OCLC RESEARCH REPORT

The Realities of Research Data Management

PART ONE

A Tour of the Research Data Management (RDM) Service Space

Rebecca Bryant, Brian Lavoie and Constance Malpas

OCLC

The Realities of Research Data Management

Part One: A Tour of the Research Data Management (RDM) Service Space

Rebecca Bryant
Senior Program Officer

Brian Lavoie
Research Scientist

Constance Malpas
Strategic Intelligence Manager & Research Scientist

March 2017

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

ISBN: 978-1-55653-525-3
doi:10.25333/C3PG8J
OCLC Control Number: 979562234

ORCID iDs
Rebecca Bryant https://orcid.org/0000-0002-2753-3881
Brian Lavoie https://orcid.org/0000-0002-7173-8753
Constance Malpas https://orcid.org/0000-0002-9312-8294

Please direct correspondence to:
OCLC Research
oclcresearch@oclc.org

Suggested citation:
Bryant, Rebecca, Brian Lavoie and Constance Malpas. 2017. *A Tour of the Research Data Management (RDM) Service Space*. The Realities of Research Data Management, Part 1. Dublin, Ohio: OCLC Research. doi:10.25333/C3PG8J

CONTENTS

INTRODUCTION

Research data management (RDM) has emerged as an area of keen interest in higher education, leading to considerable investment in services, resources and infrastructure to support researchers' data management needs. This is the first in a series of reports by OCLC Research which examines the context, influences and choices higher education institutions face in building or acquiring RDM capacity—in other words, the infrastructure, services and other resources needed to support emerging data management practices.

Our findings are derived from detailed case studies of four research universities, hailing from four distinct national contexts: the University of Edinburgh (UK), the University of Illinois at Urbana-Champaign (US), Monash University (Australia) and Wageningen University & Research (the Netherlands).

In this introductory report, we provide some brief background on the emergence of RDM as a focus for research support services within higher education; present a simple framework for navigating the contours of the RDM service space; describe the methodology we employed for assembling our findings and discuss the key elements of RDM capacity acquisition these findings address; and offer a preview of the next report in the series.

Background

Research data is fundamental to scholarly inquiry, providing the raw material for empirical investigation and inference.

The scale of research data sets and the technologies used to generate them have changed dramatically over time—compare, for example, the astronomical observations recorded by Tycho Brahe in the sixteenth century with a sextant and quadrant to those of the massive Sloan Digital Sky Survey four hundred years later—culminating in the wide-spread adoption of new, Big Data-driven modes of research. This development, combined with an emerging array of data-intensive computational research techniques in the humanities and social sciences, has changed the face of research data in 21st-century scholarship, and, by extension, the process of assembling, managing and curating research data as well.

Research data is fundamental to scholarly inquiry, providing the raw material for empirical investigation and inference.

In addition to changes in the scale and prevalence of research data itself, other factors have channeled attention to the importance of research data management. The scholarly record is evolving into a deeper, more comprehensive documentation of scholarly activity, extending beyond reporting results in journal articles and monographs to include a range of scholarly outputs generated both during the research process and in the aftermath of publication.[1]

Research data is a prime example: retention and long-term curation of data sets is becoming part of scholarly practice in many disciplines, both to support replication of published findings and to facilitate reuse for new research inquiries.

The benefits of curating research data as part of the scholarly record have been recognized by funders, government agencies and research institutions, who have issued statements of principle, guidelines and in some cases, directives, promoting research data management as a key element of good scholarship. For example, the US National Science Foundation (NSF) requires the inclusion of a data management plan with all grant proposal submissions, noting that "[i]nvestigators are expected to share with other researchers … the primary data, samples, physical collections and other supporting materials created or gathered in the course of work under NSF grants.[2]

In the United Kingdom, Research Councils UK issued a set of principles to guide the development of individual Council data policies, based on the notion that "[p]ublicly funded research data are a public good, produced in the public interest, which should be made openly available with as few restrictions as possible in a timely and responsible manner."[3] And the European Commission, in a 2011 communication, noted that "[i]nformation produced, collected or paid for by public organisations across the European Union is a key resource in the information economy. At the moment, its full potential is far from being realised," at the same time announcing the creation of an Open Data Portal "that will give direct access to a range of datasets from across the EU."[4]

Research universities also have a stake in research data management, as the publication and reuse of important data sets have the potential to add to the scholarly reputation of the institutions from which they originate. Many universities are now interested in documenting the full range of scholarly outputs produced by affiliated faculty and students, including data sets, often utilizing research information management (RIM) systems such as Elsevier Pure or Symplectic Elements to record and describe these outputs.

Tracking the creation and reuse of research data can have practical implications in countries such as the UK where the outcomes of research assessment exercises have material consequences for allocation of public research funding. The Higher Education Funding Council for England (HEFCE), one of four UK funding bodies responsible for carrying out the Research Excellence Framework (REF) assessment program, notes that, [v]ia the REF, we fully recognise data as an equally valid form of research output, and, through our open access policy for the next REF, we plan to reward research environments that deliver open access to a wider set of outputs than just journal articles and conference papers." [5]

As research data becomes the subject of new forms of funder requirements, a key aspect of open science and part of a redefined scholarly record, many institutions have developed RDM service bundles that help their faculty and students manage research data sets effectively both during the research process itself as well as afterwards. These services extend well beyond the provision of persistent storage for data sets, ranging from educational and outreach programs to raise awareness among researchers regarding the importance and basic elements of RDM, to sophisticated data curation services that ensure long-term, persistent identification, understandability and accessibility of research data. To better understand institutional investments in the RDM service space, it is useful to begin with a visualization and description of the general categories of RDM services.

Navigating the RDM Service Space

Research data management services cover a lot of territory, and it is difficult to draw firm boundaries around this service space. In order to trace a rough characterization of RDM's major components, our research team reviewed the RDM service offerings at more than a dozen research universities in North America, Europe and Australia. Our sample included a range of institution types, in a variety of national settings. Our goal was to detect patterns in the array of RDM-related services currently deployed or in development by research institutions. Through this process, we identified three categories of RDM service types that manifested with some regularity across institutions (figure 1), although not all institutions deployed services in all categories, and the specific services offered within categories varied from institution to institution. Despite these differences, the three RDM service categories provide a useful heuristic for visualizing the scope of the RDM service space.

Research Data Management Service Categories

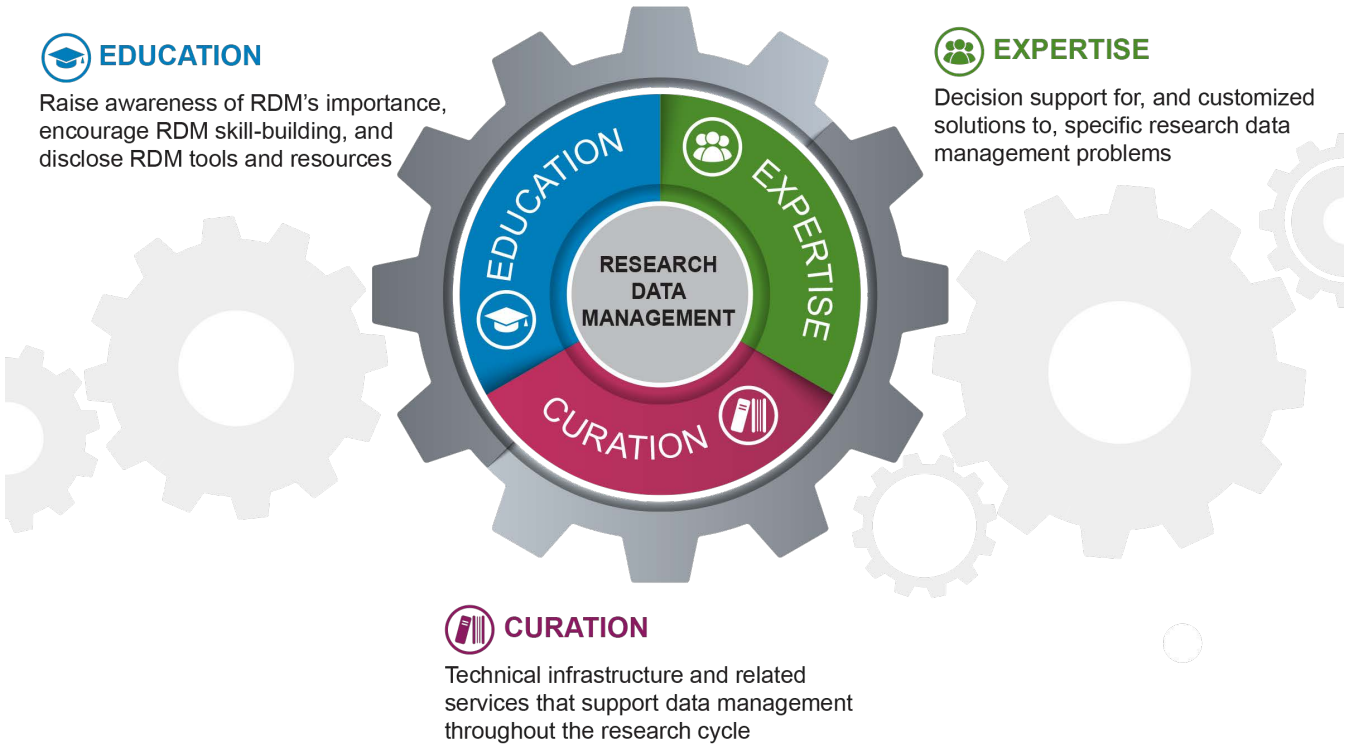

FIGURE 1. RDM SERVICE CATEGORIES

Education services are aimed at educating researchers and other stakeholders on the importance, and in some cases, the necessity, of responsibly managing their data and making arrangements for its long-term curation. A key aspect of these services is to raise awareness of the general scholarly interest in ensuring that research data is available for future use in order to document, replicate and build on published findings. More focused educational resources aimed at acquainting researchers with data management norms and practices associated with particular disciplines may be woven into this broader perspective. Education RDM services also make researchers aware of relevant data management policies and requirements imposed by funders, national or international agencies, and even the researcher's own institution.

Another important aspect of Education RDM services is to acquaint researchers with the rudiments of good data management practices. This might include tools and advice on crafting effective data management plans, guidelines for creating descriptive metadata to promote discovery of archived data sets, and workshops or training sessions aimed at RDM skill-building. Curated lists of RDM resources (e.g., discipline-

specific repositories, freely available data management tools, informational tutorials) are particularly important here. These lists may be of general interest, or may be tailored to the needs of particular research or disciplinary areas.

Although the most important impetus for good data management practices might be compliance with funder requirements, Education RDM services sometimes point out additional incentives for researchers to responsibly manage their data. These might include better documentation and management of the research process, and enhanced reputations for both the researcher and the institution for sharing important data sets for reuse.

Generally speaking, Education RDM services serve to *raise awareness* on the part of researchers regarding: the importance of good data management both in service to open science and to meet compliance obligations; promulgating basic data management practices and skills; the wide array of internally and externally provided RDM resources available; and articulating key incentives for researchers to invest the attention needed to secure their data.

Expertise services provide decision support and customized solutions for researchers working through specific research data management problems. Such services may include “helpline” resources to which RDM-related questions may be directed; direct consultation with data or liaison librarians; and customized RDM support services, such as metadata creation, data preparation and mediated deposit.

Another example of Expertise-related RDM services are training programs for internal staff with current or future responsibilities for supporting data management at the institution. Often termed as “train-the-trainer” programs, these initiatives focus on creating the expertise needed to support the kinds of services described above.

Two distinguishing features of Expertise services are worth noting. First, such services tend to be based on direct, person-to-person interaction between data management experts and the researcher, rather than unmediated resources such as LibGuides or online, self-paced tutorials. Second, Expertise services tend to operate in parallel with the research process itself, or put another way, they are consumed by researchers at various stages of the research cycle. This is in contrast, for example, with Education services, which can be consumed independently of any particular research process.

In sum, Expertise-related RDM services are *human-mediated capacities* geared toward solving specific data management problems encountered by researchers *during the research process*. Moving beyond the awareness-raising function of Education services, Expertise-related services deploy the specialized knowledge of data librarians, technologists and other support staff to ensure that individual RDM needs and requirements are met.

Curation services supply technical infrastructure and related services that support data management throughout the research cycle. This includes active data management (managing data during the research process) as well as long-term data stewardship (care of and access to the data after the research activity has concluded).

RDM Curation services cover a range of functions, including persistent storage, assignment of unique identifiers, access controls, metadata creation and management, versioning, and long-term preservation. They

can address various stages of the research cycle, beginning with near-term support (coinciding with active use of the data in the research process); extending to medium-term support (covering a finite period—e.g., three to five years—following the conclusion of the research process); and potentially extending still further to long-term support (continuing for an extended period after deposit).

Curation of research data sets can require significant infrastructure investment, often for dedicated RDM systems distinct from other campus repository services such as the institutional repository. Even if local infrastructure outlays are reduced by using externally provided data curation resources, difficult challenges may arise in developing effective "campus bridging" strategies that link distributed curation services into a cohesive workflow.[6]

While RDM Curation services are infrastructure-intensive, policy is an equally important element in implementing these services. Retention policy is a good example of the complexities and competing interests that can be at play: will all data sets be accepted for deposit, or will an appraisal process be implemented? Will deposited data sets be retained indefinitely, or will their retention be reviewed after a prescribed period? What criteria might lead to de-accession? Do local retention policies ensure compliance with requirements imposed by funders and other external agencies? Other important RDM policy areas include, but are not limited to, metadata requirements, access restrictions and privacy assurances for sensitive data.

In sum, RDM Curation services offer the *technical functionality* needed to manage data sets throughout the research life cycle, supported by local and/or distributed curation infrastructure. It is important to note that these functionalities are often scoped by local policy choices, with the consequence that the nature of RDM Curation services will differ from institution to institution.

While mutually reinforcing, the Education, Expertise and Curation service categories are, from an implementation standpoint, separable and may be deployed independently or in combination. For example, a university with a relatively low research intensity may find that the provision of RDM Education services alone adequately meets the needs of its faculty and students. In contrast, a research-intensive university may choose to offer services in all three categories. In light of this, it is important to emphasize that full coverage of all three components of our RDM framework should not necessarily be the goal of every institution setting out to acquire RDM capacity; nor should the lack of services in one or more categories be construed as a mark against the quality of RDM support at a particular institution.

The Education-Expertise-Curation (E-E-C) framework depicted in figure 1 represents a view of the *potential* extent of RDM services. There is no reason to believe that the mix of RDM service categories, or the selection of individual services within each category, will be similar in institutions with different research profiles. Indeed, choosing which RDM services are important to an institutional context should be the first decision when considering RDM capacity acquisition.

… it is important to emphasize that full coverage of all three components of our RDM framework should not necessarily be the goal of every institution setting out to acquire RDM capacity; nor should the lack of services in one or more categories be construed as a mark against the quality of RDM support at a particular institution.

Exploring RDM Capacity Acquisition: Four Case Studies

Because the functional requirements for RDM capacity, and the optimal strategy for acquiring that capacity, will differ across institutions, there is no single exemplar for RDM capacity acquisition that can serve as a template for others to follow. Nevertheless, the *aggregated* experiences of research universities that have deployed various forms of RDM capacity provide useful markers, lessons and decision points for other institutional contexts.

The importance of leveraging the RDM capacity-building experiences of institutions that have entered this service space has been marked in the literature by a number of studies based on survey approaches. The work of Carol Tenopir and colleagues,[7] as well as the Association of Research Libraries SPEC Kit on research data management services,[8] are leading examples of work of this kind. Survey-based data on the development of RDM services in higher education institutions are useful sources for taking the temperature of the current state of RDM in the community at large.

Survey-based approaches are complemented by detailed case studies that probe deeply into the specific institutional context surrounding the development of RDM services. Neil Rambo's description of New York University Health Sciences Library's experiences in developing RDM capacity is a noteworthy example;[9] another is Gary Brewerton's account of Loughborough University's partnership with Figshare and Arkivum to build a platform for RDM services.[10]

As with the survey-based literature, important lessons can be abstracted from case studies in RDM development. But the value of these case studies increases with the number available: in this way, the details of the individual case studies are complemented by a diversity of circumstances and contexts from which to draw them. In our view, more case studies of RDM capacity acquisition are needed, and for this reason, we chose a case study approach for our study.

In our study, the acquisition of RDM capacity is treated as an institutional problem; as such, we focus on the response of the institution in its entirety to the need for RDM capacity, rather than the response of a particular campus unit in isolation.

Our analysis is based on detailed case studies of the RDM capacity acquisition efforts of four research universities: the University of Edinburgh in Scotland; the University of Illinois at Urbana-Champaign in the US; Monash University in Australia; and Wageningen University & Research in the Netherlands. Data for the study was gathered through desk research and, most importantly, a series of interviews with key staff from each of the universities participating in the study. Our purpose is to construct four in-depth pictures of the process of RDM capacity acquisition, set in four distinct institutional and national contexts.

In our study, the acquisition of RDM capacity is treated as an institutional problem; as such, we focus on the response of the institution in its entirety to the need for RDM capacity, rather than the response of a particular campus unit in isolation. However, we do place special emphasis on the role of the academic library in supporting institutional RDM goals, given that this is a topic of particular interest to our organization and its members. Findings from our analysis will be presented in a series of short reports, each dealing with a different aspect of RDM capacity acquisition:

Findings from our analysis will be presented in a series of short reports, each dealing with a different aspect of RDM capacity acquisition:

Report 1	**Introduction:** This report provides brief background on the emergence of RDM within research universities, presents the E-E-C framework for navigating RDM services, describes our methodology, and previews the series.
Report 2	**Scope:** As figure 1 suggests, the RDM service space encompasses a wide array of services, running the gamut from awareness-raising outreach to technical data curation processes. Moreover, research data management as a service category has yet to mature; it is still in a very dynamic state and much uncertainty remains over future developments. It is not enough to decide to acquire RDM capacity; institutions must also determine what subset of the service space is most pertinent to meeting current and future RDM requirements. This second report in the series will consider the scope of the RDM capacity of the four research universities, examining each institution's coverage of the E-E-C categories, and their selection of specific services within each cluster.
Report 3	**Incentives:** Fundamental to any capacity acquisition process are the incentives that prompt decision-makers to take action. Two interrelated issues of particular importance are: the nature of the institutional problem that the acquisition of RDM capacity is perceived to solve; and the internal and external sources of pressure, petitioning and other inducements that motivate action to address that problem. This third report in the series will explore the incentives that inspired the acquisition of RDM capacity on the part of the four research universities described in our case studies, and seek both the general patterns and context-dependent circumstances that shaped these incentives.
Report 4	**Sourcing and scaling:** Another key facet of RDM capacity acquisition is determining where that capacity will be sourced: will it be built and maintained internally, or will it be acquired from an external provider? Similarly, the question of scale must be addressed: should RDM solutions be deployed as institution-scale services, or should they be organized at scales above the institution? Our final report of the series will examine the sourcing and scaling choices made by the four research universities in regard to their acquisition of RDM capacity. What variables are being maximized in these decisions: efficiency/cost reduction? Reliability/trust? The choices that individual institutions make with respect to sourcing and scaling likely depend on a host of factors, including local staffing and infrastructure, availability of cooperatively sourced or nationally provisioned services and willingness to pay for commercially sourced solutions.

While other aspects of RDM capacity acquisition are also important, our focus on incentives, scope and sourcing/scaling rests on the belief that they represent fundamental decision points in the path toward meeting institutional RDM needs: deciding to act (incentives), deciding what to do (scope), and deciding how to do it (sourcing and scaling).

Our case studies are intended to draw out points of convergence and divergence in how these choices manifest themselves, and are subsequently addressed and resolved, in four distinct institutional settings.

The results of our study are not intended to offer a comprehensive picture of RDM capacity choices among research institutions, nor to support grand generalizations on optimal RDM acquisition strategies. Rather, we aim to provide a detailed look at how four institutions, operating in four different institutional and national contexts, are acquiring RDM capacity to meet institutional needs in this area. We are hopeful that readers will see something of their own institutional context in these case studies, and benefit accordingly in thinking about their local RDM capacity acquisition choices.

A Story Unfolds: Building on Earlier OCLC Research Studies

The issue of RDM capacity acquisition, and research data management generally, does not exist in isolation. As discussed above, the emergence of RDM as an area of interest for research universities and other stakeholders traces back, in part, to trends that are re-configuring established patterns of research practice and scholarly communication. OCLC Research has been exploring these trends—and their implications—with a program of work focused on the evolution of the scholarly record in the 21st century. Our report, *The Evolving Scholarly Record* (ESR), characterized general trends and influences that are reshaping the nature and scope of the scholarly record.[11] This report was followed by *Stewardship of the Evolving Scholarly Record: From the Invisible Hand to Conscious Coordination*, which discussed how re-definition of the scholarly record would likely lead to parallel evolutions in stewardship models.[12]

Our analysis of RDM capacity acquisition is an extension of this earlier work, focusing the general insights and findings from the earlier ESR reports onto a specific stewardship domain: research data management. Research data is arguably the most visible aspect of the evolution of the scholarly record, presenting both challenges and opportunities for academic libraries, funders, publishers and other stakeholders.

… our focus on incentives, scope and sourcing/scaling rests on the belief that they represent fundamental decision points in the path toward meeting institutional RDM needs: deciding to act (incentives), deciding what to do (scope), and deciding how to do it (sourcing and scaling).

RDM capacity acquisition also connects with other OCLC Research work examining other aspects of research data management. This work includes practical steps toward establishing RDM policies and services on campus;[13] sustainability strategies for research data and data repositories;[14] and data reuse.[15] Research data management is also an important element of new OCLC Research work in the area of research information management (RIM): “the integrated management of information about the research lifecycle, and about the entities which are party to it.”[16] In short, RDM is an important domain for libraries and other stakeholders, and

this is reflected in the numerous OCLC Research studies that directly or indirectly address it. Our study of RDM capacity acquisition is but the latest addition to OCLC Research's accumulating body of work on research data management. Of course, while OCLC Research has addressed many aspects of RDM, it is important to acknowledge that this is still a very dynamic, fluid domain; there is much of the RDM story still to be told.

A Preview of the Next Report: Scope

As we have noted, research data is an issue of growing importance for research universities and others interested in research practice, scholarly communication and the scholarly record. In response, RDM services are emerging to meet institutional needs in this area. Although RDM has coalesced into a recognizable set of services, the borders of the RDM service space are still fuzzy, and specific implementations vary from institution to institution. The E-E-C framework discussed above is intended as a simplified, high-level view of this space, acknowledging the variations in service implementation existing within these categories.

The E-E-C framework is orthogonal to the topics which will follow in this series—scope, incentives, sourcing/scaling—and will be used to frame the discussion in each report. The next report will describe the scope of RDM services deployed at each of the four research universities participating in our study. In particular, we will consider:

- What segment of the space defined by the E-E-C framework is deployed at each university? How are these services implemented (i.e., features/functionality)?
- Does the institutional RDM service offering emphasize one or more of the E-E-C categories? If so, why?
- What is the relationship between the scope of the local institutional RDM service offering and external alternatives? Are local RDM services offered in parallel with, or complementary to, external RDM services?

…while OCLC Research has addressed many aspects of RDM, it is important to acknowledge that this is still a very dynamic, fluid domain; there is much of the RDM story still to be told.

In short, our next report will focus on the nature of the RDM capacity that has been acquired by each of the four research universities, and highlight, where possible, the key decision points that shaped the contours of this capacity.

Conclusion

A paper published in the *American Economic Review* (AER) titled "Replication in Empirical Economics: The Journal of Money, Credit, and Banking Project," reported findings from a study in which the authors tried to collect data and computer programs from economists whose work had been recently published, accepted for publication, or was currently under editorial review for publication in the Journal of Money, Credit, and Banking, and then replicate the reported findings with the resources provided.

The results of the exercise were not encouraging: in many cases, they could not obtain the necessary data and code from researchers, and even when they did, some results were still not reproducible. The authors concluded that "the economics profession needed to improve the replicability of empirical results by persuading

journal editors to secure a copy of data and computer code used by authors of empirical research prior to publication, when their leverage is most effective."

This article appeared thirty years ago.

Today, we see an acknowledgement of research data as a key part of the research and scholarly communication processes, as well as the scholarly record…

Three decades later, the emergence of services dedicated to the collection, curation and ongoing accessibility of research data, as well as outreach aimed at raising awareness about the importance of responsible data management, confirms the soundness (if perhaps belated recognition) of the authors' conclusion.

Today, we see an acknowledgement of research data as a key part of the research and scholarly communication processes, as well as the scholarly record: for replication/reproducibility, as the authors of the AER paper noted, but also for compliance with funder requirements, research assessment, data sharing and reuse and greater openness in documenting the process of scholarly endeavor.

In this climate, the acquisition of RDM capacity will only grow as a matter of practical interest to decision-makers both inside and outside the academy. We hope that this series of reports provides useful insight to those tasked with navigating the process of acquiring RDM capacity.

ACKNOWLEDGMENTS

The authors extend special thanks to our interview informants: Heidi Imker and Beth Sandore Namachchivaya (University of Illinois at Urbana-Champaign); David Groenewegen (Monash University); Dominic Tate and Jeremy Upton (University of Edinburgh); and Jacquelijn Ringersma and Ellen Fest (Wageningen University & Research).

We note with appreciation that the libraries participating in this project include three members of the OCLC Research Library Partnership (Illinois, Monash, and Edinburgh), and the home institution of an OCLC Europe, the Middle East and Africa (EMEA) Regional Council member (Hubert Krekels, Library Director at Wageningen). Thanks are also due to Lorcan Dempsey, Vice President, Membership and Research, Chief Strategist, for his role encouraging further work in the area of research data management and to our OCLC colleagues Erin M. Schadt, Jeanette McNicol and JD Shipengrover for their help in publishing this report.

NOTES

1. Lavoie, Brian, Eric Childress, Ricky Erway, Ixchel Faniel, Constance Malpas, Jennifer Schaffner, and Titia van der Werf. 2014. *The Evolving Scholarly Record*. Dublin, Ohio: OCLC Research. http://www.oclc.org/content/dam/research/publications/library/2014/oclcresearch-evolving-scholarly-record-2014.pdf.
2. National Science Foundation. "Dissemination and Sharing of Research Results." https://www.nsf.gov/bfa/dias/policy/dmp.jsp.
3. Research Councils UK. RCUL "Common Principles on Data Policy." http://www.rcuk.ac.uk/research/datapolicy/.
4. European Union Open Data Portal: http://data.europa.eu/euodp/en/data/; European Commission. 2011. *Open data: An Engine for Innovation, Growth and Transparent Governance.* COM/2011/0882 final. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0882:FIN:EN:PDF.
5. Higher Education Funding Council for Education. "Opening Up Research Data." http://blog.hefce.ac.uk/2015/09/01/opening-up-research-data/.
6. Almes, Guy T., and Ralph J. Zottola. 2014. *Big Data in the Campus Landscape: Basic Infrastructure Support.* p. 6. [Louisville, CO]: Educause Center for Analysis and Research. https://library.educause.edu/~/media/files/library/2014/9/ewg1406-pdf.
7. LIBER. 2016. "Research Data Services in Europe's Academic Research Libraries." http://libereurope.eu/blog/2016/10/13/research-data-services-europes-academic-research-libraries/.
8. Fearon, David Jr., Betsy Gunia, Sherry Lake, Barbara E. Pralle, and Andrew L. Sallans. 2013. *SPEC Kit 334: Research Data Management Services (July 2013).* Washington, DC: Association of Research Libraries. http://publications.arl.org/Research-Data-Management-Services-SPEC-Kit-334/.
9. Marcum, Deanna. 2015. "New Issue Brief by Neil Rambo Explores Research Data Management." *Ithaka S+R Blog*. Posted 22 October. http://www.sr.ithaka.org/blog/new-issue-brief-by-neil-rambo-explores-research-data-management/.
10. Brewerton, Gary. 2015. "Research Data Management: A Case Study." *ARIADNE: Web Magazine for Information Professionals. (12 October)* http://www.ariadne.ac.uk/issue74/brewerton.
11. Lavoie et al (see n. 1).
12. Lavoie, Brian and Constance Malpas. 2015. *Stewardship of the Evolving Scholarly Record: From the Invisible Hand to Conscious Coordination.* Dublin, Ohio: OCLC Research. http://www.oclc.org/content/dam/research/publications/2015/oclcresearch-esr-stewardship-2015.pdf.
13. Erway, Ricky. 2013. *Starting the Conversation: University-wide Research Data Management Policy*. Dublin, Ohio: OCLC Research. http://www.oclc.org/content/dam/research/publications/library/2013/2013-08.pdf; Erway, Ricky, Laurence Horton, Amy Nurnberger, Reid Otsuji, and Amy Rushing. 2016. *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.

14. Erway, Ricky. 2012. *Lasting Impact: Sustainability of Disciplinary Repositories.* Dublin, Ohio: OCLC Research. http://www.oclc.org/research/publications/library/2012/2012-03.pdf; Erway, Ricky and Brian Lavoie. 2012. "The Economics of Data Integrity." *In Curating for Quality: Ensuring Data Quality to Enable New Science*. Edited by Gary Marchionini, Christopher A. Lee, and Heather Bowden, 53-56. Chapel Hill, NC: University of North Carolina at Chapel Hill. Full report: http://www.oclc.org/research/publications/all/economics-of-data-integrity.html. Chapter preprint: http://www.oclc.org/content/dam/research/publications/library/2012/erway-dataintegrity.pdf; Erway, Ricky, and Amanda Rinehart. 2016. *If You Build It, Will They Fund? Making Research Data Management Sustainable.* Dublin, Ohio: OCLC Research. http://www.oclc.org/content/dam/research/publications/2016/oclcresearch-making-research-data-management-sustainable-2016.pdf; Lavoie, Brian. 2012. "Sustainable Research Data." In *Managing Research Data.* Edited by Graham Pryor. London: Facet Publishing. http://www.oclc.org/research/publications/all/sustainable-research-data.html.

15. Faniel, Ixchel, Eric Kansa, Sarah Whitcher Kansa, Julianna Barrera-Gomez, and Elizabeth Yakel. 2013. "The Challenges of Digging Data: A Study of Context in Archaeological Data Reuse." *JCDL 2013 Proceedings of the 13th ACM/IEEE-CS Joint Conference on Digital Libraries,* 295-304. New York, NY: ACM. doi.org/10.1145/2467696.2467712. Preprint available at http://www.oclc.org/content/dam/research/publications/library/2013/faniel-archae-data.pdf; Faniel, Ixchel M., Adam Kriesberg, and Elizabeth Yakel. 2012. "Data Reuse and Sensemaking among Novice Social Scientists." *ASIS&T 2012 Annual Meeting Proceedings*, 1-10. Baltimore, Maryland, USA. Preprint available online at http://www.oclc.org/content/dam/research/publications/library/2012/faniel-data-reuse-sensemaking.pdf; Kriesberg, Adam, Rebecca D. Frank, Ixchel M. Faniel and Elizabeth Yakel. 2013. "The Role of Data Reuse in the Apprenticeship Process." *ASIS&T 2013 Annual Meeting Proceedings*, Montréal, Québec, Canada. Preprint available online at http://www.oclc.org/content/dam/research/publications/library/2013/faniel-data-reuse-apprenticeship.pdf.

16. Dempsey, Lorcan. 2014. “Research Information Management Systems – A New Service Category?” *Lorcan Dempsey’s Weblog*. Posted 26 October. http://orweblog.oclc.org/research-information-management-systems-a-new-service-category/.

For more information about OCLC Research's work on research data management, please visit: **oc.lc/rdm**

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: 978-1-55653-525-3
DOI: 10.25333/C3PG8J
RM-PR-215833-WWAE 1703