From Project to Strategic Vision: Taking the Lead in Research Data Management Support at the University of Sydney Library

Belinda Norman
University of Sydney Library

Kate Valentine Stanton
University of Sydney

Abstract

This paper explores three stories, each occurring a year apart, illustrating an evolution toward a strategic vision for Library leadership in supporting research data management at the University of Sydney. The three stories describe activities undertaken throughout the Seeding the Commons project and beyond, as the establishment of ongoing roles and responsibilities transition the Library from project partner to strategic leader in the delivery of research data management support. Each story exposes key ingredients that characterise research data management support: researcher engagement; partnerships; and the complementary roles of policy and practice.

Introduction

The University of Sydney is Australia's oldest university. Established in 1850, it was built on the egalitarian values of the former penal colony, breaking with the traditions of Britain's ancient universities by admitting students on academic merit rather than on the basis of religion or social class[1]. Classes began in 1852 with 38 students and three academic staff,[2] and the phrase "research data management" was nowhere to be heard. Jump ahead to 2014 and the University is one of Australia's leading research institutions – with 51,394 student enrolments and 3,350 academic staff[3] across 16 faculties and schools. Research, teaching and learning takes place across nine campuses and is supported by the University of Sydney Library at 11 library locations. Increasingly "research data management" is a phrase that is discussed and understood by researchers, librarians, administrators, and students, and which describes policy, practices, services and job titles.

Throughout the University of Sydney's history, there is a single organisation that has endured continuously – the Library.. The establishment of ongoing research data management services "expand the role of the library in the knowledge creation process" (Tenopir, Birch, & Allard, 2012) and will continue to position the Library as a trusted partner in research and scholarly communication.

Seeding the Commons

The University of Sydney Library's research data management evolution has been fuelled largely by funding received from the Australian National Data Service (ANDS)[4]. Supported by the Australian Government, the ANDS mission is to enable the transformation of research data that are unmanaged, disconnected, invisible and single-use to structured collections that are managed, connected, discoverable and reusable.

> 'The purpose of this activity is to enable richer research, more accountable research, more efficient use of research data, and improved provision of data to support policy development' (Groenewegen & Treloar, 2013).

One of the key ways ANDS has delivered on this mission has been through the widespread funding of projects across the Australian research sector. In 2012 the University of Sydney Library began the ANDS-funded Seeding the Commons project, building on work undertaken by a number of ANDS-funded software development projects.

The Seeding the Commons project focused on embedding skills and services within the institution to support cultural change among researchers and the units that support them, rather than the development of technical systems. Project deliverables included:

- Raising the visibility of University of Sydney research data collections through the creation of 55 metadata records to be published to ANDS' national research data discovery service, Research Data Australia;[5]

- Creation of a research data management website;[6]

- Creation of a research data management planning checklist;[7]

- Creation of internal procedures documentation to embed the skills and knowledge developed through the project;[8] and

- Development and delivery of research data management training resources.

Seeding the Commons provided the Library with its first real opportunity to engage with research data management support, a major area of strategic development for academic libraries. For example, in the Australian context, 87.5% of surveyed academic libraries have current or planned guidance services for research data management (Corrall, Kennan, & Afzal, 2013). Library representatives had participated in previous ANDS projects and had been involved in cross-institutional activities around research data management, but Seeding the Commons gave the Library its chance to take the lead.

A guiding principle was established early in the project: the Library had to embrace the opportunities presented by this project to establish ongoing research data management support services for researchers. To achieve this goal it was essential that the project be integrated with the existing Faculty Liaison Librarian (FLL) network.[9] This integration served two purposes: it allowed the project team to leverage the existing robust relationships FLLs have with researchers, and allowed the FLL cohort to develop research data management literacy to support the development of post-project services.

The following three stories illustrate milestones, each occurring a year apart, in the evolution from a Library-led project to a strategic vision for Library leadership in supporting research data management. The first two stories describe activities undertaken during the Seeding the Commons project. The third describes work done to implement a University-wide research data management policy, post-project. Each story exposes key ingredients that characterise research data management support: researcher engagement; partnerships; and the complementary roles of policy and practice.

A Tale of Two Researchers:
Who are the Researchers; What do they Need;
What do they Want?

The scene: a small project team interview a researcher in the Department of Archaeology, surrounded by data gathered as part of her 20-year research career. She describes this data as the unique and irreplaceable records of shared human history – diagrams, maps and notebooks recording step-by-step discoveries as she uncovered layers of ancient history. Most of the data is in paper format. The researcher is very tech savvy and she labours away at the arduous process of digitising her life's work, but time and resources are lacking. In addition, this researcher has inherited legacy data

---

6　Research Data Management Guidelines: http://sydney.edu.au/research_support/data/

7　Research Data Management Planning Checklist: http://sydney.edu.au/research_support/data/data-management-planning.shtml

8　Toolkit and Resources: http://sydney.edu.au/research_support/data/toolkit-and-resources.shtml

9　Faculty Liaison Librarian means a specialist information professional that is assigned to a faculty, school or department to support and facilitate the research and teaching needs of that unit.

collections from colleagues in her department who have retired. Their data collections are also valuable and cannot be lost. And so, she says, the number of boxes grows year by year. This researcher's dataset represents the "long tail" of research data and highlights the issue of "legacy" data.

In the same week, the project team interview a researcher in the School of Geosciences who has been sharing his research data for decades and who has established data sharing as an ethos and business-as-usual research practice for the research group that he leads. The research group maintains a website[10] where companion datasets for published research papers are made available upon publication of the paper. Open source software for manipulation of the data is also available and readme documentation accompanies all shared resources. Data sharing has, in his view, led to increased citations to the published papers.

This story of two researchers shines a light on the central player in research data management activities – the researcher – and the diversity in research practice. Success in research data management support depends heavily on successful researcher engagement to ensure that services are researcher-centric and closely aligned with research practice. The following sections expand on the engagement activities of the project team, and the lessons learned from those conversations.

The Engagement Model

The Seeding the Commons project successfully engaged researchers through a model of collaborative engagement, developed in partnership with the FLL network. This collaborative model allowed the project to build upon the Library's long-established position as a trusted agency on campus.

The two researchers described above were participants in one of the major deliverables of the Seeding the Commons project – raising the visibility of research data collections at the University through the creation of 55 descriptive metadata records for publication to Research Data Australia. This deliverable depended upon finding researchers who were willing to be interviewed about their research data collections. The project team was not constrained by any requirements around how these researchers and collections would be selected. The decision was made to focus on researchers who were interested and to use the FLL network to help identify and engage with them.

The project communication plan focused on the Library to achieve the twin objectives of raising awareness of research data management among library staff and drawing upon the established connections that FLLs have within their client groups. The team visited individual faculty-based service teams to discuss the project within their faculty context. The project team engaged FLLs in a conversation about their client groups, asking: what do you know about the research activities of your client groups? Do you know of any researchers who are innovative and willing to try new things? Do you regularly participate in committee meetings where project aims could be communicated and researcher participants engaged? What are the best ways to reach your client groups?

These conversations resulted in the project team partnering with the FLLs to engage with researchers via an array of communication channels including:

- Presenting to researchers as part of research skills sessions offered by FLLs;

- Presenting to a number of faculty research committees;

---

10 EarthByte Group: http://www.earthbyte.org/

- Meeting with individual Associate Deans (Research); and

- Advertising the project via faculty newsletters and mailing lists.

As a result of these FLL-driven communications, responses were received from approximately 30 researchers from a range of research disciplines spanning the sciences, arts, health sciences, architecture and medical sciences.

The model of researcher engagement clearly branded Seeding the Commons as a Library project. Because of this association, the project team was able to build upon the relationship of trust that exists between the Library and academic communities, and to open the door for engagement with researchers. The positive relationships FLLs had with their clients and the depth of knowledge about their research milieu were evident during the data interview process. Ensuring that the FLL for a particular client group was an active participant during data interviews helped the project team and the researcher navigate unfamiliar subject matter.

Lessons Learned from Researcher Engagement

The engagement with researchers through structured data interviews allowed the project team to begin the work of identifying data generating behaviours, data support needs and the aspirations of researchers across a variety of disciplines. These interviews allowed the team to start conversations about data reuse and communicate the benefits and requirements of managing and sharing research data. The following four lessons summarise what was learned through these conversations:

1. Researchers universally face challenges in managing their research data

The archaeology researcher illustrates the types of challenges routinely faced by researchers in managing their data: conversion to digital formats, maintenance of digital data, digitisation of physical data, responsibility for legacy data, scarcity of resources, support services and expertise, and issues around the long term curation of research data.

2. Researchers support data sharing in principle

Every researcher we spoke to saw the benefits to knowledge creation offered by the sharing and reuse of research data. We spoke to researchers whose research practices routinely include data sharing. For example, a pioneering researcher in the School of Chemistry led a project in which all research data was published online as it was generated, and collaboration and contribution was invited from around the globe.[11]

3. The definition of data, research practices and the type of data management challenges vary greatly across disciplines

There can be no one-size-fits-all practice or technology with regards to data management and sharing, no one process that will suit every researcher. Geneticists have been sharing gene sequence data via Genbank[12] for 30 years and this has become accepted practice. However, researchers working with human participants face a range of issues related to data retention and sharing due to requirements from human research ethics committees, codes of research practice and legislation. Meeting the many and varied research data needs of researchers will be an ongoing challenge for institutions.

---

11 Open Source Drug Discovery: http://sydney.edu.au/research/spotlight/toddosdd.shtml
12 Genbank: http://www.ncbi.nlm.nih.gov/genbank/

4. Researchers want and value research data management support services

One of the questions posed to researchers during the data interviews was: "Once research outputs have been published and the retention period for your data has passed, what will become of your data? Will it be destroyed?" Not a single researcher said yes, which raises the question: how will this data be looked after? In their interviews, researchers were asked a final question: "If there were one thing you could have to support you in managing your data, what would it be?" Some said more storage, some said digitisation, some said time, and some said resources. Many said "we need people" and "we need support staff". One senior research leader said "we need a *librarian*."

Party in the Library:
Let's Celebrate How Far We've Come

The scene: a crowd gathers one warm spring afternoon in the Library's seminar and exhibition space to hear the Associate Dean (Research) of the Faculty of Health Sciences and the Deputy Vice Chancellor (Research) share stories about their own struggles with research data management over their research careers. One tells of spreading paper survey results out on the grass to dry after a flood, the other describes how much of her research data is now locked up in obsolete digital formats. The University Librarian and the Deputy Vice Chancellor (Research) speak of the importance of effective data management support and the necessity of partnerships in achieving that goal. The event has been organised to celebrate the launch of the best practice guidelines website[13] and data management planning checklist[14] created as part of the Seeding the Commons project, and is attended by researchers, associate deans (research), librarians, research support staff, ICT staff and ANDS representatives. Stories are shared and connections are established over champagne and canapés.

This story is significant because it expresses the shared commitment to research data management support by the Library, Research Portfolio and ICT. Research data management support requires collaborative approaches so that resources can be used effectively; the unique expertise of the Library, Research Portfolio and ICT can be harnessed, and the outcomes for researchers can be as simple and user-friendly as possible. The story also marks the completion of the Seeding the Commons project.

The Research Portfolio

Developing a strong partnership with the Research Portfolio was a high priority for Seeding the Commons, as this is the central research agency on campus providing strategic direction, policy, administration and support services for researchers. The project would not have succeeded without buy-in from the Research Portfolio. The Research Portfolio provided essential knowledge of the many institutional research processes and access to senior research committees. These relationships have been instrumental in assisting Seeding the Commons team members build a working knowledge of research development processes at an institutional level and beyond. This knowledge has been invaluable in enabling communication with researchers and a richer understanding of the current research landscape.

---

13 Research Data Management Guidelines: http://sydney.edu.au/research_support/data/
14 Research Data Management Planning Checklist: http://sydney.edu.au/research_support/data/data
   -management-planning.shtml

Information and Communications Technology (ICT)

The Seeding the Commons project team worked closely with project managers, faculty managers and business analysts in ICT to share information, gain access to researcher groups, explore and understand existing ICT support for research data management through capital infrastructure projects, and generally contribute ICT processes and perspectives into a developing understanding of research practices at the University. The Seeding the Commons project team also worked closely with ICT staff working on ANDS projects, attending sprint meetings for a range of data capture projects, and to provide advice on the creation of research data metadata. Project team members were invited to attend regular meetings with Faculty IT Managers to exchange information and develop a working knowledge of activities undertaken by ICT counterparts. This collaborative engagement with ICT was the first of its type for the Library.

Challenges and Opportunities

Working across multiple agencies was a challenging learning experience for the project team. The Library, ICT and Research Portfolio have their own distinct cultures, and much of the partnership development focused on articulating shared goals and establishing boundaries for departmental roles. The creation of strong partnerships across institutional units was one of the most valuable outcomes of the project, and will ensure that research data management support will continue to be informed by collaboration between the University's research and teaching support units.

Making it Local:
Realising a Post-Project Library Strategic Vision and Institutional Policy

The scene: gathered around a table in a light-filled meeting room overlooking Sydney's Royal Botanic Gardens, a group of researchers from the Sydney Conservatorium of Music are engaged in an animated conversation about the nature of research data for composers of creative works. This is the first time many of these researchers have applied the term "data" to their creative practice. Facilitating the conversation are former Seeding the Commons project team members, in new post-project research data management roles. Also involved in the discussion is the Conservatorium Library Manager and ICT Manager. The task at hand is guiding a group of researchers through the process of creating their own local research data management policy.

The scene above demonstrates research practice in meaningful conversation with research policy; it illustrates support staff from the Library, Research Portfolio and ICT working in partnership and engaging with researchers. It shows how the Library has strategically positioned itself to take up emerging opportunities to support best practice in research data management and guide the development and implementation of University policy.

The University's Research Data Management Policy[15] (the Policy) was created to meet the requirements of the Australian Code for the Responsible Conduct of Research (the Code), created by the two major national funding bodies, the Australian Research

---

15 University of Sydney Research Data Management Policy 2014:
    http://sydney.edu.au/policies/showdoc.aspx?recnum=PDOC2013/337

Council (ARC) and the National Health and Medical Research Council (NHMRC) (2007). The Policy was also created to help realise the benefits offered by effective management and sharing of research data. The Policy has three main headlines:

1. **Data sharing:** The Policy affirms the Code's encouragement to researchers: "Research data should be made available for use by other researchers unless this is prevented by ethical, privacy or confidentiality matters." (Australian Research Council & National Health and Medical Research Council, 2007);

2. **Data management planning:** The Policy mandates creation of a data management plan for all funded research projects; and

3. **Local policy:** All faculties must translate the requirements of the Policy into a local policy, termed "local provisions" in the University's Policy Framework.[16]

Local Provisions: Process, Outcomes and Responsibilities

Local provisions are binding statements of principles that apply to a particular unit, school or faculty, rather than the University as a whole. They serve to operationalize higher-level policy in a way that is responsive to local priorities, practices and infrastructure. As part of the implementation of the Policy, a facilitated process to develop local provisions was undertaken with two pilot faculties – Engineering and Information Technologies and the Sydney Conservatorium of Music.

Led by the Research Portfolio's Policy Officer, in collaboration with the Library's Research Data Manager, the pilot involved regular meetings with a working party made up of faculty researchers and the Associate Dean (Research). The aims of the project leaders were to communicate the underlying principles of managing and sharing research data, suggest ways that the faculties could establish processes in support of those principles, and gain a deeper understanding of research practices in those faculties.

After an unstructured, discursive approach was undertaken with the first pilot group, the project leaders decided to try a more structured approach with the second group. A local provisions document template was created that mirrored the major sections of the Policy. The template document was then broken down into sections and organised around key data management concepts, such as "metadata and documentation." Each section of the draft was introduced to the working party through a presentation that defined the concepts, benefits and requirements within that section. This approach allowed researchers and the project team to come to a shared understanding of the topics discussed. Alongside the drafting of the local provisions document, the presentations enabled fruitful discussion and provided training in research data management literacy.

Both faculties involved in the pilot established ground-breaking requirements for their researchers that went beyond the requirements of the Policy. For example:

- Research data management planning will be mandatory for all higher degree by research students and all internal grant applicants;

- Data sharing – all higher degree research students and researchers leaving the University are required to have a record created in the Library's Research Data Registry;

---

16 University of Sydney Policy Framework: http://sydney.edu.au/legal/policy/university/framework.shtml

- Data management training will be integrated into existing research methods training for higher degree research students.

The creation of local provisions also provided an opportunity to articulate research data management support roles and responsibilities within those faculties, in which the Library will take a leading role.

Library Strategic Vision for Ongoing Roles

Upon completion of the Seeding the Commons project, the Library committed to leadership in research data management support through the creation of a full-time continuing role for the project's Research Data Manager. This role is responsible for the development of a Library strategic vision for research data management leadership, to take up the opportunities presented by the implementation of the Policy, local provisions and other initiatives.

The strategic vision establishes four key themes:

1. Research data management planning

Goal: Support researchers to meet the requirements and realise the benefits of research data management planning, as required by the Policy.

2. Data management training

Goal: Support robust research data management practices through strategic research data management training support.

3. Research data curation

Goal: Enable the management and discovery of University of Sydney research data outputs. This goal involves the Library taking business ownership and service development responsibilities for the metadata registry.

4. Advocacy and awareness

Goal: Raise awareness about and facilitate engagement with research data management practices. This goal supports the Policy's data sharing recommendation.

The strategic vision enabled the Library to clearly articulate Library roles within the local provisions development process. These roles map to the four themes framed above and are being established partly as FLL roles and partly as a role to be played by the specialist Research Data Manager position.

Library service team leaders who participated in the local provisions development pilot projects were asked to share their experiences at an information forum for FLLs about the Policy and local provisions development activities. The Team Leader, Faculty of Engineering and Information Technologies described the work of the local provisions working party as a true collaboration between the Library and Faculty. The Conservatorium Library Manager said that in his long professional career, this was the first time he experienced the Library engaging and collaborating with Faculty as equals.

Conclusion

Across the span of two and a half years research data management support in the Library has progressed from a small-scale project to a strategic vision for leadership. The Library has successfully taken up the opportunities presented by the Seeding the Commons project to embed research data management services in FLL roles and to take a significant leap toward positioning the Library as a trusted partner in research. Many challenges remain as the work of implementing the strategic vision begins: to maintain partnerships; continue to effectively engage researchers; and identify and develop the resources and skills required to meet the heterogeneous research data support needs of contemporary researchers.

Acknowledgements

References

Australian Research Council & National Health and Medical Research Council. (2007). *Australian code for the responsible conduct of research*. Retrieved from http://www.nhmrc.gov.au/_files_nhmrc/publications/attachments/r39.pdf

Corrall, S., Kennan, M.A., & Afzal, W. (2013). Bibliometrics and research data management services: Emerging trends in library support for research. *Library Trends*, *61*(3), 636–674. doi:10.1353/lib.2013.0005

Groenewegen, D., & Treloar, A. (2013). Adding value by taking a national and institutional approach to research data: The ANDS experience. *International Journal of Digital Curation*, *8*(2), 89–98. doi:10.2218/ijdc.v8i2.274

Tenopir, C., Birch, B., & Allard, S. (2012). *Academic libraries and research data services: Current practices and plans for the future.* Retrieved from Association of College and Research Libraries website: http://www.ala.org/acrl/sites/ala.org.acrl/files/content/publications/whitepapers/Tenopir_Birch_Allard.pdf