

ISBN: 1-55653-507-4
978-1-55653-507-9
1603/215545, OCLC