Case Studies in Capacity Acquisition: institutional strategies for sourcing Research Data Management (RDM) services

Constance Malpas, OCLC Research
Research Data and the Evolving Scholarly Record

• Growing attention to research data as a ‘first-class’ research output, requiring long-term stewardship and discoverability

• Persistent identification and preservation necessary to support reproducible research

• Digital research environment increasing the volume and diversity of research data products
Library roles in Research Data Management (RDM)

• Training and **education** on the importance of data management, compliance with funder Data Management Plan and open data requirements

• Guidance on appropriate **disciplinary repositories** for discoverability and preservation

• Metadata support for **deposit** of data sets

• Selection, **curation** of research data in local repositories
Project Team

Brian Lavoie
Research Scientist

Rebecca Bryant
Senior Program Officer
Research Library Partnership

Constance Malpas
Research Scientist
Strategic Intelligence Manager
Scholarly record evolving alongside research practices

- from print-centric to **digital** and networked
- from books and journals to **data sets**, computer models, presentations...
- from concentration in library collections to **distribution across the network**

http://oc.lc/esr
Framing the Scholarly Record …
Research Questions

• What are the primary incentives that motivate university investment in RDM services?
• What factors determine the scope of RDM service provision in different institutional settings?
• How are services sourced to meet institutional needs? What informs the build, buy or partner decisions?
Case Study Approach

• Detailed examination of local circumstances and institutional decision-making
• Insight on how national and institutional contexts shape local service portfolios
• Leverages OCLC membership and Research Library Partnership
Case Study Partners

Public research universities with global reputation

- University of Edinburgh (UK)
- University of Illinois (US)
- Monash University (Australia)
- Wageningen University (Netherlands)
Project Timeline

- **Jul 2016 - Sep 2016**: Develop framework (Complete)
- **Oct '16 - Feb 2017**: Document case studies (Complete)
- **Mar - Aug 2017**: Publish report series (In progress)
- **Sep 2017**: Community engagement (Planning)
RDM Service Categories

**EDUCATION**
Tools and services to build researcher awareness of good RDM practice

**CURATION**
Technical infrastructure and services supporting RDM throughout entire research lifecycle

**EXPERTISE**
Decision support and customised solutions for specific RDM needs
Sourcing: a spectrum of choices

RDM LibGuides
Externally sourced workshops (Data Carpentry etc.)
MANTRA MOOC

Specialist data librarians
Data Curation Centre, 4TU Data Centrum
DataQ

Local research data repository
Metadata support for deposit in external repositories
Consortium or commercial repositories
RDM expertise, innovation a university differentiator
RDM services reduce administrative burden on university researchers.

Research Data Service

Catalyzing management of research data for efficiency today and accessibility tomorrow.

Contact us

Need to publish research data for a project, publication, or grant?

ILLINOIS DATA BANK

A secure file-based data repository for Illinois research. Deposits are professionally managed and curated by Research Data Service staff at the University Library.
RDM featured as a core library competency

Managing research data

Data collection, storage and dissemination

Benefits of data planning
Find out more

Data management at Monash University
RDM hub is a central university service

Data Management Support Hub

Where questions and answers meet

On the Data Management Support Hub you find a growing collection of practical guidance, support and advice for managing your research data.
University of Edinburgh:
RDM Service Offering Profile

Overview
Research data management services at the University of Edinburgh are housed within the Information Services unit, which includes the library, information technology, study spaces, and learning technology services. RDM-related services are offered through the Research Support sub-unit, and are divided into two areas:

- **Research Data Management:** “Find out how to create data management plans, where to store your data during your project, and how to share and deposit your final data at the end of your project.”
- **Data Library and Consultancy:** “Help researchers discover, use, manage and learn about datasets for analysis, learning and teaching.”

The homepage for RDM services is at:
http://www.ed.ac.uk/information-services/research-support/data-management

A description of the Research Data Service is available in an online booklet:

Leadership/staff: Dominic Tate and Jeremy Upton

History and milestones

![Timeline of RDM Service Offering Profile](image)

Service Offerings Report
Monash University

Australian national context
In July 2012 the National Health and Medical Research Council (NHMRC) announced its policy on the dissemination of research findings, which requires publications resulting from NHMRC supported research to be deposited in an open access institutional repository and/or made available in another open access format within twelve months of initial publication. In January 2013, the Australian Research Council (ARC) implemented a similar Open Access Policy. Neither of these policies specifically mentions the public dissemination of datasets. The ARC does have a separate statement on Research Data Management, which states that researchers “have a responsibility to consider the management and future potential of their research data.”

In 2007 the Australian government released the Australian Code for the Responsible Conduct of Research, which states that it “considers data management planning an important part of the responsible conduct of research and strongly encourages the depositing of data arising from a Project in an appropriately publicly accessible subject and/or institutional repository.” Since 2014 ARC requests that PIs outline plans for the management of data in project proposals, but ARC does not mandate open data. NHMRC issued a similar statement in 2015, encouraging but not mandating the sharing and reuse of research data sets.

The Australian National Data Service (ANDS) was formally established in 2008 as a partnership led by Monash University in collaboration with the Australian National University (ANU) and the Commonwealth Scientific and Industrial Research Organization (CSIRO). It is funded by the Australian Government through the National Collaborative Research Infrastructure Strategy (NCRIS) in order to collaboratively support the management, preservation, and dissemination of Australian research data. ANDS provides resources to educate researchers and librarians on the development of data management plans.

Key information
The Monash University Library provides education, advice, and repository infrastructure for data management, deposit, and dissemination. It works collaboratively with partners on campus and across Australia, including the Monash eResearch Centre, Monash eSolutions (IT support) and the Monash Research Office.

- [https://www.monash.edu/library/researchdata](https://www.monash.edu/library/researchdata)
- Leadership: David Groenewegen, Director, Research, Monash University Library, david.groenewegen@monash.edu

History and milestones

![Timeline of Service Offerings Report](image)
Service Offerings Report
University of Illinois at Urbana-Champaign Research Data Service

US National context
In the United States, the National Institutes of Health (NIH) began requesting data sharing plans from PIs submitting proposals in excess of $500,000 in October 2003. The National Science Foundation (NSF) followed in Spring 2010 when it announced that it, too, would begin requiring research data management plans (DMPPs) in future grant proposals. As a result, there has been an emphasis in the US on the development and sharing of data management plans as a part of the grant proposal and funding apparatus. The DMPTool, a free service to help researchers and institutions to create high-quality data management plans, was developed through collaborative community effort and first released in October 2011.

In February 2013, the White House Office of Science and Technology Policy released a public access memo, directing Federal agencies supporting research to develop a plan to support increasing public access to publicly-supported research. Specifically, the OSTP memo called for the public availability of federally funded research outputs following a twelve month post-publication embargo period. Both peer-reviewed publications and datasets were specifically included in the memo. While NSF and NIH are the two largest research funders in the United States, there are numerous other federal agencies providing research support. With each federal agency responding on its own timeline with its own public access plan, a complex landscape of differing requirements and systems has emerged. University libraries, research offices, as well as individual researchers work to monitor this complex and changing landscape. SPARC and Johns Hopkins University Libraries have developed a community resource to aggregate this information.

Key information
The University of Illinois Research Data Service (RDS), http://researchdataservice.illinois.edu/, is housed in the Illinois University Library, in partnership with several other campus units, including the Office of the Vice Chancellor for Research, Provost, School of Information Sciences, Technology Services, and the National Center for Supercomputing Applications (NCSA).

- RDS leadership: Heidi Imker, PhD, Director imker@illinois.edu

History and milestones

RDM Service Offering
Wageningen University and Research Centre (WUR)

Institutional research data management policy

Excerpt of institutional policy quoted here:

“Wageningen Graduate Schools (WGS) requires that PhD candidates create a Data Management Plan for their research. Chair groups are also required to have a Data Management Plan for their group.

Once PhD candidates have created their DMP, it becomes an appendix to their research proposal and may be reviewed by Wageningen Graduate Schools. Each graduate school has slightly different requirements on how to submit the DMP with other documents.”

Primary URL for RDM services


RDM leadership
Jacqueline Rengersma, Head of Digital Production Centre

History and Milestones

OCLLC Research 1 12 October 2016
Selected Findings

• Institutions are implementing a mix of locally-developed, cooperatively-sourced and commercially-licensed solutions

• Benefits and tradeoffs of outsourcing services vary with institutional context and culture

• There is shared concern about sustainability of locally developed RDM capacity, given diverse disciplinary needs and uncertainty about long-term costs of large-scale storage

• All participants see increasing need for services around ‘active data management’ and researcher workflow support
Key Insight

• Strategic choices about which RDM services to internalize or externalize can increase institutional differentiation, e.g.

  Edinburgh’s reputation for data-driven research and innovation is reinforced by internalizing most EEC components

  Wageningen has externalized curation to national and group providers to focus on local expertise services
Forthcoming reports

I. Overview of RDM services and orientation to Education, Expertise, Curation framework

II. Scope of RDM services at our case study sites

III. Incentives motivating RDM service development

IV. Factors influencing choices about procurement (source), deployment (scale) of RDM services
What’s Next?

• First report to be published in March, others will follow this year
• Case study profiles will be shared as underlying data set for project
• Webinars featuring member stories from OCLC Research Library Partnership
Join the conversation

• Does your library offer RDM services in the Education, Expertise or Curation categories?

• Are you implementing any cooperatively-sourced RDM services? (E.g., shared Dataverse, consortial repository, DMP tools, formal expertise networks)

• What are the primary incentives driving your RDM service strategy? (E.g., compliance with funder mandates, opportunity to innovate, distinctive competency or infrastructure capacity)

• What have you stopped doing (or reduced) in order to support RDM?
Together we make breakthroughs possible.

Let’s stay in touch:

malpasc@oclc.org
@ConstanceM