

The Realities of Research Data Management

PART TWO

Scoping the University
RDM Service Bundle

Rebecca Bryant, Brian Lavoie and Constance Malpas

The Realities of Research Data Management

Part Two: Scoping the University RDM Service Bundle

Rebecca Bryant

Senior Program Officer

Brian Lavoie

Research Scientist

Constance Malpas

Strategic Intelligence Manager & Research Scientist



© 2017 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/>



September 2017

OCLC Research

Dublin, Ohio 43017 USA

www.oclc.org

ISBN: 978-1-55653-023-4

doi:10.25333/C3Z039

OCLC Control Number: 1004431329

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. *Scoping the University RDM Service Bundle*. The Realities of Research Data Management, Part 2. Dublin, OH: OCLC Research.

<https://doi.org/10.25333/C3Z039>.

CONTENTS

Introduction.....	5
The RDM Service Space.....	7
Decision Point: Scoping the Local RDM Service Bundle.....	8
Case Studies in Scoping	8
University of Edinburgh	8
Education	10
Expertise	10
Curation	11
Reflections	12
University of Illinois at Urbana-Champaign	13
Education	14
Expertise	15
Curation	15
Reflections	17
Monash University	17
Education	19
Expertise	20
Curation	20
Reflections	22
Wageningen University & Research.....	22
Education	24
Expertise	25
Curation	25
Reflections	26
Scoping the RDM Service Bundle: Takeaways	27
Acknowledgments	31
Notes	32

FIGURES

Figure 1. RDM Service Categories: A Tour of the Research Data Management (RDM) Service Space. The Realities of Research Data Management, Part 1.....	7
Figure 2. RDM Service Bundle Highlights: University of Edinburgh	9
Figure 3. RDM Service Bundle Highlights: University of Illinois at Urbana-Champaign	14
Figure 4. RDM Service Bundle Highlights: Monash University.....	18
Figure 5. RDM Service Bundle Highlights: Wageningen University & Research	23

INTRODUCTION

Research data management (RDM) is both a matter of keen interest and an ongoing challenge in higher education. Proliferating data mandates, open science advocacy, replication of published results, and efforts to thwart academic fraud have all focused attention on data management practices during the research process, and long-term data stewardship practices once the research process has ended. This, in turn, has led to the development of services, infrastructure and other resources to support RDM needs at research universities.

The Realities of Research Data Management is a series of four reports looking at the context, influences and choices research universities face in building or acquiring RDM capacity. 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).

Our first report¹ explored the background surrounding the emergence of RDM and introduced a simple framework for navigating the RDM service space. In this, our second report, we look at the RDM service bundles of the four universities to better understand how the local RDM service bundle is shaped by the complex interplay of internal and external factors, institutional requirements and local choices.

While many universities agree on the importance of RDM in supporting emerging practices and expectations that shape 21st century scholarship, strategies for implementing RDM fork repeatedly from this common departure point. Indeed, an early decision point a university faces in the process of acquiring RDM capacity is to *scope the local RDM service bundle*—to decide what mix of RDM services and resources are best suited to address the needs of local researchers.

RDM is not a monolithic set of services duplicated across universities; it is a customized solution shaped by a range of internal and external factors operating on local decision-making.

In reviewing the experiences of our four case study institutions, a common theme emerged: RDM is not a monolithic set of services duplicated across universities; it is a customized solution shaped by a range of internal and external factors operating on local decision-making.

An important corollary to this finding is that scoping an RDM service bundle sufficient to meet institutional needs does not necessarily mean implementing the full range of services within the RDM service space.

In this report, we explore these findings in the context of our case study partners, examining how each university scoped their local RDM service bundle in light of each institution's broader institutional and external environments. In each case, the choices made—and the resulting RDM service bundle—reflect the particular circumstances of the university in question. While each

university is unique, other institutions may see something of themselves in one or more of the universities we examine, and the case studies may serve as models or patterns to inform local RDM planning in other contexts.

...scoping an RDM service bundle sufficient to meet institutional needs does not necessarily mean implementing the full range of services within the RDM service space.

More generally, the case studies provide insight into how four research universities responded to the emergence of a new scholarly practice—research data management—by scoping a set of services to support local researchers in meeting the requirements of this practice.²

The RDM Service Space

The first report in this series presents a simple framework for thinking about the RDM service space in its entirety (figure 1). The framework divides RDM services into three categories: *Education*, *Expertise* and *Curation*. These categories summarize a wide array of specific services that may be deployed as part of a university's RDM service bundle.

RDM covers a range of complementary yet distinct service categories, with an even wider range of specific services existing within each category. Enumerating all of them does not translate into a checklist of required services a university must deploy in order to build a credible RDM capacity. For example, a university may not feel that offering locally built Curation resources, such as data repository services, is necessary given local circumstances. Nor is the decision necessarily a binary

“offer-or-do-not-offer” one; instead, the decision may be one of striking the appropriate emphasis. For example, in deploying Expertise services, a university may choose to establish a general help-line email account to handle RDM-related inquiries, rather than a more elaborate strategy involving data librarians or discipline-specific liaisons providing face-to-face consultation.

The key point is that while general discussions of RDM and its importance to research universities may reference RDM conceptually as a monolithic block of services, moving from concept to practice requires much more precision and customization. Acquiring RDM capacity involves making choices about the services needed to support local data management needs. In other words, a key decision point in the RDM capacity acquisition process is to scope the local RDM service bundle.

Research Data Management Service Categories

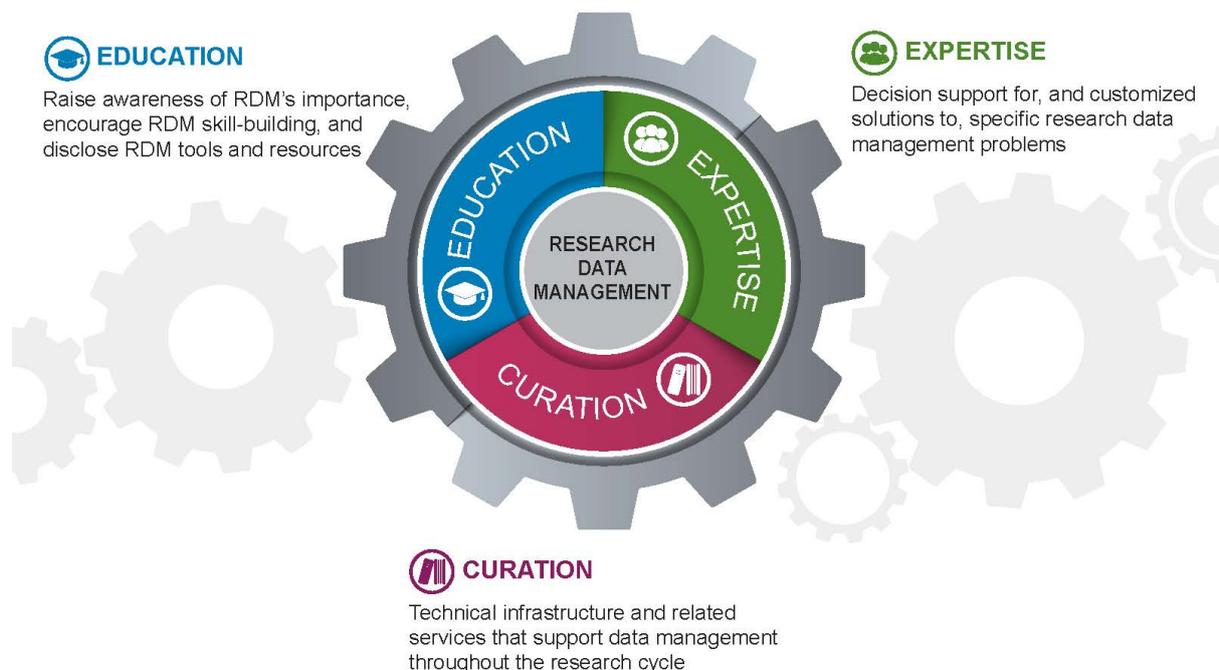


FIGURE 1. RDM SERVICE CATEGORIES: A TOUR OF THE RESEARCH DATA MANAGEMENT (RDM) SERVICE SPACE. THE REALITIES OF RESEARCH DATA MANAGEMENT, PART 1³.

Decision Point: Scoping the Local RDM Service Bundle

Research data management is an emerging challenge both for research universities and the individual researchers affiliated with them. Moreover, for many institutions, the need to respond to the challenge is perceived both immediate and imperative. Noting the changing nature of the scholarly record, CNI Executive Director Clifford Lynch describes RDM as an “enormous ongoing problem.”⁴ *The international New Media Consortium Horizon Report: 2017 Library Edition* cites RDM as a key factor driving technology adoption in academic and research libraries in the next one to two years.⁵ It is not difficult to find testimonials to the importance of research data, both as a means of documenting past research and as a source of raw material for future research, coupled with calls for action. But how does this translate into an institutional RDM service bundle that effectively meets the needs at a particular institution?

From an institutional perspective, the first step in addressing RDM is *deciding to act*—in other words, to take steps to meet the RDM needs prevailing at the university. This is largely a question of responding to incentives, a topic we will take up in our next report.⁶ The second major decision point is deciding *what to do*—in other words, scoping the RDM service bundle. This is what we address in this report.

We introduce the concept of an *RDM service bundle*, which is the range of local RDM services offered by a university to its researchers. An important point to clarify is the boundary of “local.” In its narrowest interpretation, local RDM services can mean those which are built, hosted and deployed locally. Alternatively, it could also include RDM services that are provided externally, for which the university arranges access for affiliated researchers. For the purposes of this report, we choose the latter interpretation.

The fabric of RDM services, infrastructure and other resources extends well beyond the

research university; in many, if not most, cases, the bundle of locally sourced RDM services is embedded deeply in this broader network of services. Consequently, we define a local RDM service bundle as the full spectrum of services available to local researchers, including those provided locally, as well as external services to which the university brokers or facilitates access.

Our goal in this report is to describe how each of the four case study universities has scoped their local RDM service bundle and relate those choices to internal and external factors.

Case Studies in Scoping

Like many other research universities, our case study partners—Edinburgh, Illinois, Monash and Wageningen—chose to act in response to emerging RDM needs, and subsequently scoped a local RDM service bundle to meet those needs. In this section, we explore the nature of these service bundles and relate their scoping to influencing factors in both the internal and external environments.

UNIVERSITY OF EDINBURGH

The University of Edinburgh, founded in 1582, is one of Scotland’s four ancient universities.⁷ Among other affiliations, it is a member of the Russell Group, a group of public research universities in the UK, as well as the League of European Research Universities and the Coimbra Group, both associations of research-intensive European universities. Edinburgh is one of the Sutton Trust 13 universities, considered the most prestigious research-intensive universities in the UK; it also ranks highly in various international university league tables, such as the QS World University Rankings and the Times Higher Education World University Rankings.

Edinburgh’s efforts to acquire RDM capacity align with factors linked to both institutional priorities and external context. Edinburgh adopted an institution-wide research data management policy in 2011, which leads off by noting that “Research data will be managed to the highest standards throughout the research data lifecycle as part of the University’s

RDM Service Bundle Highlights University of Edinburgh



FIGURE 2. RDM SERVICE BUNDLE HIGHLIGHTS: UNIVERSITY OF EDINBURGH

commitment to research excellence.”⁸ Significantly, the commitment to RDM in Edinburgh’s data policy is backed up by clear pledges to supply the necessary resources to meet this goal.

The policy goes on to stipulate that “the University will provide training, support, advice, and where appropriate guidelines and templates” to support RDM, as well as “mechanisms and services for storage, backup, registration, deposit and retention of research data assets.” In short, Edinburgh’s data policy not only calls for effective RDM practices, but pledges to provide services and other resources across the Education-Expertise-Curation spaces (see figure 2) to ensure that these practices are implemented.

In addition to the institutional RDM policy, Edinburgh’s acquisition of RDM capacity is also impacted by several important external factors. Research Councils UK, an influential funder in the UK, released a set

of data management principles in 2011, noting that “publicly 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 keeping with this principle, “[d]ata with acknowledged long-term value should be preserved and remain accessible and usable for future research.”⁹

Additionally, Edinburgh, like all UK universities, is subject to the reporting requirements of the Research Excellence Framework, a national initiative to evaluate research quality in UK higher education; research data sets can be submitted to this process for evaluation along with more traditional outputs such as articles or monographs.¹⁰ In this way, Edinburgh’s prioritization of RDM, evidenced by the institutional data policy, complements, and is reinforced by, external guidelines and requirements.

RDM services at Edinburgh are housed within the [Information Services Group](#), which includes the library, information technology, learning and teaching technologies, study spaces, as well as [EDINA](#), a center of digital expertise based at the university. RDM-related services are consolidated under the [Research Data Service](#), “a suite of tools and support that helps staff and students across the university be effective with their research data before, during and after their research project is completed.” RDM services are provided through a matrixed organization, involving the library, campus IT and EDINA; however, the university desired a single point of contact for users, which is achieved by consolidating the RDM service bundle under the single Research Data Service.

Edinburgh’s RDM service bundle includes all three service areas described in figure 1, briefly described below.

Education

Edinburgh’s Education RDM services emphasize course- or workshop-based training. [MANTRA](#) is an online self-paced tutorial that introduces the importance, benefits and basic concepts of research data management. Additionally, Edinburgh offers five in-person [courses or workshops](#) that address more focused aspects of RDM, covering topics that range from creating data management plans to managing sensitive research data.

Edinburgh has collaborated with the University of North Carolina at Chapel Hill to create [Research Data Management and Sharing](#), a free, five-week MOOC (massive open online course) offered through the Coursera platform. The course covers a wide range of introductory topics, including data management needs across the research data lifecycle, data management plans and best practices for curating data sets. For a fee, a “Statement of Accomplishment” can be received upon successful completion of the course. Feedback on this course has been quite positive: as of June 2017, the average rating was 4.5 out of 5 from more than 80 responses.

In addition to these general educational resources, the Research Data Service also offers customized training sessions “on any aspect of RDM within your discipline, school, institute or research group.”¹¹

An interesting aspect of the Education component of Edinburgh’s RDM service bundle is that MANTRA and the RDM MOOC are community resources, rather than targeted exclusively to Edinburgh’s affiliated researchers. These resources are recognized as key educational services within the broader RDM community: for example, many university RDM web sites point to MANTRA as a recommended introduction to RDM. Of course, both MANTRA and the RDM MOOC were developed with funding from organizations with broad higher education interests (Jisc and the Institute of Museum and Library Services, respectively).

Key takeaway: National centers of RDM capacity are important supplements to the local RDM service bundle.

Yet Edinburgh’s participation in the development and maintenance of these resources, and their subsequent take-up by the broader RDM community, has established Edinburgh as a global leader in the provision of RDM Education services—in fact, services developed by Edinburgh have become part of the Education components of other universities’ RDM service bundles.

Expertise

The Research Data Service’s Expertise component consists of a general email contact point, to which researchers can submit RDM-related questions. The names and contact details of Research Data Service staff are also provided,¹² and researchers are encouraged to contact Research Data Service staff to discuss specific RDM needs.

Although not directly part of Edinburgh, the UK [Digital Curation Centre](#) (DCC) complements and extends Edinburgh's locally provided Expertise services. The DCC is a national center of expertise in digital curation, which provides services, research, and advice on data management and related issues. In addition to tools like DMPonline, a web-based data management planning tool, DCC provides "tailored support" for research organizations, including "development of new services, advocacy and awareness raising programmes, or support to customise tools, guidance and training."¹³ Edinburgh was a partner in the Jisc-funded consortium that launched the DCC, and the organization is now headquartered on the Edinburgh campus.

Edinburgh's leadership in the establishment of the DCC, and its ongoing role as the DCC's host institution, is another example of how Edinburgh has invested in the development of RDM services that benefit the broader RDM community. In this sense, Edinburgh's efforts in shaping its local RDM service bundle, at least in the Education and Expertise areas, have included investment in resources whose benefits reach well beyond the local university community.

Curation

Edinburgh offers an array of data curation services extending across the research life cycle. [DataStore](#) is a central file store for active research data, with a free allocation of 0.5 terabytes, and the ability to purchase additional space on demand. DataStore can be linked up with another service, [DataSync](#), a tool for synchronizing and sharing research data with internal or external collaborators. Currently, DataStore is the most used service in Edinburgh's RDM service bundle.

For long-term archiving, Edinburgh offers two data curation services. [DataShare](#) is an online repository of data sets produced at Edinburgh. Deposited data sets receive persistent identifiers (DOIs and handles), a suggested citation, and are publicly discoverable through Google and other search engines. DataShare commits to retain archived data sets indefinitely. In parallel with DataShare,

Edinburgh is also in the process of deploying its [DataVault](#) service, a private archival storage service for data sets that are no longer active, but must be retained and cannot be published. Access to deposited data sets is restricted to the data set creator or their designated representative. This service is especially useful for data sets that cannot be shared, but still must be preserved to comply with funder and/or university requirements.

Complementing the university's data archiving services, [Edinburgh's Pure](#) research information management system serves as a data asset registry for documenting data sets created at the university. Pure metadata records are automatically linked to Edinburgh's Research Explorer service, which provides a public portal to the university's research activities; in this way, research data set descriptions are displayed as part of a researcher's online profile, alongside other research outputs.

Key takeaway: Many research institutions choose to offer RDM Curation services in parallel with, rather than subsumed in, the institutional repository.

It is useful to note that the [Edinburgh Research Archive](#) (ERA), the university institutional repository, is not part of the research data curatorial infrastructure. ERA is a repository of original research "written by, or affiliated with, academic authors, or units, based at Edinburgh that have sufficient quality to be collected and preserved by the library, but which are not controlled by commercial publishers. Holdings include full-text digital doctoral theses, master's dissertations, project reports, briefing papers and out-of-print materials."¹⁴ ERA is considered Edinburgh's "legacy" repository, and is now used primarily for storing dissertations; currently, Pure is used as Edinburgh's institutional repository for new research publications.

In considering Edinburgh's RDM Curation services, a key feature is that important services like DataShare and DataVault are locally deployed and locally managed systems, with deposited data hosted on campus. The data sets deposited in these services are also locally focused, consisting primarily of those created at Edinburgh, by Edinburgh-affiliated researchers.

Yet there is an important external focus as well: DataShare is an open-access repository, supporting public discovery of deposited data sets. The quality and value of the data sets accessible through DataShare are linked both to Edinburgh's curatorial care and institutional reputation:

"What makes users more likely to use datasets they find in our repository? We hope that the careful guidance we give data producers about how to make their data usable, the quality checks we do upon submission, as well as the rule that every data item must include some human-readable documentation, reaffirms our commitment and reassures users that they are accessing and using quality data. We also hope that the University of Edinburgh's reputation for world-class research lets them know that they are getting authentic, state-of-the-art research data outputs from the DataShare repository."¹⁵

While Edinburgh's curatorial RDM services aim to meet the data archiving needs of local researchers, it is clear that the benefits from these services are expected to extend beyond the university itself, reaching into the broader scholarly community.

Reflections

Edinburgh's RDM service bundle was shaped by a strong—and early—university commitment to data management, coupled with an external context that mandated responsible data management and sharing. The service bundle demonstrates Edinburgh's stated commitment in its data policy to provide the necessary resources to ensure that research data are managed with the highest standard of care. While acknowledging the

many extra-institutional RDM services and resources that local researchers can and do use (physicists, in particular, tend to use discipline-based RDM services), Edinburgh nevertheless would like its locally deployed RDM services to be seen as the best choice for local researchers to meet their RDM needs, rather than as a last resort when external options are not available. At the least, regardless of where data created at the university is deposited, Edinburgh wants to ensure it is registered in its Pure system.

While the Edinburgh service bundle is currently well-formed and stable, there is nevertheless acknowledgement that the service bundle exists in a fluid RDM environment, and that flexibility and willingness to adapt are needed. In light of this, Edinburgh's RDM staff constantly engage with their user community and campus administrators through a variety of channels, to help set future direction for the service bundle.

Lingering uncertainty over future RDM needs—and consequent changes to the RDM service bundle—increases the pressure to listen. One factor that helps mitigate uncertainty is that all of Edinburgh's RDM services are shaped around university- and funder-framed requirements, which have helped to create some cross-disciplinary alignment of RDM expectations.

Key takeaway: Shaping the local RDM service bundle can have spillover effects in the form of resources that are available to and highly used by the broader RDM community.

The Edinburgh RDM service bundle was developed first and foremost to meet the RDM needs of local researchers and fulfill the objectives of the university data policy. But Edinburgh also sees itself as a global leader in the RDM service space, with a role in

supporting external stakeholders beyond the local university community. This is exemplified by the ongoing provision of the MANTRA tutorial and other educational sources, as well as Edinburgh's leadership in the establishment (and subsequent hosting) of the Digital Curation Centre. In this sense, Edinburgh's RDM service bundle is seen not as a self-contained resource serving local researchers exclusively, but instead as a "center of excellence" in a broader network of RDM services and resources.

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

The University of Illinois at Urbana-Champaign is a public research university and the flagship campus of the University of Illinois system. Illinois is categorized as an R1 institution according to the Carnegie Classification of Institutions of Higher Education, indicating the highest level of research activity. It is a member of the Big Ten Academic Alliance, an academic consortium of the 14 universities in the Big Ten Conference, as well as the Association of American Universities, a group of leading research universities in North America, and the Universities Research Association, an international consortium of research universities. Illinois is ranked 36th in the most recent Times Higher Education World University Rankings,¹⁶ and 47th in the latest *US News & World Report* Best Global Universities rankings.¹⁷

The external context surrounding Illinois's RDM service bundle played a key role in its subsequent development. In 2003, the US National Institutes of Health (NIH), a key funder in the medical sciences, began requesting data sharing plans in grant proposals exceeding \$500,000¹⁸; seven years later, the US National Science Foundation (NSF) also began requiring data management plans in grant proposals.¹⁹ Consequently, US research universities have begun to emphasize support for the development of data management plans as part of the grant submission process. In Illinois's case, emerging funder requirements like those of NIH and NSF translated into a need to develop

an RDM service bundle that demonstrated a robust capacity to manage data in compliance with funder requirements.

The emergence of individual agency funder requirements was buttressed in 2013 by the release of a White House Office of Science and Technology Policy (OSTP) memo directing federal agencies acting as funders to develop plans to increase public access to publicly funded research.²⁰ "Digital scientific data" was specifically called out as one of the research outputs covered by the directive. Unfortunately, as each agency responds with its own plan on its own timeline, a complex and dynamic landscape of funder requirements is emerging, which, in turn, poses challenges for scoping a local RDM service bundle in response.

In addition to these external factors, Illinois's RDM service bundle was also shaped by an institutional priority to address a perceived gap in support for data management. An internal review of campus-wide information technology initiated by the offices of the chancellor and the provost revealed that while Illinois produced many publications and data sets, the institution was doing little to manage the data—a serious gap for a university heavily engaged in data-intensive research. This review set the stage for the development of an RDM service bundle at Illinois.

Key takeaway: Ongoing fluidity and uncertainty in the RDM service space remains a challenge in scoping RDM service bundles.

RDM services at Illinois are collected under the [Research Data Service](#) (RDS), which is housed in the university library, and operates in partnership with the Office of the Vice Chancellor for Research, the Office of the

RDM Service Bundle Highlights University of Illinois at Urbana-Champaign

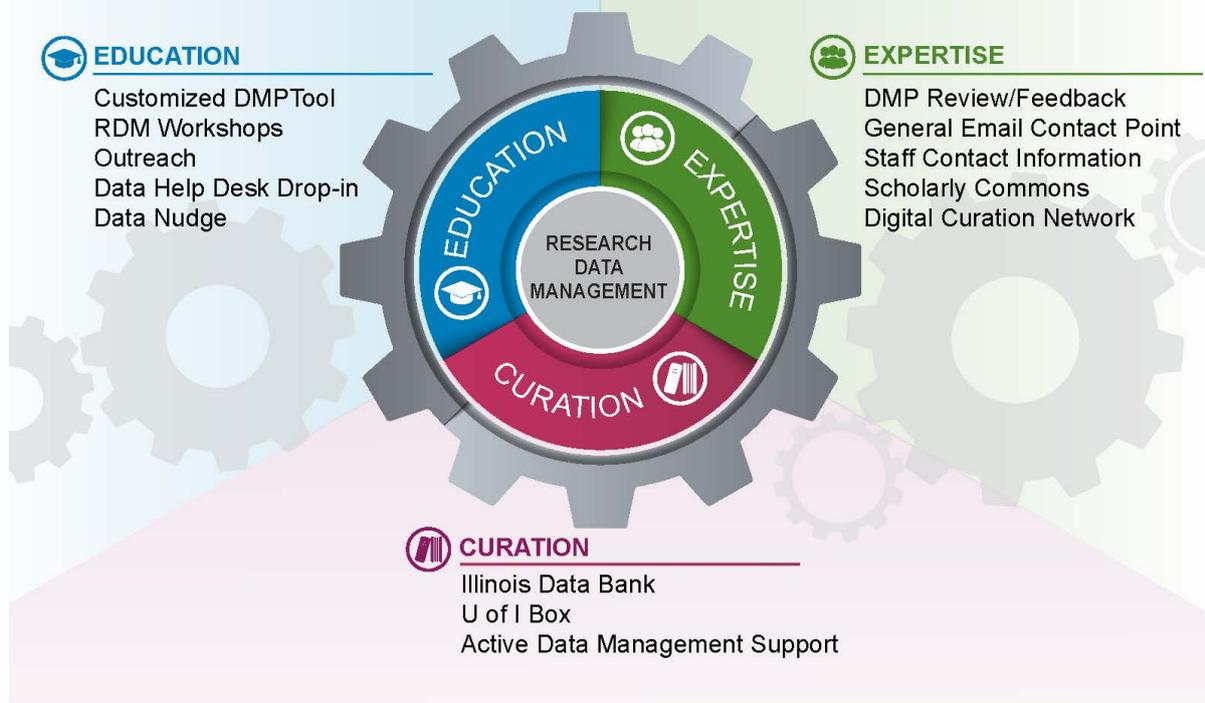


FIGURE 3. RDM SERVICE BUNDLE HIGHLIGHTS: UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Provost, the School of Information Sciences, Technology Services, and the National Center for Supercomputing Applications. RDS “is a campus-wide program that provides the Illinois research community with the expertise, tools and infrastructure necessary to manage and steward research data.”²¹ Illinois’s RDM service bundle includes Education, Expertise and Curation components, briefly described below.

Education

The Illinois RDM service bundle features an array of Education services aimed at guiding researchers through the process of [developing a data management plan](#) (DMP), including basic information on the importance of creating a DMP, best practices for producing DMPs and a customized version of the [DMPTool](#) for Illinois researchers. The Research Data Service also provides introductory information on [good practices for managing data](#) during the research process,

including advice on naming data sets, version control, file formats and descriptive metadata.

RDS offers a variety of [RDM-related workshops](#) and other training opportunities. Workshop topics range from an introduction to RDM, to more focused issues such as data documentation, data sharing and data workflows. RDS also offers customized RDM workshops on request for departments, research teams or student groups. In AY2016, RDS held more than 40 workshops to more than 500 attendees.²²

The Education services offered in the Illinois RDM service bundle are reinforced by outreach efforts to raise awareness among researchers of the importance and benefits of good data management practices. Outreach efforts take a number of forms, including interweaving RDM training into data/software carpentry workshops and maintaining [Data Help Desk Drop-in Hours](#), during which researchers can seek help on a variety of data

management issues. Other channels include an [RDS twitter account](#) and a [Data Nudge](#) service, which sends researchers monthly data management reminders and tips.

The emphasis on outreach to researchers, via channels like the Data Help Desk, the RDS twitter account or the Data Nudge service, is a distinctive feature of the Education services in the Illinois RDM service bundle. The idea of “pushing out” Education services in these and other ways is based on a recognition that data management can be an afterthought for many researchers; therefore, outreach is a key ingredient for ensuring that Education services are highly visible to those that would benefit from them.

Expertise

Expertise services feature prominently in the Illinois RDM service bundle. Researchers developing DMPs can submit draft versions to RDS for “[fast, free and confidential feedback](#),” including input from library subject specialists in order to provide a disciplinary perspective on the DMP. Functionality for requesting feedback is incorporated into the customized Illinois DMPtool template.

General support is provided through a dedicated RDS email account; in addition, the names and contact information for all RDS staff are listed as [contact points](#). Illinois researchers are encouraged to contact RDS for any questions or assistance with their RDM needs, such as choosing an appropriate data repository, organizing data sets or storing data.

RDS works in partnership with the university library’s [Scholarly Commons](#), a “technology enriched space for faculty, researchers, and graduate students to pursue research and receive expert copyright, data, digital humanities, digitization, scholarly communications, and usability consultation services.”²³ The Scholarly Commons also offers services to help researchers identify, purchase, format and use research data sets.

Organizationally, the Scholarly Commons falls within the library’s [Office of Research](#),

which provides services in support of campus collaborations to “create, manage, use and publish research information,”²⁴ such as researcher information systems (Pure) and scholarly communication and publishing support.

Illinois also participates in the recently launched [Data Curation Network](#), a collaboration between six US universities to share their respective data curation expertise, thereby creating a “network of expertise” for RDM and enabling the partners to act collectively to curate a greater variety of data (i.e., type, format, discipline) than a single institution could manage.

Expertise services included in the Illinois RDM service bundle have several distinctive features. Consultation with RDS is positioned as covering virtually any RDM need that might arise over the course of the research cycle, from the development of a data management plan, to managing active data during the research process, to securing long-term storage and curation. Rather than enumerating specific tasks or issues for which advice or consultation can be provided, Illinois’s Expertise services are for the most part open-ended in their scope. For example, researchers seeking help on saving and sharing their data are simply invited to contact RDS staff, who can “outline the options and go over resources to ensure your data is accessible for years to come.”²⁵ Finally, the Expertise services in the Illinois RDM service bundle often incorporate library subject specialists, blending general RDM advice with more focused disciplinary expertise.

Curation

The centerpiece of the Curation services in the Illinois RDM service bundle is the [Illinois Data Bank](#), “a public access repository for publishing research data from the University of Illinois at Urbana-Champaign.”²⁶ Developed and hosted locally, and launched in August 2016, the Data Bank is optimized for data set curation, focusing on preserving and providing ongoing access to data sets created by Illinois-affiliated researchers. Access to the Data Bank’s contents is open

to anyone, and, therefore, deposited data must be unencumbered by access restrictions or privacy concerns. Deposited data must also be in a “final state,” with no expected revisions. Each published dataset undergoes a curatorial review by RDS staff to “ensure metadata completeness and dataset discoverability.”²⁷

Use of the Data Bank is optional: researchers are advised that “[i]f your discipline already has a trusted repository, we recommend you deposit where your community knows to look,” and that the Data Bank “is available to campus researchers for when discipline-specific repositories do not exist.”²⁸ Under current policy, the Data Bank commits to retaining data for a minimum of five years, after which archived data sets are subject to preservation review and possible deaccession.

While the Data Bank is intended to support open sharing of research data sets, another service, [U of I Box](#), is aimed at supporting private or restricted sharing of data among research collaborators. U of I Box is an Illinois-branded instance of [Box](#), a cloud-based commercial content management and file sharing service. Illinois faculty, students and research teams can securely store either active or final data, and share it within or outside the Illinois network.

Key takeaway: Local RDM curatorial services don't need to be positioned as a preferred or first choice solution for local researchers.

Illinois also offers an [active data management \(ADS\) service](#), operated as a partnership between RDS, Illinois's National Center for Supercomputing Applications, and the campus Tech Services unit. ADS provides “mid-scale storage needs such as online backups, hosting of large datasets, operational data storage and scratch space for analysis, mid-

term data archiving and more.”²⁹ ADS was a response to a need for mid-range storage beyond what was available through Illinois's Box instance.

[IDEALS](#) is the university's DSpace-based institutional repository service. Currently, IDEALS is optimized for scholarly outputs such as publications, presentations and posters. Data sets fit somewhat awkwardly in this environment—file size, for example, is an issue. The differing requirements, functionalities and commitment to permanence for data curation (as opposed to managing PDF files and similar outputs) necessitate a separate repository for data sets. Prior to the launch of the Data Bank, IDEALS served as a last-resort repository for smaller data sets that could not find a home elsewhere; at this time, there are no plans to migrate these legacy data sets into the Data Bank. Like Edinburgh's Curation services, those in the Illinois RDM service bundle have both an inward- and outward-facing aspect: inward in the sense that the focus is on data sets created and managed by Illinois-affiliated researchers, and outward in the sense that, for the Illinois Data Bank, the goal is to provide “maximum public access”³⁰ to the research data outputs of the Illinois community.

Key takeaway: The RDM service bundle is not just a “back-office” component of a university's research support infrastructure; it often plays an important public-facing role.

Similarly, Illinois's Curation services are inward-focused in that they are intended to be “responsive to the Illinois research community,”³¹ yet, at the same time, contribute to the “advancement of scholarship and the public good in ways that are consistent with the US President's Office of Science and Technology Policy (“OSTP”) Public Access Memo of 2013.”³² In this sense, Illinois's Curation services—in particular, the Data Bank—serve the dual

purpose of supporting local researcher needs for data management and sharing, and, at the same time, elevate the profile and discoverability of the university's collective research effort by improving access to the full range of scholarly outputs produced on campus.

Reflections

The Illinois RDM service bundle emerged from a combination of external pressures, such as the OSTP memo on public access, and internal priorities, such as the need to correct a perceived gap in capacity for supporting data management. The resulting service bundle provides a comprehensive solution for meeting campus RDM needs, with services that extend over all three components of the RDM service space: Education, Expertise and Curation. The greatest early investment among the three,

however, is in Curation services, with the development and hosting of the Illinois Data Bank, combined with the curatorial activities undertaken by RDS staff in support of the Data Bank: review of deposited data sets, data set description, and linking deposited data sets to related scholarly outputs, like articles or software, archived elsewhere.

Most of the components of the Illinois RDM service bundle are located and managed on campus, making for a fairly self-contained local RDM solution. However, it is significant that Illinois positions its central curatorial service—the Data Bank—as just one repository in an eco-system of many repositories, to be used when external options do not exist or are otherwise not suitable. Moreover, the library does not see itself primarily as a “compliance agent” for the university, with a goal of preserving data sets on a large scale; this reflects new thinking within the university on how much data actually can or should be preserved.

A significant factor that shapes both current and future scoping of the Illinois RDM service bundle is uncertainty over how to adapt to future evolution of the RDM space. A key concern is how Illinois can avoid needless duplication of RDM services and infrastructure,

while, at the same time, provide an appropriate level of localization so that the RDM service bundle remains relevant to Illinois-affiliated researchers.

Illinois has provided a strong local RDM solution for its constituents, but committing significant additional investment for local RDM capacity may not be warranted, especially if collaboration and “above the institution” solutions end up being the future in the RDM space. Illinois embraces this possibility, and, indeed, its participation in the Data Curation Network is but one example of how components of the RDM service bundle can be lifted above the institution and deployed at scale. Future scoping of the Illinois RDM service bundle will be influenced by close observation of trends in cooperation and scaling in the broader RDM service space.

MONASH UNIVERSITY

Monash University is Australia's largest university by enrollment. It is a member of the [Group of Eight](#), a consortium of Australia's leading research-intensive universities.

Monash is also a member of several international higher education associations, including the [Association of Southeast Asian Institutions of Higher Learning](#)—a cooperative organization aimed at assisting its membership “to strengthen themselves through mutual self-help and to achieve international distinction in teaching, search and public service”³³—and the only Australian member of the [M8 Alliance](#), a group of 25 academic institutions committed to developing solutions to global health challenges. Monash ranks in the top 100 of the Times Higher Education and QS World University Rankings, as well as the Academic Ranking of World Universities.³⁴

Development of Monash's RDM service bundle occurred in an Australian environment that paid increasing attention to the issues of RDM and open data. In 2007, the Australian government released the Australian Code for the responsible Conduct of Research, which notes that “the responsible conduct of research includes the proper management

RDM Service Bundle Highlights Monash University

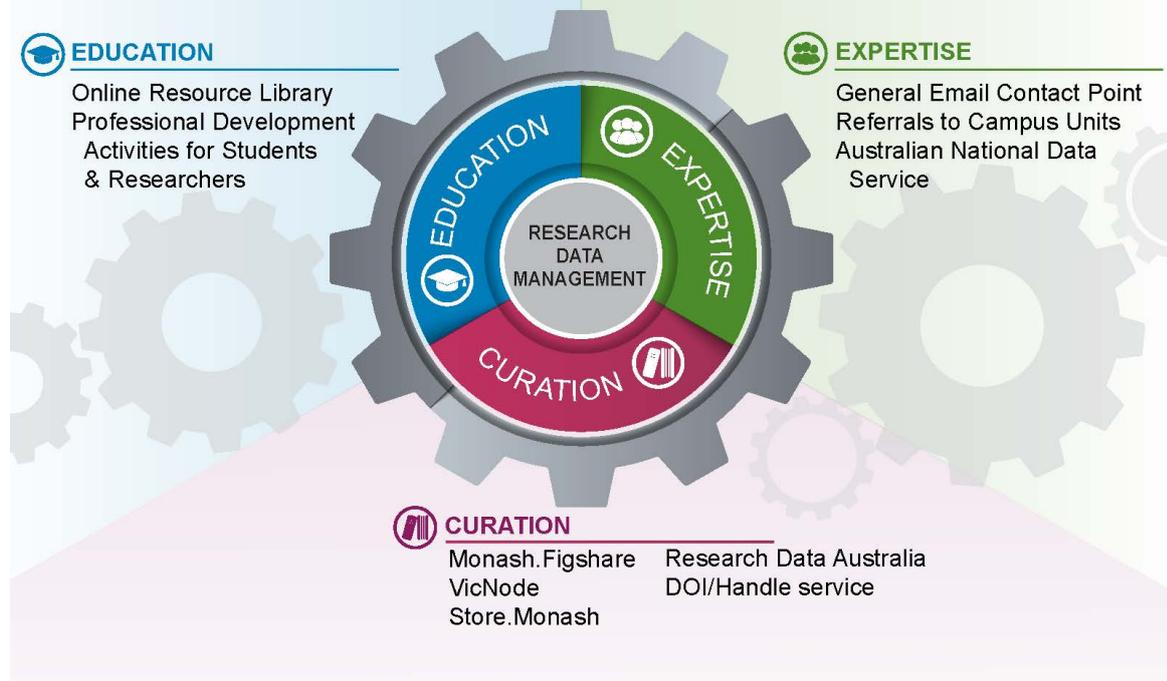


FIGURE 4. RDM SERVICE BUNDLE HIGHLIGHTS: MONASH UNIVERSITY

and retention of the research data,” and calls for the development of institutional policies on data retention, with a recommended minimum five-year retention period.³⁵

Beginning in 2014, the Australian Research Council (ARC) requests that principal investigators outline plans for data management in funding proposals;³⁶ the Australian National Health and Medical Research Council (NHMRC) issued a statement the following year acknowledging the importance of data sharing, and encouraging researchers to provide access to data produced in the course of NHMRC-supported research.³⁷ While these policies all promote attention to data management as part of the research process, none of them mandate open data. Nevertheless, their encouragement to deposit data in open repositories aligns with broader Australian support for open access, evidenced by open-access mandates for research publications released by NHMRC and ARC in 2012 and 2013, respectively.³⁸

The Monash University Research Data Management Policy, released in 2010, reflects an internal prioritization of RDM and is intended to “ensure that research data is stored, retained, made accessible for use and reuse, and/or disposed of, according to legal, statutory, ethical and funding bodies’ requirements.”³⁹ Yet the impetus for acquiring RDM capacity predates this policy: in particular, the university librarian recognized RDM as an emerging area and initiated the shaping of an RDM service bundle before a clear demand for such a service had emerged from Monash researchers. It was thus foresight, rather than explicit pressure from either internal or external sources, that provided the initial catalyst for development of the Monash RDM service bundle.

Monash’s first RDM coordinator was appointed in 2008—the first to hold such a post at an Australian university—with an initial responsibility of speaking to faculty about RDM and gathering support for creating an RDM strategy.

Monash credits several factors that informed their approach when developing and successfully implementing their local RDM service space. These include:

- Recognition on the part of senior university management that RDM is important, leading to a heavy and sustained internal prioritization
- University-wide effort to address RDM through collaboration and a shared responsibility
- Broad view of what is meant by research data, recognizing its manifold formats and manifestations
- Internal decision that the university would be a leader and early adopter in the RDM space⁴⁰

The [RDM service bundle at Monash](#) is housed in the library and provides education about, advice on and repository infrastructure for research data management. The library works in partnership with other campus units, such as the Monash e-Research Centre, as well as external organizations to provide RDM services to the Monash community.

Education

Monash offers [online resources](#) providing guidance on a variety of RDM-related issues, such as data planning (including research data planning checklists), sharing and disseminating data (including links to institutional and external data repository options), choosing among data storage options, ownership and rights (including Australian copyright law and student ownership of data under the Monash Intellectual Property Framework), and data security. These resources provide a practical introduction to key RDM topics and are supplemented by links to external resources where relevant.

In addition to self-directed online guidance, the library also offers [professional development activities](#) aimed at cultivating data management skills among researchers. These include short presentations as well as

more lengthy seminars, addressing topics such as “Research data management: Essentials for graduate research students” and “Data management and IP for supervisors.” These resources are offered on a standalone basis or embedded into other curricula; many of the sessions can also be customized for particular university groups. Monash has expanded its RDM training offerings for graduate students as part of its Doctoral Program Professional Development⁴¹ and is currently investigating options to make data skills training available “anytime, anywhere.”⁴²

Monash also offers a library of RDM-related information, “including links to relevant research data policy, procedures, strategy and guidance documents, workflows and templates and tools developed by Monash University in building research data management capability.”⁴³ Among these resources are a set of case studies, including an RDM policy implementation pilot project in the Faculty of Pharmacy and Pharmaceutical Sciences; RDM scenarios for graduate students in different disciplines; and a case study on the improvement of RDM at Monash undertaken by the UK Digital Curation Centre.⁴⁴

Taken together, the Education resources offered by Monash provide a wide-ranging introduction to the concepts and practice of RDM. Several features of these resources stand out. First, the resources seem particularly well-suited for researchers with no prior familiarity with RDM or why it is important to their work, guiding readers through basic information needed to form a general understanding of what RDM is, why it is important and beneficial, and what can be done about it. Second, a distinctive aspect of Education resources at Monash is that they are customized for use by graduate students. This emphasis on RDM training for students recognizes that RDM needs differ for researchers at different career stages, but also that data management skills are an important part of academic training, “essential in the research sector and transferable to a range of other workplaces.”⁴⁵

Expertise

In addition to general Education resources, Monash also coordinates services that provide researchers with advice and consultation on a wide range of RDM-related matters. One resource is a dedicated email account for researchers to submit RDM-related inquiries.⁴⁶ Monash researchers are also referred to a variety of campus units for expertise on specific issues directly or indirectly related to RDM.

For example, researchers are directed to the Monash eResearch Centre for advice on technical solutions for storing and sharing data, to the library for advice on copyright, and to the Records and Archives Service for advice on retention, appraisal and de-accession.⁴⁷

In addition to internal services, the [Australian National Data Service](#) (ANDS) also serves as a source of expertise on RDM-related topics. Established in 2008, ANDS is a national service aimed at making “Australia’s research data assets more valuable for researchers, research institutions and the nation.”⁴⁸ ANDS is a government-funded collaboration led by Monash, in partnership with the [Australian National University](#) and the [Commonwealth Scientific and Industrial Research Organisation](#) (CSIRO).

Among other services, ANDS provides advice and expertise on RDM-related issues, including data management, policy, sharing and reuse; it also brokers relationships between RDM stakeholder institutions.⁴⁹ In addition to expertise, ANDS offers an array of educational and skill-building resources, as well as Curation services.

A distinctive feature of the Expertise component of the Monash RDM service bundle is that it is diffused over many campus units, extending well beyond the library, which serves as the administrative home of Monash’s RDM services. As mentioned above, one of the factors Monash enumerates as contributing to the success of its RDM service bundle is that RDM is “treated as a multi-faceted issue requiring a university-wide effort,

collaborative approaches and a strong sense of shared responsibility.”⁵⁰ This is evidenced by the array of campus units that contribute expertise on RDM-related issues, a decentralized approach that leverages the distinctive knowledge of different components of the university community.

Curation

In 2015, Monash launched [monash.figshare](#), a local implementation of the “figshare for institutions” service—a customizable portal for institutional research outputs provided by [figshare](#), a commercial digital repository company. Monash.figshare is described as Monash’s institutional data repository, and is intended to promote the research outputs of the university, including research data and other scholarly materials. Since Monash has chosen to host data locally, monash.figshare serves as an interface and uploading process for storage resources coordinated by Monash.

Key takeaway: An RDM service bundle includes not just what is built and deployed locally, but the full range of services, sourced locally and externally, that the institution manages, or to which it brokers access.

In addition to data storage, monash.figshare also supports active data management and collaboration. Monash is investigating options for long-term archiving of data sets deposited into monash.figshare. Currently, data sets in monash.figshare are backed up in [VicNode](#), the Victoria-based node of the Australian government’s [Research Data Services](#), which, among other infrastructure services, operates a network of high-capacity storage nodes.

In addition to monash.figshare, Monash also provides a variety of storage options and resources for particular disciplinary needs,

as well as active data management. [Store.Monash](#) is Monash's implementation of the [MyTardis](#) data management platform, and supports the capture and storage of data generated from scientific instrumentation. The [OzFlux Repository](#), built on the Eddy data management platform developed by Monash's eResearch Centre, captures and stores atmospheric data from network of collection points around Australia and New Zealand. Monash also uses [LabArchives](#), a commercial cloud-based electronic lab notebook platform, for managing and storing research data online, as well as Monash Google Drive for small-scale data storage.

Complementing the Curation services offered directly through the university, Monash researchers also have access to several services offered through [ANDS](#). One such service is [Research Data Australia](#), a portal for discovering research data produced by Australian researchers. While not a storage service, Research Data Australia collects descriptions of and links to data sets, which are themselves held by 106 contributing research organizations around Australia, including Monash. A project was recently completed at Monash to automate the upload of monash.figshare metadata to Research Data Australia.

ANDS also offers a [DOI minting service](#) to support data citation; a [Handle service](#) to assign globally unique identifiers to data sets and other research outputs; and [Research Vocabularies Australia](#), a service for publishing, managing and discovering controlled vocabularies “commonly used by or relevant to Australian researchers.” These vocabularies serve as a “common language for the concepts in data sets”, and improve “discovery, linking, understanding and reuse of research data.”⁵¹

The Monash University Research Repository has served as the campus institutional repository for several years but will be decommissioned in 2017. Monash researchers now deposit open-access copies of their publications using the [MyResearch](#) portal, a

locally branded instance of Elsevier's Pure research information management system. Researchers are directed to deposit other types of research outputs (i.e., gray literature, working papers, media and datasets) in monash.figshare. Legacy content in the Research Repository will be migrated.

Key takeaway: Even institutions that scope their RDM service bundles to include services in each of the three components of the RDM service space—Education, Expertise and Curation—can differ in the extent or emphasis placed on a particular component.

In shaping the Curation component of its RDM service bundle, Monash relies heavily on an externally focused approach. Curatorial capacity is drawn from a number of externally sourced platforms: principally figshare, but also LabArchives, Google Drive, VicNode and various services provided through ANDS. Moreover, Monash has played a leading role in developing several “above the institution” RDM resources and solutions, including the ANDS service, as well as the MyTardis and Eddy data management platforms.

These last two examples also speak to the special attention Monash has paid to the RDM needs of disciplines collecting data sets directly from scientific instrumentation. Leveraging standardized, externally sourced solutions, while at the same time accommodating the specific RDM needs of different communities on campus—like those disciplines dependent on instrument-intensive data collection—is a key trade-off Monash must face in managing its RDM Curation services: “balanc[ing] the efficiencies that come from a well-managed portfolio of standardised tools with researchers' requirements for flexible and innovative

solutions that support research;” in this sense, the university “does not adopt a ‘one size fits all’ approach.”⁵²

Reflections

Monash was an early adopter of RDM services and has since developed a mature RDM service bundle that covers the Education, Expertise and Curation components of the RDM service space. Impetus for acquiring RDM capacity at Monash emerged primarily from an internal view that RDM was an issue of emerging importance and that the university needed to respond. The resulting RDM service bundle is positioned as a first-choice solution for Monash researchers, rather than a stop-gap or “last resort” service when external options are not available.

A distinctive feature of the Monash RDM service bundle is that it is deeply embedded in the broader, external eco-system of RDM services. This is particularly true of Monash’s Curation services, which are sourced from a variety of commercial, national and government entities. The external focus with which the RDM service bundle is imbued also reflects a sense that the continued development and maturation of the RDM service space will be accomplished through cooperation and engagement with other institutions, such as Monash’s participation in ANDS, and its development of several data management platforms that have been deployed by institutions across Australia. At Monash, leadership in RDM requires engagement in a broader community of RDM stakeholders: “sharing information, expertise and tools with other organisations to build national capability, not just through the lead agent role in the Australian National Data Service (ANDS), but also as an institution.”⁵³

Although Monash has developed a relatively mature set of offerings in its RDM service bundle, take-up of these services is still building, and outreach about RDM to the university community remains a priority. Monash has found that when conversing with researchers about RDM, it is important to

have solutions in place to support those conversations.

It is not sufficient to simply tell researchers about the importance of RDM and why they should be addressing it, while providing no solutions or gesturing toward as yet unrealized “vaporware.” In this sense, deployment of an RDM service bundle in advance of strong demand for such services is, in Monash’s view, a sensible strategy: when researchers are ready to seek support services from the library, the library must be prepared with a mature offering for them.

In Monash’s view, a key ingredient for success in outreach is confidence when engaging with researchers. Liaison staff need to be credible when making the case for RDM, and must go beyond general discussion of “things to know about RDM”, instead providing researchers with practical RDM support relevant to their workflows and disciplines. Monash sees this as a challenging area, requiring continual skill-building on the part of liaison librarians.

Addressing this need may require shifting resources within the library, discontinuing some library activities to free up capacity to boost engagement around research support.

WAGENINGEN UNIVERSITY & RESEARCH

[Wageningen University & Research](#) is a public university with a focus on life and agricultural sciences. It consists of the university, along with nine research institutes, which carry out applied and field-based research. Wageningen is a member of several Netherlands-based higher education consortia, including the [Association of Universities in the Netherlands](#) (VSNU), which consists of the 14 Dutch research universities, and the [4TU.Federation](#), which includes four leading Dutch technical universities (Wageningen, Delft University of Technology, Eindhoven University of Technology and the University of Twente). In 2016, Wageningen was ranked 65th in the Times Higher Education World University Rankings,⁵⁴ and received the top ranking for its agriculture and forestry programs in the QS World University Rankings.⁵⁵

RDM Service Bundle Highlights Wageningen University & Research

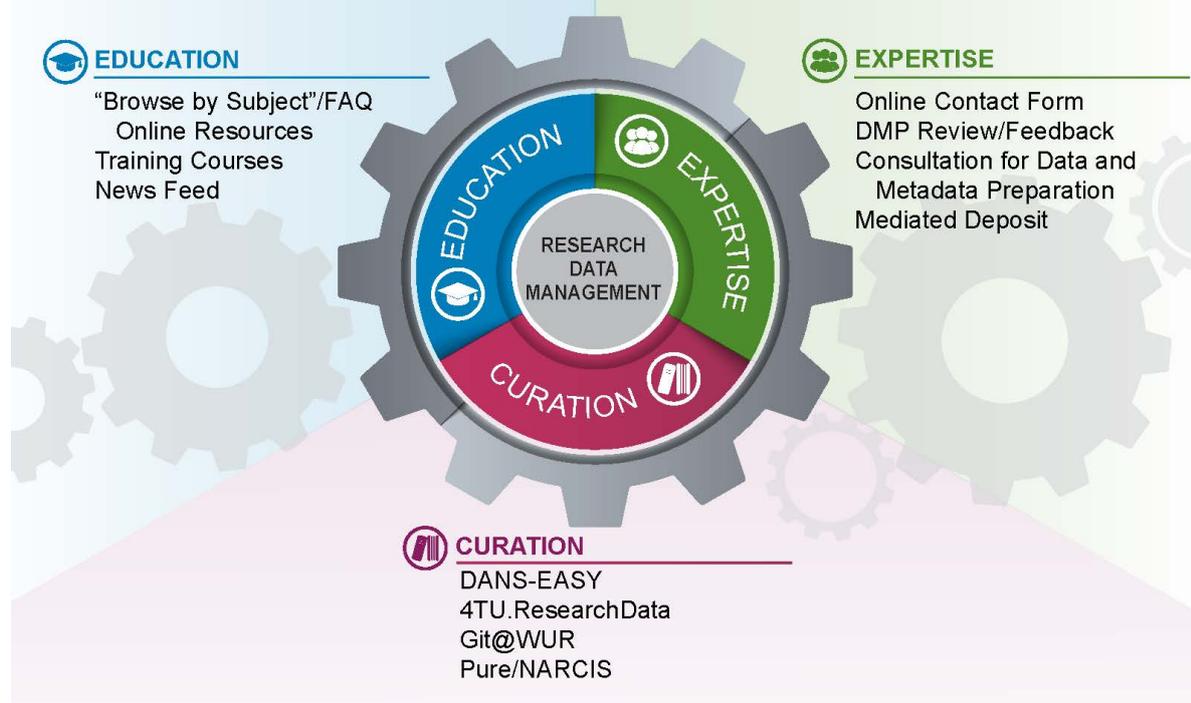


FIGURE 5. RDM SERVICE BUNDLE HIGHLIGHTS: WAGENINGEN UNIVERSITY & RESEARCH

The external context surrounding Wageningen’s RDM service bundle includes a number of national- and consortial-scale RDM initiatives. The [National Coordination Point Research Data Management](#) (LCRDM) is a national-scale effort to coordinate RDM policy and infrastructure planning, initiated by VSNU and [SURF](#), an organization for technology collaboration in Dutch education and research.

LCRDM’s mission is “to prepare, facilitate and monitor the development and implementation of Research Data Management Policy for scientific research in the Netherlands and to ensure the exchange of knowledge and experience on a national level so that the Netherlands can achieve an efficient and effective development and implementation of research data management.”⁵⁶ Currently, LCRDM focuses on facilities/infrastructure, legal issues, finance, researcher support and awareness/outreach.

In addition, [Data Archiving and Networked Services](#) (DANS), jointly sponsored by the [Royal Netherlands Academy of Arts and Sciences](#) and the [Netherlands Organization for Scientific Research](#), operates [DataverseNL](#) (a distributed network of Dutch data repositories), [EASY](#) (an online archiving system) and [NARCIS](#) (a discovery service for Dutch research outputs, experts and news); DANS also provides a range of RDM training and consultancy services.

Wageningen is directly involved in several group-scale collaborations around RDM. The 4TU consortium operates a shared data archive—[4TU.ResearchData](#)—for its members. Data in the archive are discoverable through the NARCIS discovery platform. In addition, the [UKB library consortium](#), representing 13 Dutch university libraries along with the national library, maintains a Research Data Management working group, which supports RDM knowledge exchange across the UKB membership.

Overlaying both the national- and group-scale coordination is the Netherlands Code of Conduct for Scientific Practice, which among other principles, prescribes that data sets are to be retained for a minimum of ten years to promote reuse and to permit validation of published findings through replication.⁵⁷ Additionally, several well-publicized cases of academic fraud in the Netherlands involving fabricated data and other irregularities incentivized Dutch research institutions to prioritize leadership and guidance on RDM, and likely accelerated efforts at Wageningen and other Dutch institutions to develop RDM service bundles.⁵⁸

Key takeaway: Cross-unit responsibility for the RDM service bundle is common.

In addition to the external context surrounding Wageningen's RDM service bundle, some institutional prioritization has also been placed on promoting data management. In response to a 2013 proposal from the Graduate Schools, Wageningen released a new policy, which takes effect in 2017, requiring PhD candidates and chair groups (research teams) to develop data management plans for their projects.⁵⁹

Wageningen's RDM services are positioned as part of the university's core research infrastructure and are collected into a web presence called the [Data Management Support Hub](#). The Data Management Support group is composed of experts from the library, IT Services, Document Management and Logistics, and Corporate Governance and Legal Services departments.

The cross-unit allocation of RDM responsibilities reflects the university's desire to position RDM both as an institutional priority and an institution-level policy. While the library is tasked with organizing Wageningen's RDM capacity, and bears ultimate responsibility for it, the

RDM service bundle is a collaborative effort within the university, and is branded at the university level.

Education

Wageningen's RDM service bundle features Education resources with practical guidance on specific data management issues and related topics. A ["Browse by Subject" page](#) serves as a starting point for information resources on a variety of topics, such as data management planning, data documentation and storage solutions. For example, the data management planning page provides a brief overview of why data planning is important; sets out the institutional requirements for PhD students and chair groups regarding DMPs; and provides links to guides, templates and examples.⁶⁰ The Browse by Subject resource is supplemented by a [Frequently Asked Questions page](#), which provides short answers to focused questions about RDM, such as "where can I store my current research data?" and "what are the minimum data retention periods?"

Wageningen offers a [one-day training course](#) through the library on data management planning; the course is targeted to PhD students and postdoctoral researchers, and participants earn credit (0.4 ECTS) from the Graduate Schools upon completion. Follow-up sessions provide an opportunity for participants to discuss and refine their DMPs in a small group setting. Additional RDM training courses, such as "Essentials 4 Data Support," are accessible through the 4TU consortium's Centre for Research Data; customized presentations for specific research groups are also available.⁶¹

The Data Management Support Hub provides a news feed on its home page featuring updates on RDM-related happenings in the institutional, national and international community, such as training opportunities, workshops and conferences.

An interesting feature of the Education resources in Wageningen's RDM service bundle is that source code management and

curation are specifically addressed as part of the data management process. For example, in the FAQ resource, one question is “How can I safely store and manage my software source code?” while in the “Browse By Subject” resource, a section is provided on managing source code within Wageningen’s GitLab implementation, Git@WUR.⁶²

The Education component of the RDM service bundle is focused on gathering together information on fairly granular RDM-related topics and presenting the information in an easily consumable form. With the exception of the one-day training course, there is little emphasis on self-directed tutorials or structured workshops and seminars. Rather, Education resources are largely directed toward identifying specific RDM-related problems or needs and supplying the information necessary to solve them. This is especially evident with the FAQ resource. In fact, the tagline of the Data Management Support Hub itself is “Where questions and answers meet.” While there is some general information on the importance and benefits of RDM, the focus is very much on helping the researcher navigate specific RDM-related tasks.

Expertise

Expertise is emphasized strongly within Wageningen’s RDM service bundle; indeed, the Data Management Support Hub is positioned as one of the university’s array of “Expertise and Services” supporting the research endeavor at Wageningen. A distinctive feature of the Expertise component of Wageningen’s RDM services is its ubiquity across the service bundle, beginning with a “I have a question” button, prominently displayed on the Data Management Support Hub home page, that connects researchers via email to Data Management Support staff.⁶³ Similarly, each topic linked under the “Browse by Subject” resource has an equally prominent contact form at the top of the page. Throughout the Data Management Support Hub, researchers are repeatedly encouraged to engage directly with Data Management Support staff.

In addition to general support, Wageningen provides several specific Expertise services. Researchers developing a data management plan can submit it for review to Data Management Support staff and receive feedback and advice.⁶⁴ Data Management Support staff also provide consultation on data set documentation and metadata creation, and mediate deposit into either the DANS-EASY or 4TU.ResearchData repositories.⁶⁵ Staff can also link data sets stored in these repositories—or source code stored in Git@WUR—to publications registered in Wageningen’s Pure implementation.⁶⁶

With the exception of the specific services enumerated above, Wageningen’s Expertise services are for the most part open-ended in scope: researchers are encouraged to seek assistance and advice from Data Management Support staff for issues pertaining to all facets of the RDM process. The Support Hub describes its staff as a “team of experts”⁶⁷ who are available to support researchers in meeting their RDM needs. An important component of this support is helping Wageningen researchers identify RDM service options within the broader consortial and national context, and connecting researchers to the right solution.

Curation

The Curation component of the Wageningen RDM service bundle is scoped with a reliance on externally provided services. The university does not maintain a locally hosted data repository; instead, researchers are encouraged to utilize discipline-focused repositories as a first choice, or, alternatively, to deposit their data into the national-scale DANS-EASY archive or the consortial-scale 4TU.ResearchData repository. Mediated deposit into the DANS and 4TU repositories is available from Wageningen Data Management Support staff; researchers provide the data files, data documentation (a “readme.txt” file that describes the data and a “methodology.txt” file that describes the data collection process); and a limited set of discovery metadata. Data Management

Support staff then manage the process of depositing the data into either the DANS or 4TU repository.⁶⁸

Key takeaway: Scoping an RDM service bundle sufficient to meet institutional needs does not necessarily mean implementing the full range of services within the RDM service space.

Wageningen provides researchers with Curatorial support in conjunction with data deposit in repositories other than DANS or 4TU; however, in these cases, it is limited to advice on data preparation. Actual preparation of data files for deposit, and the deposit process itself, are the responsibility of the researcher.⁶⁹

Wageningen researchers enjoy cost-free storage capacity (with some exceptions) in the 4TU.ResearchData repository through Wageningen's membership in the 4TU consortium. Under current service terms, data deposited in the 4TU repository is maintained for a minimum of 15 years, with the option for extension.⁷⁰

In addition to Curation resources for research data sets, Wageningen also offers an instance of GitLab (Git@WUR), a source code repository management application. Source code stored in Git@WUR can be assigned a persistent identifier and then linked to research publications reporting findings based on the code.⁷¹

Wageningen encourages researchers to register all data sets produced in the Pure research information management system, which surfaces them in the university's experts/staff profiling and publications system and the national NARCIS discovery portal.

Data Management Support staff also monitor external repositories for data sets produced by Wageningen researchers; these are then registered in Pure as well.

The distinguishing feature of the Curation component of the Wageningen RDM service bundle is that it has been almost entirely externalized: rather than building locally operated resources, Wageningen has relied on national- and consortial-scale resources available in the wider RDM service ecosystem.

Moreover, Wageningen acknowledges the importance of discipline-based RDM solutions, and recommends them as a first choice for researchers. This is partly motivated by a recognition that its researchers cannot be forced to use a local solution, particularly if it requires a disruption of current practices and workflows, or results in a lower level of service than what is available externally. But, at the same time, it also reflects local priority:

Key takeaway: No RDM service bundle is an island—all are connected, to a greater or lesser degree, to the broader, external RDM service ecosystem.

Wageningen found that its researchers valued active storage and collaboration environments the most in regard to their data, while long-term archiving ranked as a matter of less concern.⁷² Wageningen estimates that roughly half of its researchers receive adequate RDM support within their disciplinary domains. But regardless of where and with whom data curation is sourced, Wageningen encourages registration of data sets in the local instance of Pure.

Reflections

The Wageningen RDM service bundle is scoped partially with locally developed resources (chiefly in its Education and Expertise components) and

partially with externally sourced services (chiefly in its Curation component). *Support* is a key element of the locally sourced components of the service bundle: the Wageningen RDM web presence repeatedly emphasizes the availability of Data Management Support staff to provide advice and consultation on any RDM need that might arise in a researcher’s scholarly workflow.

This serves to elevate the prominence of the Expertise component of Wageningen’s RDM capacity *vis-à-vis* Education and Curation, suggesting that engagement with researchers—answering questions, advising in the choice of external repository options, assisting in data documentation and deposit—is the key deliverable of the Wageningen RDM service bundle.

Wageningen takes the view that building and managing data curation infrastructure requires special expertise, which the university has not yet developed internally. This strongly influences the choice to externalize much of the Curation aspect of the RDM service bundle. This strategy is aided by the fact that Wageningen is a small institution operating in a small country with a small national higher education system.

Key takeaway: An RDM service bundle includes not just what is built and deployed locally, but the full range of services, sourced locally and externally, that the institution manages, or to which it brokers access.

Personal networks across institutions are extensive—“everybody knows everybody”—and this facilitates understanding of the overall national context and collective RDM capacity. In such an environment, the option to engage in collaboration with trusted partners, and to develop “above the institution” solutions, is particularly efficacious.

Looking to the future, curation of big (e.g., multi-terabyte) data sets is an important area that the scoping of the RDM service bundle must evolve to accommodate. At Wageningen, animal and plant sciences are important disciplinary specialties where huge data sets are not uncommon. In this and other areas, Wageningen expects to rely on external collaboration rather than building internal capacity, with an emphasis on partnerships within the 4TU consortium.

Wageningen’s RDM service bundle is deeply embedded in the external RDM service ecosystem, particularly with consortial- and national-scale curatorial resources. The importance of externally sourced RDM services to its researchers is reflected in Wageningen’s recommendation that discipline-based repositories should be sought as a first choice for data curation, with the DANS and 4TU repositories serving as alternatives if a suitable disciplinary solution is not available. But the fact that the Curation component of the service bundle is largely entrusted to outside services does not diminish the important role retained locally: identifying and brokering access to appropriate RDM solutions for Wageningen researchers, and, in some cases, mediating use of these services to ensure that researcher needs are met.

Scoping the RDM Service Bundle: Takeaways

This report explores the nature of the RDM capacity acquired by four research universities in four different national contexts, highlighting, where possible, key internal and external factors that shaped the contours of this capacity. Along the way, we have highlighted a number of general observations, or takeaways, derived from the four case studies detailed in this report.

In this concluding section, we collect these observations together and supplement them with some additional explication. We acknowledge that our sample size is small and therefore inferences or extrapolations drawn from the case studies should be considered

with caution. Nevertheless, we believe that the case studies, taken together, yield some useful starting points for thinking about general trends in scoping university RDM service bundles:

- **RDM is not a monolithic set of services duplicated across universities. It is a customized solution shaped by a range of internal and external factors operating on local decision-making.**

Scoping an RDM service bundle is not a binary question of whether or not to acquire RDM capacity, but a much more nuanced question of which specific RDM services are needed to support local needs.

- **Even institutions that scope their RDM service bundles to include services in each of the three components of the RDM service space—Education, Expertise and Curation—can differ in the extent or emphasis placed on a particular component.**

The relative emphasis placed on each of the three service areas will depend on a host of factors, including the university’s research profile, local prioritization, and the availability of external resources and services.

- **Scoping an RDM service bundle sufficient to meet institutional needs does not necessarily mean implementing the full range of services within the RDM service space.**

This is a corollary to the previous observation: because local circumstances influence both the relative emphasis on Education, Expertise and Curation, as well as the sets of services implemented within each of these areas, the “minimum viable product,” in terms of a university’s RDM service bundle, will differ from institution to institution. Optimizing the RDM service bundle to suit local circumstances often involves being selective of the services offered, rather than comprehensive.

- **An RDM service bundle includes not just what is built and deployed locally, but the full range of services, sourced locally and externally, that the institution manages, or to which it brokers access.**

Scoping an RDM service bundle is just as much about situating it within a broader landscape of external services and resources as it is about building and deploying services locally. Often, a key component of an RDM service bundle is a service supplied by an external provider, with access to that service facilitated by the local university.

- **Shaping the local RDM service bundle can have spillover effects in the form of resources that are available to and highly used by the broader RDM community.**

Services and resources developed for the local RDM service bundle, if made available for general use, can become widely used across universities. This is particularly evident with Education resources, like MANTRA or RDM LibGuides, which are often cross-linked across service bundles. In this way, services or resources originating in one RDM service bundle can be incorporated into many service bundles, even without direct cross-institutional collaboration.

- **Many research institutions choose to offer RDM Curation services in parallel with, rather than subsumed in, the institutional repository. However, research information management (RIM) systems can be an important component of the RDM service bundle.**

In many cases, the RDM service bundle is scoped as a distinct set of services supporting a distinct class of research output—data sets—instead of being subsumed within a set of more general repository services aimed at curating a wide range of research outputs. Reasons for implementing dedicated data curation infrastructure include file size, metadata

requirements, preservation commitment and storage requirements. On the other hand, RIM systems, like Pure, are often used as data set registries, as well as to surface data sets in researcher profiles.

- **Ongoing fluidity and uncertainty in the RDM service space remains a challenge in scoping RDM service bundles.**

Although many universities have developed mature, stable RDM service bundles, there is little sense that the environment in which these service bundles are situated is equally mature or stable. The RDM service space is still dynamic, making planning for future development of and investment in RDM services a fraught exercise. Uncertainty surrounds the availability of reliable external RDM solutions, opportunities for cross-institutional collaboration, researcher priorities and requirements, and funder/government data policies.

- **Local RDM curatorial services don't need to be positioned as a preferred or first choice solution for local researchers.**

Given the extensive network of discipline-, consortial- and national-scale RDM services, many institutions have scoped their local RDM service bundles to be complementary to, rather than parallel with, these external options. Several considerations are important here: have local researchers already established workflows that incorporate external RDM services? Are sufficient resources available to build and manage local solutions? Is being perceived as a "center of excellence" in RDM an important aspect of the institution's reputation or research profile?

- **Despite institutional memberships in a variety of cooperative peer groups, the RDM service bundle retains a local branding and at least some institution-scale components.**

Each of the RDM service bundles examined in this report are embedded in universities participating in a variety of inter-institutional associations. However, none of the universities have completely relinquished the provision of RDM services to "above the institution" or cooperative solutions; all (with the notable exception of the Curation component of Wageningen's service bundle) have implemented some form of local solution in the Education, Expertise and Curation service areas. Moreover, the service bundle itself is branded as a local university offering. Some institutions may develop an institution-scale RDM service bundle by choice (e.g., to cultivate a reputation as a leading RDM provider); others may find institution-scale sub-optimal and a result of untapped opportunities for collaboration and consolidation in the RDM service space.

- **National centers of RDM capacity are important supplements to the local RDM service bundle.**

A number of national-scale data curation centers have emerged that provide RDM-related services in Education, Expertise and Curation. Three of the four universities examined in this report benefit from such centers: the UK Digital Curation Centre (DCC), the Australian National Data Service (ANDS) and the Dutch Data Archiving and Networked Services (DANS). These centers serve to deepen and extend the scope of local RDM service bundles, and enhance the ability of a university to selectively specialize in locally deployed RDM services.

- **Cross-unit responsibility for the RDM service bundle is common.**

None of the RDM service bundles discussed in this report are the sole responsibility of a single campus unit. Although the service bundle may be housed and administered within a single unit like the library, the services provided

are the result of partnerships across multiple campus units. Consequently, the RDM service bundle is often branded at the university level, rather than at the level of a single campus unit. In many cases, scoping the RDM service bundle will be an exercise in cross-campus consultation and collaboration, leading to distributed responsibilities across multiple campus partners.

- **The RDM service bundle is not just a “back-office” component of a university’s research support infrastructure; it often plays an important public-facing role.**

The case studies make clear that the RDM service bundle—in the form of openly accessible data repositories, registries of locally produced data sets, and curated linkages to publications and staff profile systems—is usually scoped not just to meet researcher data management needs, compliance requirements and other administrative

priorities, but also to enhance the visibility and accessibility of an important segment of the university’s research output in the broader scholarly community.

- **No RDM service bundle is an island— all are connected, to a greater or lesser degree, to the broader, external RDM service ecosystem.**

Generally speaking, RDM service bundles are not self-contained. Although they differ in the degree to which they incorporate external services and resources, they are scoped to leverage some connection to the external RDM service eco-system, creating a network of interdependence—varying in intensity from institution to institution—across the RDM service space. In this sense, an important RDM-related service is helping local researchers navigate the service options available to them both locally and externally.

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 Jacqueliijn 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. 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, OH: OCLC Research. doi:10.25333/C3PG8J.
2. For even more detail into each of the university's national, institutional and consortial contexts, and their RDM service space offerings, see the OCLC Research Supplemental: Bryant, Rebecca, Brian Lavoie and Constance Malpas. 2017. *The Realities of Research Data Management: Selected University RDM Service Profiles*. Dublin, OH: OCLC Research. doi:10.25333/C3T92S.
3. See note 1.
4. Lynch, Clifford. 2017. "Updating the Agenda for Academic Libraries and Scholarly Communications." *College and Research Libraries*. <http://crl.acrl.org/index.php/crl/article/view/16577/18023>. 78, no. 2: 129
5. Adams Becker, S., M Cummins, A. Davis, A. Freeman, C. Giesinger Hall, V. Ananthanarayanan, K. Langley, and N. Wolfson. 2017. *The NMC Horizon Report: 2017 Library Edition*. <http://cdn.nmc.org/media/2017-nmc-horizon-report-library-EN.pdf>, 18-19.
6. Although deciding to act precedes deciding what to do, we chose to address the latter first in our report series, because we felt it was useful to describe the RDM service bundles at each of our case study institutions at the outset, in order to provide context for discussions in future reports.
7. The other three are University of Glasgow, University of St Andrews and University of Aberdeen.
8. University of Edinburgh. "Research Data Management Policy." <http://www.ed.ac.uk/information-services/about/policies-and-regulations/research-data-policy>.
9. "RCUK Common Principles on Data Policy" <http://www.rcuk.ac.uk/research/datapolicy/>. In July 2016, RCUK joined several other stakeholder groups (Higher Education Funding Council for England, Universities UK, and Wellcome Trust) in publishing Concordat on Open Research Data, a set of principles for managing, curating and sharing research data.
10. In 2014, a new UK Open Access Policy was released, requiring that journal articles and conference proceedings be deposited in repositories, discoverable and freely accessible in order to be eligible for consideration in the next Research Excellence Framework exercise. Currently, this policy does not apply to research data.
11. The University of Edinburgh. "Data Training." Research Data Service. <http://www.ed.ac.uk/information-services/research-support/research-data-service/training>.
12. The University of Edinburgh. "Contact." Research Data Service. <http://www.ed.ac.uk/information-services/research-support/research-data-service/contact>.
13. DCC. "Tailored Support." <http://www.dcc.ac.uk/tailored-support>.
14. The University of Edinburgh. "Communities in DSpace." ERA. <https://www.era.lib.ed.ac.uk/>.
15. The University of Edinburgh. "Trustworthy Digital Repository." Research Data Service. <http://www.ed.ac.uk/information-services/research-support/research-data-service/sharing-preserving-data/data-repository/trustworthy-digital-repository>.
16. The World University Rankings 2016-2017. <https://www.timeshighereducation.com/world-university-rankings/2017/world-ranking>.
17. U.S. News & World Report. "Best Global Universities Rankings." <https://www.usnews.com/education/best-global-universities/rankings>.

18. National Institutes of Health (NIH). 2003. Final NIH Statement on Sharing Research Data. <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-03-032.html>.
19. National Science Foundation. "Dissemination and Sharing of Research Results." <https://www.nsf.gov/bfa/dias/policy/dmp.jsp>.
20. The White House: President Barak Obama. "OSTP Public Access Policy Forum." Office of Science and Technology Policy. <https://web.archive.org/web/20161222160413/https://www.whitehouse.gov/administration/eop/ostp/library/publicaccesspolicy>.
21. University of Illinois at Urbana Champaign (UIUC). "Research Data Service." <http://www.library.illinois.edu/rds/>.
22. IDEALS. "RDS Campus Annual Report 2016." <https://www.ideals.illinois.edu/handle/2142/91571>.
23. UIUC. "Scholarly Commons." http://www.library.illinois.edu/sc/about_us/.
24. UIUC. "Office of Research." <http://www.library.illinois.edu/administration/research/>.
25. UIUC. "Saving and Sharing Your Data." Research Data Service. <http://www.library.illinois.edu/rds/save/>.
26. UIUC. "Illinois Data Bank." <https://databank.illinois.edu/>.
27. UIUC. "Curation Procedures." Illinois Data Bank. https://databank.illinois.edu/help#curation_procedures.
28. See note 24.
29. UIUC. "Active Data Storage." Research Data Services. <http://www.library.illinois.edu/rds/active-data-storage-overview/>.
30. UIUC. "Illinois Data Bank Policy Framework and Definitions." <https://databank.illinois.edu/policies>.
31. *ibid.*
32. *ibid.*
33. The Association of Southeast Asian Institutions of Higher Learning (ASAIHL). <http://www.asaihl2016.upm.edu.my/home.php>
34. Times Higher Education World University Rankings 2016-2017. <https://www.timeshighereducation.com/world-university-rankings/2017/world-ranking>; QS Top Universities. <https://www.topuniversities.com/university-rankings/world-university-rankings/2018>; and Academic Ranking of World Universities. <http://www.shanghairanking.com/World-University-Rankings/Monash-University.html>.
35. National Health and Medical Research Council (Australia). 2007. Australian code for the responsible conduct of research: revision of the Joint NHMRC/AVCC statement and guidelines on research practice. Canberra: National Health and Medical Research Council. <https://www.nhmrc.gov.au/files/nhmrc/file/publications/synopses/r39.pdf>, p. 2.1
36. Australian Government: Australian Research Council (ARC). "Research Data Management." <http://www.arc.gov.au/research-data-management>.
37. Australian Government: National Health and Medical Research Council (NHMRC) "NHMRC Statement on Data Sharing." <https://www.nhmrc.gov.au/grants-funding/policy/nhmrc-statement-data-sharing>.
38. NHMRC Open Access Policy: <https://www.nhmrc.gov.au/grants-funding/policy/nhmrc-open-access-policy>; and ARC Open Access Policy: <http://www.arc.gov.au/arc-open-access-policy>.

39. Monash University. "Research Data Management Policy." Last reviewed 24 November.2013. https://www.monash.edu/data/assets/pdf_file/0011/797339/Research-Data-Management-Policy.pdf.
40. Monash University. 2017. "Research Data Management at Monash." <https://www.monash.edu/library/researchdata/about>.
41. Monash University. 2017. "Professional Development." In Graduate Research. <https://www.monash.edu/graduate-research/future-students/phd/professional-development-option>.
42. Monash University. 2017. "Research Data Skills Development." In Library: Managing Research Data. <https://www.monash.edu/library/researchdata/skills>.
43. Monash University. "Resources and Activities." In Library: Managing Research Data. <https://www.monash.edu/library/researchdata/resources>.
44. Monash University. "Case Studies: Research Data Management Policy Implementation Project." In Library: Managing Research Data. <https://www.monash.edu/library/researchdata/resources/case-studies>.
45. See note 42.
46. Monash University. "Advice and Planning." In Library: Managing Research Data. <https://www.monash.edu/library/researchdata/advice>.
47. *ibid.*
48. Australian National Data Service (ANDS). <http://www.ands.org.au/about-us>.
49. ANDS. "Trusted Partnerships." <http://www.ands.org.au/about-us/what-we-do/trusted-partnerships>.
50. Monash University. "Research Data Management at Monash." In Library: Managing Research Data. <https://www.monash.edu/library/researchdata/about>.
51. ANDS. "Research Vocabularies Australia." In Online Services. <http://www.ands.org.au/online-services/research-vocabularies-australia>.
52. See note 50.
53. *ibid.*
54. Times Higher Education World University Rankings 2016-2017. <https://www.timeshighereducation.com/world-university-rankings/2017/world-ranking>.
55. QS Top Universities. Wageningen University. <https://www.topuniversities.com/universities/wageningen-university>.
56. National Coordination Point Research Data Management (LCRDM). <https://www.surf.nl/en/lcrdm/about-lcrdm>.
57. Vereniging van Universiteiten (VSNU) [Association of Universities in the Netherlands]. 2014. *De Nederlandse Gedragscode Wetenschapsbeoefening: Principes van goed wetenschappelijk onderwijs en onderzoek*. [The Dutch Code of Conduct of Science: Principles of Good Scientific Education and Research] DEN HAAG: VSNU. [http://www.vsnu.nl/files/documenten/Domeinen/Onderzoek/Code_wetenschapsbeoefening_2004_\(2014\).pdf](http://www.vsnu.nl/files/documenten/Domeinen/Onderzoek/Code_wetenschapsbeoefening_2004_(2014).pdf), 8. [In Dutch]
58. See Jhattacharjee, Yudhijit. 2013. "The Mind of a Con Man." *The New York Times*. 26 April. <http://www.nytimes.com/2013/04/28/magazine/diederik-stapels-audacious-academic-fraud.html?pagewanted=all>; and Cole, Graham D., and Darrel P. Francis. 2014. "Cole And Francis on Delivering Reliable Science and Guidelines after Poldermans Fraud." *European Heart Journal*. 35(36): 2435-40. <http://www.sin-nl.org/cole-francis-delivering-reliable-science-guidelines-poldermans-fraud/>.

59. Wageningen University and Research. "Data Management Planning."
<http://www.wur.nl/en/Expertise-Services/Data-Management-Support-Hub/Browse-by-Subject/Data-Management-Planning-1.htm>.
60. Ibid.
61. 4TU.Centre for Research Data. "Essentials 4 Data Support Course." In Training and Events.
<http://researchdata.4tu.nl/en/training-events/>.
62. Wageningen University and Research. Manage Your Source Code with Git@WUR."
<http://www.wur.nl/en/Expertise-Services/Data-Management-Support-Hub/Browse-by-Subject/Manage-your-source-code-with-GitWUR.htm>
63. Wageningen University and Research. "Data Management Support Hub."
<http://www.wur.nl/en/Expertise-Services/Data-Management-Support-Hub.htm>.
64. Wageningen University and Research. "Data Management Support Services."
<http://www.wur.nl/en/Expertise-Services/Data-Management-Support-Hub/Data-Management-Support-Services.htm>.
65. Wageningen University and Research. "Publishing Your Dataset in a Repository."
<http://www.wur.nl/en/Expertise-Services/Data-Management-Support-Hub/Browse-by-Subject/Publishing-your-dataset-in-a-repository.htm>.
66. See note 64.
67. See note 63.
68. See note 65.
69. Ibid.
70. 4TU.Centre for Research Data. "Costs." <http://researchdata.4tu.nl/en/about-4turesearchdata/costs/>.
71. See note 62.
72. Wageningen is currently investigating the possibility of acquiring local solutions for active data management.

For more information about OCLC Research's work on research data management, please visit: oclc.org/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: 978-1-55653-023-4
DOI: 10.25333/C3Z039
RM-PR-215833-WWBE 1709