
Copyright 2011 The Authors
Getting research data ‘out there’: Collaborative solutions to identifying, describing and making research data more visible.

Abstract

At QUT research data refers to information that is generated or collected to be used as primary sources in the production of original research results, and which would be required to validate or replicate research findings (Callan, De Vine, & Baker, 2010). Making publicly funded research data discoverable by the broader research community and the public is a key aim of the Australian National Data Service (ANDS). Queensland University of Technology (QUT) has been innovating in this space by undertaking mutually dependant technical and content (metadata) focused projects funded by ANDS.

Research Data Librarians identified and described datasets generated from Category 1 funded research at QUT, by interviewing researchers, collecting metadata and fashioning metadata records for upload to the Australian Research Data commons (ARDC) and exposure through the Research Data Australia interface. In parallel to this project, a Research Data Management Service and Metadata hub project were being undertaken by QUT High Performance Computing & Research Support specialists. These projects will collectively store and aggregate QUT’s metadata and research data from multiple repositories and administration systems and contribute metadata directly by OAI-PMH compliant feed to RDA. The pioneering nature of the work has resulted in a collaborative project dynamic where good data management practices and the discoverability and sharing of research data were the shared drivers for all activity. Each project’s development and progress was dependent on feedback from the other. The metadata structure evolved in tandem with the development of the repository and the development of the repository interface responded to meet the needs of the data interview process. The project environment was one of bottom-up collaborative approaches to process and system development which matched top-down strategic alliances crossing organisational boundaries in order to provide the deliverables required by ANDS.

This paper showcases the work undertaken at QUT, focussing on the Seeding the Commons project as a case study, and illustrates how the data management projects are interconnected. It describes the processes and systems being established to make QUT research data more visible and the nature of the collaborations between organisational areas required to achieve this. The paper concludes with the Seeding the Commons project outcomes and the contribution this project made to getting more research data ‘out there’.
Introduction

Data is at the heart of any research activity. It is generated or collected, sorted, organised and structured. It is mined, analysed, combined, compared, re-analysed, and must be kept for the validation of research findings. It has intrinsic value as the building blocks of new knowledge. Very often it offers additional, temporal value such as when the data represents a snap-shot in time that can never be repeated. The creation and collection of data often represents a considerable investment of time and research funds and it must be properly managed. In an ideal world, research datasets would be highly visible and discoverable to the global research community through searchable interfaces via quality metadata records.

An objective of the Australian National Data Service (ANDS) is to establish an Australian Research Data Commons (ARDC), with its accompanying discovery portal Research Data Australia. The aim is to “support the discovery of, and access to, research data held in Australian universities, publicly funded research agencies and government organisations for the use of research.” (Australian National Data Service, 2010b). Research Data Australia is a “mesh of searchable web pages describing (and where possible linking to) Australian research data collections." (Australian National Data Service, 2010a).

Queensland University of Technology has contributed to this national objective through several ANDS funded and internally funded projects. These projects focus on the integration of the research generation space and repository space, on the automated capture of metadata, the aggregation of metadata and the discovery and description of research datasets at QUT, through to contributing metadata to Research Data Australia.

The university is benefiting not only from the goal of making QUT research data more visible, but also from the evolution of a culture of research data management.

This paper showcases the work undertaken at QUT, using the Seeding the Commons project as a case study to:

- Describe the processes and systems being established to make QUT research data more visible.
- Illustrate how the data management projects are interconnected.
- Discuss the evolving nature of collaboration between organisational areas at QUT required for project success.
Note the outcomes of the Seeding the Commons project.

1.1 The Data Management, Discovery and Reuse Cycle

The data management cycle is illustrated in Figure 1. This vision guides QUT’s work in the dataset management space.

The first half of the cycle is the domain of organising, creating and collecting research data. This part of the cycle includes the processes for: the storage of data; the registration (organisational identification) of data; the description of data (creation of metadata records using automatic and manual methods); and the exposure of the metadata and research data to discovery mechanisms dependant on established and access rights.

Once the data is found and well described the second part of the cycle - discovery and reuse - can start. This includes the steps of: data discovery; data evaluation (through review of the metadata records); accessing the data (directly or through negotiated access with the data custodian) and potential reuse of the data. Should the reuse of data result in the creation of a derived dataset, the cycle would start again.

This simple diagram illustrates how the processes are interconnected. Good data management in the first half of the cycle can bring this valuable research data into the light. Poor data management keeps it in the shadows.
1.2 QUT’s work in this space

QUT has been innovating in the first half of this cycle, through projects funded by ANDS under the National Collaborative Infrastructure Strategy (NCRIS) and Education Investment Fund (EIF) as well as internally funded projects. These projects are known as the *Seeding the Commons* and *Metadata Exchange Hub* projects respectively. The development of a *Research Data Management Service* is an internally funded project.

In a large environment such as a University, services supporting research can sometimes be fractured across different departments and areas of expertise. The two ANDS funded projects mentioned above spanned every facet of the data management, discovery and re-use cycle. Because of this, the projects pulled together groups of experts who, while part of the same Division (the Division of Technology, Information and Learning support – TILS), had traditionally operated independently in service provision: the High Performance Computing and Research Support unit (HPC & RS) and library liaison and researcher support services.

The projects were discreet on paper, but highly inter-dependent in practice. Deliverables relied upon close cooperation between staff from the Library and HPC & RS. Research Data Librarians were employed for the *Seeding the Commons* project, and they collaborated closely with HPC & RS specialists. Both of these groups were also reliant upon cooperation with the university’s Office of Research, which provides services around grant management, research administration and ethics approvals.

A collaborative approach to managing research support services was modelled top-down by managers. A Research Data Management Steering Committee includes stakeholders from across divisional areas, and oversees all of the Research Data Management projects within TILS to ensure a cohesive approach.

This shared approach reflects the QUT University’s policy to guide and encourage good data management, which is a shared responsibility for all involved with research data, from support services staff and managers, through to research directors and the individual researchers themselves. (Queensland University of Technology, 2010)
1.3 Project Overviews

Seeding the Commons: Project Overview

The objective of the QUT Seeding the Commons project was to identify QUT research data associated with Australian Research Council (ARC) and National Health and Medical Research Council (NHMRC) funded research. Once identified, metadata was captured and stored and RIF-CS compliant metadata records created. A test set of these metadata records were then made available to ANDS for display in Research Data Australia. Overall the project assisted QUT in developing the capability to better manage its research data assets and contribute to a change in the culture around research data management and data sharing. This project was managed by two Research Data Librarians, with technical input and software development and configuration support from HPC & RS specialists.

Metadata Exchange Hub and Metadata Store: Project Overview

The Metadata Exchange Hub project, or simply the Metadata Hub project, focussed on the construction and deployment of a metadata aggregator. It was undertaken concurrently with the Seeding the Commons project and involved a) the collection and analysis of business requirements b) the choice and evaluation of software as a foundation on which to build the metadata aggregator, and c) the customisation and deployment of the chosen software.

The software that was chosen was the semantic web application VIVO, the development of which had recently attracted a lot of funding in the United States through a National Institutes of Health (NIH) grant. The VIVO software was customised and extended to meet the requirements of the Metadata Hub project. The project was undertaken in collaboration with Griffith University and at time of writing is in its final stages.

The metadata hub will initially ingest much of its data from QUT’s Data Repository which uses the XML object database called Mediaflux.

QUT’s Data Repository and Research Data Management Service: Project Overview

Development of QUT’s Data Repository is being undertaken as the first part of an internal project which will lead to the establishment of a Research Data Management Service. This development work is being undertaken by HPC & RS staff. This includes development of a website to support access to the service and development of the data repository and interfaces for researchers to deposit and access data.
1. Case Study: Seeding the Commons

The *Seeding the Commons* project, as a case study, will:

- Describe the processes and systems being established to make QUT research data more visible.
- Illustrate how the data management projects are interconnected.
- Discuss the evolving nature of collaboration between organisational areas at QUT required for project success.
- Note the outcomes and the contribution this project made to getting more research data ‘out there’

This study is from the perspective of the Research Data Librarians appointed directly to this project.

2.1 Processes

There were several key processes to this project. Many processes were inter-dependent and relied upon the progress and processes within other projects underway at QUT.

The processes for the *Seeding the Commons* project were:

- Identifying required metadata elements and sources of metadata, mapping to RIF-CS and preparing existing metadata for ingest to the Mediaflux repository
- Obtaining additional metadata: Data interviews, publications and research
- Creating metadata records
- Building the user interface
- Feeding metadata records to Research Data Australia

2.1.1 Identifying required metadata elements and sources of metadata, mapping to RIF-CS and preparing existing metadata for ingest to the Mediaflux repository

The *Registry Interchange Format - Collections and Services* (RIF-CS) metadata schema is prescribed for use with the Australian Research Data Commons (ARDC). Knowledge sharing about RIF-CS was one of the initial catalysts for tight collaboration between the Research Data Librarians and specialists from HPC & RS.
RIF-CS is “is an object-oriented, relational model” where several registry objects are described and the relationships between the objects defined. (Australian National Data Service, 2009)

Therefore the definitions of the registry objects as used by the project, and the relationships, determined the essential metadata elements required. Certain working definitions were decided, for use within *Seeding the Commons* project, and are illustrated in Figure 2, below:

![Figure 2 – Seeding the Commons Working Definitions of RIF-CS Objects](image)

The HPC & RS Specialists trained the Research Data Librarians in XML schemas to a working knowledge level, and discussed technical solutions for the input of metadata to the metadata repository (Mediaflux). This skilled the Research Data Librarians to deconstruct the RIF-CS Schema and understand the requirements for collecting and contributing metadata to the internal repository built on Mediaflux and from there to the ARDC. This process informed where to look for existing data within the university’s systems, and what unique metadata would need to be collected by interview.

In return, this collaboration informed the HPC & RS team regarding what metadata that the Mediaflux repository would need to store for the purposes of a RIF-CS feed to the ARDC, as
well as additional metadata that would be collected to enrich the records and be held locally only.

Reusing metadata already available within existing systems was a priority. The success of sustaining a future process would be dependent upon efficient workflows, elimination of duplication of effort, and data accuracy and integrity.

Within QUT’s Division of Research and Commercialisation, the Office of Research uses the Research Master system to manage research activity and grant processes. It is the primary source of pre-existing authoritative metadata describing research grants (Activities in RIF-CS) and researchers (Parties in RIF-CS) for the Seeding the Commons project.

There were initial discussions between Research Data Librarians and the Senior Systems Support Officer in the Office of Research about the sought-after metadata. This request to use Research Master data in such a way was novel, sparking another ground-up collaboration which influenced top-down management relationships.

The Research Master system contains substantial confidential data and so it was necessary to negotiate what access was required and what purposes the data would be used for. This triggered higher level, cross-Divisional discussions between senior managers from the Library and the Office of Research. This top-level collaboration and endorsement of internal data sharing will be a major contributor to an effective Research Data Management Service.

A condition of use of Research Master data was that any record to be made public through Research Data Australia would be first reviewed and approved by the Office of Research. This condition was later built into a draft workflow which prepared records for the feed to the ARDC.

A custom report was written by the Senior Systems Support Officer, exported as a text file and imported into Excel for review. This file was used later by ingesting the data into Mediaflux to create the stub records for Party and Activity objects.

Mapping of metadata to RIF-CS began with the Party and Activity data exported from Research Master. A complex mapping table was used to match data elements from Research Master to the RIF-CS schema, and identify the mandatory and optional data elements. This process also identified the data which was required to fashion a record to the minimum standard required by ANDS for the ARDC, as well as what was ‘missing’ from the Research Master report and would have to be gathered by other means.
There was a large amount of additional metadata exported from Research Master which was additional to the basic requirements of a RIF-CS record. This data formed the basis of the extension of the schema for local purposes and, simultaneously, the configuration of the Mediaflux metadata repository which was in the early stages of development.

Two activities – Seeding the Commons and work on Mediaflux metadata repository development and deployment – were now progressing in tandem, with one dependent upon the other. This is what it felt like:

Research Data Librarians and HPC & RS Specialists then co-developed workflows for metadata capture and repository which are illustrated in a simplified form by Figure 3.
2.1.2 Obtaining additional metadata: Data interviews, publications and research

Not all the metadata needed was exported from Research Master. The RIF-CS Collection Object and the additional information required could only be obtained by asking the researchers about the data associated with their research projects.

The Research Data Librarians set up meetings - ‘data interviews’ – with the chief Investigator named in Research Master against each of the research activities in scope for the project. The interviews were loosely scripted, focussed on confirming existing information (Party and Activity) and filling in the ‘missing’ metadata (Collection).

A small number of pilot interviews were undertaken in late 2009 to test a data interview instrument and workflows and to get feedback from interviewees. Feedback covered researchers’ attitudes towards data management and data sharing, through to more technical issues relating to the storage and transfer of data. During the pilot phase a HPC &
RS specialist attended the data interviews. They observed the flow of the interview – which would later inform user interface design to Mediaflux – as well as the type of information volunteered during the interview and how it translated into metadata. Faculty Liaison Librarians were also involved in the process, contributing discipline knowledge, background on the researchers and attending the data interview when possible.

During the project, 195 interviews were conducted in connection with 424 research activities. From these interviews 492 datasets were identified.

Table 1 shows some of the typical metadata gathered for a Collection (dataset).

<table>
<thead>
<tr>
<th>Interview Topic</th>
<th>Typical Metadata collected</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description and coverage of dataset</strong></td>
<td>- Title of dataset, description of dataset</td>
</tr>
<tr>
<td></td>
<td>- Keywords (author and controlled vocabularies)</td>
</tr>
<tr>
<td></td>
<td>- Data source (instruments, surveys, existing data, simulations, variables, cohort size, demographics)</td>
</tr>
<tr>
<td></td>
<td>- Temporal and spatial coverage</td>
</tr>
<tr>
<td></td>
<td>- Investigators, Primary Data Contact, Who collected the data</td>
</tr>
<tr>
<td><strong>Documentation and file information</strong></td>
<td>- What documentation exists about the data (codebooks, data management plans)</td>
</tr>
<tr>
<td></td>
<td>- Type of data (digital or analogue, quantitative or qualitative, documents, video, image, audio, GIS)</td>
</tr>
<tr>
<td></td>
<td>- Format and file sizes</td>
</tr>
<tr>
<td></td>
<td>- Specialist software required</td>
</tr>
<tr>
<td></td>
<td>- Publications associated with data (journal articles, thesis)</td>
</tr>
<tr>
<td><strong>Access and location of data</strong></td>
<td>- Data retention &amp; Disposal</td>
</tr>
<tr>
<td></td>
<td>- Location of data</td>
</tr>
<tr>
<td></td>
<td>- Can data be shared (ethics, contracts, confidentiality, IP)</td>
</tr>
<tr>
<td></td>
<td>- Embargoes, copyright, ownership</td>
</tr>
</tbody>
</table>

Table 1 – Typical Metadata Collected in a Data Interview

Additional information could be gathered from publications directly related to the research grant. Researchers often preferred that the Research Data Librarians sourced information for the descriptive records as it was written in publication and reviewing publications.
provided extra, rich, ‘authoritative’ information. However publications on their own do not replace a data interview and researchers themselves were often required to identify the specific publications. There is currently no system to link publications and research datasets which might streamline this process.

2.1.3 Creating metadata records

To create complete Party, Activity and Collection records, data was compiled from a variety of sources – both digital and hard copy. The percentage of metadata gathered was about 50/50 digital and hard copy. Party and Activity data from Research Master was mapped to RIF-CS and ingested into Mediaflux en masse, with the predefined relationships between the two objects automatically created. In the future the Metadata Hub will play a part in the metadata aggregation process, and obtaining data directly from Research Master.

An excel spreadsheet was to hold some Collection data while the Mediaflux data repository was being developed. The spreadsheet was not structured for ingest, as the structure for the metadata repository was still evolving during the data interview process, and much of the detailed and anecdotal collection information remained in written hard copy files.

Once Mediaflux was configured, Collection Objects and relationships were created manually, by the Research Data Librarians through a user interface, This is the part of the metadata collection process which is difficult to automate in most cases, as it is unknown what datasets exist or how they are related to other objects and where it does exist this information is not available in digital form suitable for ingest.

Early on in the project, an additional ‘interview’ Object was identified as essential. This is not a RIF-CS object, but a local administration object, recording interview notes and any follow requirements. Every data interview undertaken has an Interview Object and relationships to existing Party and Activity objects are created. Finally, Collection Objects are created for any datasets identified, and relationships made between Interview, Party and Activity Objects.

RIF-CS, the Registry Interchange Format for Collections and Services, is focused around Collection Object However, the workflow of the Seeding the Commons project was driven by the Interview process which was focused on the Activity Object. Relationships in RIF-CS are made by using keys, which are preferably globally unique identifiers. The ANDS Identifier Service mints globally unique handles and the HPC & RS specialist worked with ANDS to establish a machine to machine service to mint a handle for QUT Party, Collection and Activity Objects which is used in combination with a local Mediaflux identifier for any records
which are fed to Research Data Australia to be. There are local identifiers, such as the grant number assigned by Research Master, and staff ID which are only held locally.

To obtain unique national identifiers for researchers (Party records), the Research Data Librarians worked with the National Library of Australia (NLA) in early stages of the Party Infrastructure Project which is also funded by ANDS. An XML file of QUT researcher and publication information was exported from QUT ePrints with the help of QUT Digital Repository staff. The NLA Party Infrastructure project used this information to create a sample of 50 Party identifiers which were included in the QUT Party records.

2.1.4 Building the user interface

The previous section makes reference to a user interface to Mediaflux which enabled manual creation of objects and relationships between objects within the metadata repository. The configuration of Mediaflux developed alongside the metadata gathering and record creation process of the Seeding the Commons project, with each activity informing the other. Each individual use case and workflow in the user interface was developed from the interview experience and the subsequent tasks of cleansing and transcribing the data. Research Data Librarians described what the interface had to do and in what order and the HPC & RS specialists built it from scratch, on demand.

User testing of the interface clarified assumptions about the workflow. The testing was ‘real’ in the sense that real data was entered and had to work in a ‘real’ way immediately, in order for the project to achieve one of its goals – a feed of quality records to the ARDC.

New versions of the interface were being released on an hourly basis in direct response to demands for functionality in the face of project deadlines.

A functional and robust user interface emerged from an intense development phase where new versions of the interface were being released almost hourly in direct response to demands for functionality in the face of project deadlines.

It was the investment of all parties in the end goal, good communication and a preparedness to negotiate and compromise, and not least because each team member felt part of a larger, collaborative project environment that this aspect of the project was highly successful.

Figure 4 is a screen shot of a Collection Object being created in the Mediaflux user interface.
2.1.5 Feeding metadata records to Research Data Australia

The internal data model of the system contained much more data than was required for the creation of RIF-CS objects. A program was written, however, whereby objects within the system could be serialised to RIF-CS XML. This made the preparation of RIF-CS data files for Research Data Australia relatively easy. The Java api and libraries for RIF-CS provided on the ANDS web site were used for this purpose.

The user documentation created for the user interface also contained the required mappings of fields for each object to RIF-CS and established the relationships between objects in the data repository that would need to be created.

The system uses a ‘trigger’ field in the Collection Object to flag that the record has been approved for publishing to Research Data Australia. Approval of the Collection Object automatically brings the Party and Activity records along with it. All records are exported in RIF-CS format with the appropriate relationships encoded. Other objects such as the Interview Object are not included in the feed, and where there is more than one Party Object related to a Collection, only the primary contact for the dataset is passed into the feed. This
was a local decision about what content should be fed to the ARDC and made visible in Research Data Australia. However relationships to additional Party Objects are held in Mediaflux, and can be included in the feed if this decision is reviewed.

### 3 Project Outcomes

The *Seeding the Commons* project was successfully completed in early July 2010, achieving its stated goal to Identify and Describe Category 1 Research Datasets at QUT.

- A sample of shareable datasets was written up as Party, Activity and Collection Objects, and the records successfully exported in RIF-CS format.
- A feed of this data was successfully uploaded to the Australian Research Data Commons and displayed through Research Data Australia.
- Processes were documented into workflows, and rich information was collected which will inform further research data management initiatives at QUT.

A copy of the final report can be accessed on the ANDS Partners site.

Across the 424 activities where data interviews occurred, 120 projects generated no new data, or the data belonged to another institution, or the project re-used existing data. Of the remaining 304 projects 492 unique datasets were identified and varying levels of descriptions obtained. These 492 datasets were classified as to their “shareability” and this is detailed below in Table 2. Of the research datasets identified 86% cannot be shared or are currently unable to be shared. A sample of the datasets which were suitable to share was successfully described and are discoverable through Research Data Australia.

<table>
<thead>
<tr>
<th>Shareability</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Online (already available via URL)</td>
<td>10</td>
</tr>
<tr>
<td>Mediated (datasets exist now, and are accessible via the primary contact – either digitally or in hard copy. Some may be available for unmediated access in future if a data repository was available)</td>
<td>40</td>
</tr>
<tr>
<td>Potentially (datasets exist but could only be made available at project completion, after publications, after ethics checks, contract checks or after further work was undertaken to provide contextual information, metadata and to make the files sharing ready)</td>
<td>103</td>
</tr>
<tr>
<td>Not-shareable (No access due to IP, ethics, partners data, confidentiality, patents, data collection and analysis still going, dataset requires too much work to make shareable, description not to go to RDA, researcher requested not to share)</td>
<td>339</td>
</tr>
</tbody>
</table>

*Table 2: Datasets Classified by “Shareability”*
There are many barriers to sharing research data, and rich anecdotal information was captured during the interview process. Those datasets which were not shareable now (but potentially shareable in the future) and those not-shareable at all (442 in total) were coded to try to define and quantify the barriers to sharing. Table 3 shows the count of codes across the 442 datasets and the % of datasets which were flagged with each code. More detailed discussion around the barriers to sharing research data were written into the Project Issues Report, and will inform the ongoing development of the Research Data Management Service.

<table>
<thead>
<tr>
<th>Issue</th>
<th>%</th>
<th>Barrier Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethics</td>
<td>46%</td>
<td>Where ethics approval or informed consent agreements lock data down</td>
</tr>
<tr>
<td>Contracts</td>
<td>21%</td>
<td>Contracts such as ARC Linkage stops data being shared, IP, patents</td>
</tr>
<tr>
<td>Project not yet complete</td>
<td>20%</td>
<td>Data may be shared but the situation needs to be reviewed when project is complete</td>
</tr>
<tr>
<td>Data Condition</td>
<td>14%</td>
<td>Data may be shareable but would require significant work to make it usable, or is '90%' complete</td>
</tr>
<tr>
<td>Publication Cycle</td>
<td>12%</td>
<td>All publications not complete, data has ongoing analysis value to research team which sharing would undermine</td>
</tr>
<tr>
<td>Location</td>
<td>11%</td>
<td>Data could be shared, but current location unknown, or in possession of person other than primary contact, or in possession of another institution</td>
</tr>
<tr>
<td>Data Source</td>
<td>11%</td>
<td>Where data belongs to a partner / or derived datasets from the primary source cannot be shared</td>
</tr>
<tr>
<td>Researcher</td>
<td>8%</td>
<td>Where the researcher just doesn't want to share</td>
</tr>
<tr>
<td>Confidentiality</td>
<td>7%</td>
<td>Data cannot be shared due to confidentiality of data or participants</td>
</tr>
<tr>
<td>Data Quality</td>
<td>4%</td>
<td>Researcher advises that the data is of very limited use, i.e., too small a sample, too out of date, too extremely specific to the researcher's context or field of research</td>
</tr>
<tr>
<td>Competitive advantage</td>
<td>2%</td>
<td>Not sharing as having the data is what gives the competitive edge</td>
</tr>
<tr>
<td>Lab books</td>
<td>2%</td>
<td>Data is in lab books</td>
</tr>
<tr>
<td>Embargo</td>
<td>1%</td>
<td>Data has an embargo before sharing can occur</td>
</tr>
</tbody>
</table>

Table 3 – Potential Barriers to Sharing Research Data
4 Conclusion

As a result of undertaking these projects, QUT is well placed to move forward in the area of research data management and in enabling the sharing and discovery of research data.

The similarity and timing of the projects proved to be a great opportunity for Divisional and other areas of the university to work collaboratively, exchange knowledge and expertise, and contribute actively to the evolving culture of research support at QUT.

Identifying, describing and making research data more visible is an enormous undertaking which cannot be undertaken by any one professional area in isolation. Skills and knowledge need to be shared and existing resources and processes need to be used to minimise duplication of effort in the collection of metadata. Collaboration is required across all areas of the University including High Performance Computing and Research Support; the Library; Digital Repository (ePrints), the Office of Research (Ethics, Publications, Commercialisation); and Information Technology Services. This wide and inclusive approach to developing research support services will ensure that the best solutions to identifying, describing, and making research data visible are achieved.

5 Acknowledgements

The authors would like to acknowledge the work of Yi Sun, High Performance Computing and Research Support Specialist, in the development of the Research Data Librarian interface to Mediaflux as well as other technical solutions. We would also like to thank Janne Lawson, Senior Systems Support Officer, Office of Research, for her assistance with exporting data from Research Master.

The Seeding the Commons project was supported by the Australian National Data Service (ANDS). ANDS is supported by the Australian Government through the National Collaborative Research Infrastructure Strategy (NCRIS) Program and the Education Investment Fund (EIF) Super Science Initiative.

6 References


