

# NATIONAL COLLABORATIVE RESEARCH INFRASTRUCTURE STRATEGY (NCRIS)

## FINAL REPORT

2013-14  
ANNUAL REPORT 6

ANDS PROJECT PARTNERS



# Contents

<b>AUSTRALIAN NATIONAL DATA SERVICE (ANDS)</b>	<b>1</b>	<b>6</b>	<b>CO-INVESTMENT</b>	<b>34</b>
		6.1	Access and Pricing	34
<b>1 PROJECT STATUS</b>	<b>3</b>	6.2	Project Co-Investment	34
1.1 Background	3	<b>7</b>	<b>PERFORMANCE INDICATORS</b>	<b>35</b>
1.2 Major Activities, Breakthroughs, Highlights, and Issues	5	7.1	KPI Report	35
		7.2	Progress over the Life of the ANDS Project	38
<b>2 ACTIVITIES UNDERTAKEN</b>	<b>7</b>	7.3	Overall Progress	39
2.1 Research Infrastructure	7	<b>8</b>	<b>AUDIT STATEMENT</b>	<b>40</b>
2.2 Seeding the Commons	8	<b>9</b>	<b>SUMMARY OF THE CONDUCT OF THE PROJECT</b>	<b>41</b>
2.3 Frameworks and Capabilities	13	<b>10</b>	<b>EXPECTED FUTURE USAGE OF THE INFRASTRUCTURE</b>	<b>43</b>
2.4 National Collections	19	<b>11</b>	<b>EXPECTED FUTURE TRENDS</b>	<b>44</b>
2.5 Project Office	22	<b>12</b>	<b>APPENDICES</b>	<b>45</b>
2.6 Promotion	23	12.1	Confidential Information	45
2.7 Risk Management	27	12.2	Progress against activities	45
		12.3	Risk Register	50
<b>3 PROGRESS AGAINST MILESTONES</b>	<b>28</b>			
3.1 Seeding the Commons	28			
3.2 Frameworks and Capabilities	29			
3.3 National Collections	30			
<b>4 DEVIATIONS FROM THE PROJECT PLAN</b>	<b>31</b>			
<b>5 FINANCIAL AND HUMAN RESOURCES</b>	<b>32</b>			

# 1. Project Status

## 1.1 Background

Research data is simultaneously becoming more voluminous, more complex, and more vital as the very nature of research changes. Research has become more investigative because it is possible to assemble significant data collections that enable much broader problems to be addressed. Thus it is essential that research data is managed, able to be assembled, connected to other data and thus used to address problems that may well be different to the reasons for gathering the data in the first place. The Australian Government recognised the significance of data as key infrastructure in supporting research excellence and research innovation when it established the Australian National Data Service (ANDS).

ANDS has been in operation since January 2009 as part of the NCRIS initiative. Its aim of enabling more researchers to reuse research data more often required establishing partnerships beyond ANDS, and this need is continuing to increase. In May 2010, the Australian Research Data Commons (ARDC) project was announced as an EIF funded Super Science Initiative, to be managed by ANDS. An agreed Project Plan was submitted in June 2009 and accepted in September 2009. Some activity in the NCRIS funded ANDS project was transferred to the ARDC project as a result. It was subsequently agreed that ANDS operations should be extended beyond June 2011, to June 2013. In October 2012, the Collaborative Research Infrastructure Scheme (CRIS) was announced, and ANDS received an additional \$3M to maintain minimal critical infrastructure. This led to a substantial modification of the submitted 2012-13 Annual Business Plan. Each of these changes had a high impact on the activities of ANDS, and the second change substantially affected the ANDS project as agreed in the 2012-13 Business Plan. As each of these changes occurred, ANDS

continued to manage the ANDS and ARDC projects together, as they are strongly co-dependent. This report describes activity taking place against the agreed 2013-14 Business Plan, subject to the agreed variation to the plan submitted in March that described the effect of the business plan variation.

At the time of the submission of the 2009-10 Business Plan, ANDS had four programs of activity:

- *Developing Frameworks* – the frameworks that will enable research data producing institutions to capture, manage and share research data.
- *Providing Utilities* – services that reduce the cost of capture and ease the task of discovery.
- *Seeding the Commons* – improving local data capture and populating the data commons.
- *Building Capabilities* – improving Australia's capability to manage its research data.

As a result of the EIF funded ARDC project, the NCRIS ANDS project was consolidated into two programs of activity:

- *Frameworks and Capabilities* – the frameworks that will enable research data producing institutions to capture, manage and share research data; and improving Australia's capability to manage its research data.
- *Seeding the Commons* – improving local data capture and populating the data commons.

The associated ARDC project has five programs of activity:

- *Data Capture* – an institutionally based program to automate the capture of data and metadata from instruments (broadly defined) in data intensive research.

- *Public Sector Data* – a program of making more public data collections visible and available through the ARDC.
- *Metadata Stores* – an institutionally based program that enables metadata to be stored coherently across an institution that supports data management, publishing, sharing and reuse.
- *ARDC Core Infrastructure* – an ANDS driven program that puts in place the national services that enable research data to be published and discovered (it is an expansion of the *Providing Utilities* program).
- *ARDC Applications* – a program that develops tools and services to support demonstrations of the value of exploiting data in the ARDC.

Two new programs were subsequently created in 2012:

- *National Collections* – an ANDS-driven, NCRIS-funded program partnering with institutions wishing to make National Collections available, and with RDSI and its nodes to help improve storage and access to those collection.
- *International Infrastructure* – a program designed to work collaboratively with international organisations and partners to ensure a more compatible international data-sharing environment for Australian researchers.

These programs were created in response to the changing environment, rather than a changing focus within ANDS. The advent of the RDSI initiative meant that there were new opportunities to assemble and make available collections of significance to researchers, to research disciplines, to research institutions and to the nation.

ANDS' relationship with institutions meant that ANDS could focus effort to ensure that more collections of strategic significance were available on RDSI nodes where the collections are managed, connected, discoverable and increasingly usable, through better access, and possibly with new forms of access.

The strong drive in Europe and the US in particular, but also more generally, to create an environment that enables data to be shared across boundaries, provided Australia with an excellent opportunity to engage internationally, particularly through the newly forming Research Data Alliance. This international initiative should enable Australian researchers to partner more effectively through a shared research data environment.

Figure 1 shows how the NCRIS programs complement and inter-relate to the creation of the Australian Research Data Commons.

Taken together, the intent of the two investments is to:

- Create an "essential meeting place where the Australian path forward for research data management can evolve and where a vision can be achieved." (*Towards the Australian Data Commons*), developed during 2007 by the ANDS Technical Working Group.
- Enable the following capability: "Research data and research outputs from all sources can be discovered and reused across disciplines and over time through an integration of repositories and data centres supporting national and specialist discovery services." (*Towards the Australian Data Commons*).

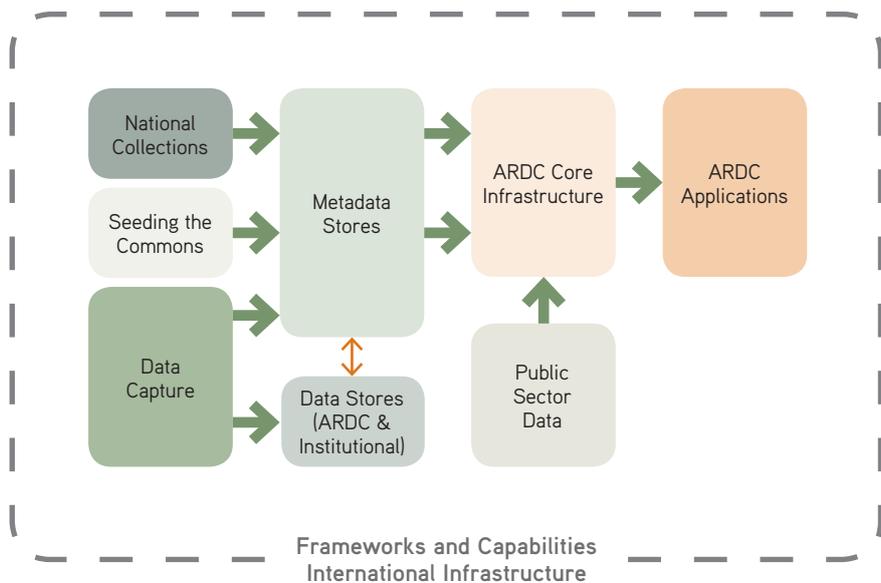


Figure 1: Relationship between Programs

- Create and populate the Australian Research Data Commons, which "will 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." (ARDC Investment Plan).

This ARDC investment statement can thus be seen as an intensification of effort in support of the second *Towards the Australian Data Commons* statement.

A consequence of the CRIS investment was that ANDS undertook a further change and simplification of its program structure. Its current programs are National Services, National Collections, Institutional Engagements and International Collaboration. This structure was used again to determine a response to the NCRIS 2013 Scheme that is beyond the scope of this report. However, whilst this simplification was gradually bought

into effect late in 2012-13, it was felt that as the bulk of the activity of ANDS was conducted under the structure depicted in Figure 1, and to ensure that the activities of ANDS can be described against the agreed Annual Business Plan, this report should use the structure depicted in Figure 1.

***This report is the final annual report of the NCRIS project. It stands alone, with an initial summary of the activity of both NCRIS and EIF project in the final year in the context of the overall project investment, before detailed description of program activities.***

## 1.2 Major Activities, Breakthroughs, Highlights, and Issues

Overarching achievements, some of which have been reported in previous Annual Reports, include:

- The Australian Research Data Commons (ARDC) has been established, and substantial progress has been made in populating it. The ARDC is a combination of the set of shareable Australian research collections, the descriptions of those collections, including the information required to support their reuse, the relationships between the various elements involved (the data, the researchers who produced it, the instruments that collected it and the institutions where they work), and the infrastructure needed to enable, populate and support the commons. In summary, all components of the ARDC exist, but not all components have been established and used at all relevant institutions, nor by all relevant researchers.
- ANDS has driven a change in the research data management uptake in Australia. ANDS is engaged with all major research institutions, and importantly they are engaged with and learning from each other's approaches.
- Research data infrastructure and research data management have been established at a significant number of research institutions. ANDS estimates that there are approximately 300 people working on data management within research institutions, which is probably a ten-fold increase compared to January 2009. Research institutions are seeing substantial value in this infrastructure (see for example quotes from the University of Adelaide and James Cook University). ANDS' investments at institutions have triggered substantial

co-investment and post-investment, with over \$2M of institutional investment made to date, and over \$3M of post-project investment. This indicates the extent to which institutions are embedding a research data infrastructure into standard operations.

- Data is overwhelmingly on the agenda in research and research infrastructure, and ANDS has helped position Australia internationally. The Research Infrastructure Roadmap produced in 2011 saw data as crucial infrastructure for research, and this has been emphasised in the draft National Research Investment Plan, which refers to the crucial role information and data play in enabling Australian research to tackle the key research challenges of the country. Very importantly, research institutional leaders are similarly seeing great value in the way that research data can give their researchers an advantage in research data partnerships and tackling larger research questions. This perception is mirrored internationally. There was a very strong emphasis on research data in the 2012 International Conference on Research Infrastructure in Copenhagen. Both Europe and the US have made significant investments in research infrastructure. ANDS has had an important role in ensuring that Australia has a leading role in international research infrastructure initiatives, particularly through the emerging Research Data Alliance, where Australia is partnering with the US and the EU. ANDS, together with some of the data-intensive capabilities, has made a significant contribution in ensuring that

Australian researchers and research institutions are engaged and leading in these global trends.

- The Australian Research Data Commons has matured and grown substantially. There are twice as many collections and three times as many contributing institutions as compared to last year, and the ARDC now covers every Field of Research. Importantly, now that the Data Citation service is available, researchers are availing themselves of the opportunity to publish their research data, using minted Digital Object Identifiers to connect their collection description to the relevant data repository, and to cite their data.

The 2013-14 business year saw a gradual transition from completion of the NCRIS and EIF projects to undertaking CRIS and NCRIS 2013 activities. The activities under NCRIS and EIF were principally to:

- Maintain and strengthen existing ANDS national data services.
- Continue institutional research data management engagement.
- Continue activity to strengthen the research data policy environment.
- Continue to increase the number of collections discoverable in Research Data Australia.
- Strengthen the National Collections available to researchers.
- Conclude the program of institutional research data infrastructure establishment – including automated metadata capture, metadata stores, and tools to support data management.
- Conclude the demonstration of how data can be used effectively using the Australian Research Data Commons.

- Develop the Research Data Alliance – in partnership with the EU and the US – which is committed to “data sharing without barriers”, particularly by holding the Plenary in Dublin, Ireland, and by attracting a increasing level of international commitment.

In this period ANDS has transitioned from project support to institutional partnerships, where there was significant focus on other parts of the eResearch landscape through the increasing activities of RDSI and NeCTAR in particular.

Two activities worthy of particular note during this year are:

- A comprehensive response to the Australian Research Council (ARC) changes in funding guidelines with regard to research data – ANDS provided individual advice and help, generated guides, held workshops and hosted meetings that enabled the ARC to respond directly to institutional queries.
- A deepening of the ability for researchers to publish their research data, with persistent identifiers for the data, and rich connections to publications, research grants, and researchers.

Thus, as previously noted, by the end of the NCRIS and EIF projects ANDS has worked across the whole sector in partnership with major research organisations and NCRIS facilities. Significant progress has been made in enabling improved data management, connectivity, discoverability, and usability by:

- Establishing the Australian Research Data Commons, a network of shared data resources.
- Populating the Australian Research Data Commons with 100,000 research data collections.
- Dramatically improving institutional research data management capacity.
- Helping to establish institutional research data infrastructure.
- Co-leading the establishment of the Research Data Alliance, improving international data exchange.

This has meant that Australian researchers, research institutions and the nation are at the forefront of the opportunities inherent in global research data intensive activity.

## 2. Activities Undertaken

### 2.1 Research Infrastructure

ANDS has continued to make progress towards its goals of providing greater support to enable researchers to work in the new world of data-intensive research – notably through the substantial increase of effort in describing data collections and making the descriptions automatically visible through Research Data Australia (RDA), the increased use of specific services (Identify My Data and Register My Data), roadshows and “boot camps” to improve the capability of institutions to manage and share their research data, engagement with specific institutions to better support their data management, and the provision of advice and documentation in various data management areas via the ANDS website. More detailed reports on progress in this area are contained in **section 10.2**. The effort placed on infrastructure development can be seen in Figure 2 where the dark green pipes and green boxes show the infrastructure being created in the ARDC project.

Complementary infrastructure established as part of the ARDC project is described in a separate progress report. Some of this infrastructure was to be established in the NCRIS ANDS project, but as a result of the ARDC project plan and a modified ANDS Business Plan, other work was undertaken under the ARDC project.

As a result of the ARDC project, the NCRIS ANDS project was consolidated to two programs of activity, and an additional program focused on national collections of research data:

- **Frameworks and Capabilities** – the frameworks that will enable research data producing institutions to capture, manage and share research data; and improving Australia’s capability to manage its research data.
- **Seeding the Commons** – improving local data capture and populating the data commons.

- **National Collections** – ensuring that the formation and curation of research data collections, together with associated services, is maintained.

The associated ARDC project had five programs of activity:

- **Data Capture** is constructing the pipes that connect data sources to the data stores and the metadata stores.
- **Public Sector Data** is connecting data held in public sector agencies to the commons either from their data and metadata stores to the ANDS portal.
- The **Metadata Stores** program is creating a set of metadata store solutions that can be deployed at research institutions.
- The **ARDC Core** program is creating the infrastructure that enables collections to be identified, harvested and discovered through the ANDS portal.
- The **ARDC Applications** program is designed to enable researchers to exploit the whole of the ARDC infrastructure to get new value from existing data.

A new program focused on **International Infrastructure** was introduced during the year to consolidate international activities and reflect an increasing focus on International Infrastructure. This program was funded by a combination of EIF funds and funding from separate contracts with the Department of Infrastructure.

The next section describes the specific research infrastructure created in the 2013-14 financial year funded under the NCRIS ANDS project.



Researcher Dr Fred Ford of CSIRO’s Australian National Wildlife Collection (ANWC). Image courtesy of CSIRO and David McClenaghan (CC BY 3.0)

## 2.2 Seeding the Commons

### 2.2.1 Overview of program

The aims of this program were to improve the fabric for data management in a way that will increase the amount of content in the data commons; and to improve the state of data capture and management across the research sector, with a focus on the tertiary education sector, CSIRO and the NCRIS Capabilities.

ANDS had a significant existing set of commitments for this program. Its major activities were:

- 40 research institutions developing local capability for managing research data and making collections visible in the ARDC through funded projects.
- regional local support through funded positions.
- National advice and direct support.

These activities were focused on growing the commons, supporting partners, and providing national advice. This advice complemented the written guides and training provided in the Frameworks and Capability program. Specific contributions to growing the data commons provided:

- advice on metadata standards and requirements to ensure that metadata is prepared in a manner consistent with ANDS' needs.
- advice on ANDS service requirements to ensure that metadata is available in the ARDC.
- advice and support on requirements to enable the creation and sustainability of research data infrastructure.
- funding of staff at partner projects and institutions to provide support for ANDS' goals.
- Identification of as much content as possible and making it discoverable.

### 2.2.2 Outline of projects

An analysis of research intensity for the major Australian research-producing institutions was undertaken in late 2009 based on the most recent publicly available data, and \$4.55M of Seeding the Commons funds were allocated in bands of \$250K, \$125K, or \$75K, as detailed below.

Institutions were sent an invitation to take part in an Expression of Interest process in late 2009.

Institution	\$,000
CSIRO	250
The University of Sydney	250
The University of Melbourne	250
The University of Queensland	250
University of New South Wales	250
Monash University	250
The Australian National University	250
The University of Adelaide	125
The University of Western Australia	125
University of Wollongong	125
Queensland University of Technology	125
Macquarie University	125
Griffith University	125
Curtin University of Technology	125
Royal Melbourne Institute of Technology	125
University of South Australia	125
University of Newcastle	125
University of Technology, Sydney	125
La Trobe University	125
Deakin University	125
University of Western Sydney	125
University of Tasmania	125
The Flinders University of South Australia	125
James Cook University	125
Charles Sturt University	75
Victoria University	75
University of New England	75
Swinburne University of Technology	75
Edith Cowan University	75
Murdoch University	75
University of Canberra	75
University of Southern Queensland	75
Central Queensland University	75
<b>Total</b>	<b>\$4,550</b>

Table 1: Institutions invited to participate in Eol

The outcomes and deliverables of the [ANDS-funded] Seeding the Commons and the Metadata Stores programs have assisted universities and research organisations with implementing infrastructure and developing local expertise to better manage their research data assets.

PICASSO, VICKI. & PHELAN, LIAM. (2014). 'THE EVOLUTION OF OPEN ACCESS TO RESEARCH AND DATA IN AUSTRALIAN HIGHER EDUCATION'. RUSC. UNIVERSITIES AND KNOWLEDGE SOCIETY JOURNAL, 11(3). PP. 122-133. THE BIOSCOPE, MAY 2012, AS QUOTED IN SHARE ISSUE 19, JULY 2014

Charles Sturt University did not respond to the EoI. The funds allocated to them will be used on other areas of the program, notably ANSTO. At the end of 2011 it was decided to further extend this program, and to offer \$75,000 to each of the following universities:

- Australian Catholic University.
- Charles Darwin University.
- Southern Cross University.
- University of Ballarat (Federation University).
- Bond University.
- University of the Sunshine Coast.
- Charles Sturt University.

All accepted the offer except for Charles Sturt University.

At June 30, 2012, ANDS had either entered into contracts (or had agreed on project descriptions) for Seeding the Commons projects at all of listed institutions. A breakdown of the progress made in relation to this is provided in **section 2.2.3**. Detailed descriptions of the contracted and agreed projects can be found in **Section 10.2**.

Activities to work more closely with partners will provide:

- review and assessment of partner projects to ensure they are completed on time and as specified.
- advice to these projects and other ANDS programs.
- targeted engagements around a coherent approach to research data management, with a focus on the specific needs and desires of the partners; where possible these will be integrated with the Metadata Stores projects.

- advice and assistance on data management and related policy and procedures.
- advice and assistance in the use and deployment of ANDS produced or funded services, applications and material.
- identification of and partnering with exemplar institutions to maximise data management; and
- analysis, reuse and redeployment assistance of the outputs of the projects funded by ANDS or drawn from the ANDS product catalogue.

ANDS will create a community of data managers through:

- continuing to support the state based or related groups of data managers established to date, with effective communication channels between them.
- training provided to Community members as required (in conjunction with Frameworks and Capabilities Program).
- capturing information about successes for dissemination within the community and beyond.
- developing relationships with equivalent activities overseas to share approaches to data management systems that can inform ANDS.

### 2.2.3 Activity/Deliverables for 2013-14

Over the past year, the last remaining 9 projects funded under the Seeding the Commons program were completed. All of these delivered collection records to the ARDC, as well as promoting the growth of data management policy within the institution. In addition to these projects, ANDS-funded work has produced advice and guidance material on data management policy and practice, which has been made available to the larger data manager community, through ANDS communication tools.

Unlike the Data Capture program, the Seeding the Commons projects tend to be similar in nature. Common elements of these projects include:

- Identification of existing data collections within an institution, with a view to enabling the exposure and description of them where possible.
- Examining existing processes, policies and workflows in data management, and creating new policies and procedures with the intention of using what is learnt to improve data management capabilities within the institution.
- Creating descriptions of data collections to increase understanding of how these descriptions can best be created.
- Building tools to enable and simplify metadata capture related to data collections.

All *Seeding the Commons*, along with *Data Capture*, *Applications* and *Metadata Stores* funded projects are now complete and involved running over 170 discrete projects with most of the 39 research institutions

based in all states and territories around Australia. These projects have delivered a number of successful outcomes; most significantly a cultural change within the institutions on how data is managed. These projects have also resulted in a number of ongoing or fixed term positions being established at the institutions within the

Library and/or eResearch Office dedicated to supporting their researchers in managing data assets.

The following table demonstrates the projects that have been completed during this reporting period. Some of the projects were reported as completed, flagged\*, in the last reporting period (2012-2013)

but awaiting the submission of final reports. As the final report was submitted, the project has again been listed in this table; these projects are flagged in the table below. Descriptions of the projects completed are listed in **section 10.2**.

Institution	Project Title	Project Agreed	Contracted	Underway	Completed
Charles Darwin University	CDUeDATA	X	X	X	X
Macquarie University	Macquarie University Seeding the Commons	X	X	X	X
Murdoch University	Integrating precision agriculture	X	X	X	X
University of New England	UNE's N.C.W. Beadle Herbarium Database	X	X	X	X
University of Queensland*	Seeding the Commons Funding	X	X	X	X
University of Southern Queensland*	Sustainable policy and procedure for capturing research data	X	X	X	X
University of Sydney*	Seeding the Commons at the University of Sydney	X	X	X	X
University of New England	N.C.W. Herbarium Beadle Database	X	X	X	X
University of Queensland	Seeding the Commons at University of Queensland	X	X	X	X
University of Western Australia	Building a Research Data Registry for the University of Western Australia	X	X	X	X
Victoria University*	Research data framework	X	X	X	X

**Table 2:** Current status of Seeding the Commons projects

By the end of the reporting period all of the 42 Seeding the Commons projects have been completed. These projects were undertaken with 39 research institutions and three national organisations (CSIRO, ANSTO, and the Australia National Corpus via Griffith University). The project undertaken at Charles Darwin University included implementing a metadata store (RedBox instance in the NeCTAR cloud) at the institutions, which although outside the scope of the Seeding the Commons program, was deemed to be of value so

was undertaken using in-kind support at the institution. The delivery of records into Research Data Australia continued to grow during the reporting period, increasing the amount of data discoverable and accessible through the data commons.

Although the amount of funding was modest compared to some of the later funding programs, the *Seeding the Commons* projects initiated a process of ANDS engaging with institutions that continues today. This has enabled ANDs to continue to engage with

institutions on new programs funded under NCRIS 2013. *Seeding the Commons* has delivered a significant change of awareness about the importance of research data issues and also captured and published a number of significant data collections to kick-start the Australian Research Data Commons (ARDC).

Members of the *Institutional Engagement* team, previously known as the *Seeding the Commons* have been working to provide engagement support for the *National Collections*, encompassing the previous

## 35 Faculty Liaison Librarians were trained in research data management, significantly improving the capability of the Library to support researchers' data management needs across all faculties.

UNIVERSITY OF SYDNEY'S SEEDING THE COMMONS PROJECT

*Public Sector Data program*, in their roles as Outreach Officers for the various universities.

ANDS is working to use the infrastructure described above to further develop a coherent approach to institutional research data using the CRIS and NCRIS2 funding. The aim of this approach is for ANDS to build on existing relationships with research institutions to help them to deliver nationally significant collections while simultaneously further achieving their research data management ambitions and needs.

The program has been able to do this in part due to the provision of outreach staffing. ANDS has extended contractual agreements to employ staff at eResearch South Australia (eRSA), QCIF (Qld), Intersect (NSW), and Edith Cowan University. These staff have enabled ANDS to have an active presence in cities outside of where ANDS offices are based. ANDS has also actively partnered with these organisations in the running and administration of projects and events which cross over both parties' agendas.

ANDS has been working to grow the community of research data managers and data librarians in Australia through the provision of well-used communication tools (message board, a partners email list and website; including a registry of ANDS funded projects). Additionally training programs have been held to promote the understanding of ANDS' expectations around data management and data citation. This work has been done in conjunction with the Capabilities program. ANDS has also been holding regular online 'catch ups' and informal events where ANDS staff are available via a virtual meeting, to discuss particular topics, and anyone with questions can join. Initially these were set up for the new Seeding the Commons projects, but due

to popular demand similar sessions were added for the Metadata Stores projects. As projects have been completed, the focus of these sessions changed from the funded projects, to broader discussions for new people moving into the area, many of whom had not previously been working on ANDS funded projects or research data management activities. Informal state based sessions have also been held to help grow the community and promote information sharing.

The main national community event was the *eResearch Australasia* conference, at which ANDS staff were able to meet with institutional and other NCRIS capability partners. As has been reported in previous years, presentations were again given either by ANDS staff or in collaboration with our partners. These included participation in general and showcase sessions, panel discussions, Birds of a Feather (BoF) session and workshops. A large number of presentations at the conference related to ANDS funded projects given by our partners.

As briefly mentioned above, other activities have focused on continuing to build a community of data managers needed to support ANDS and institutional goals into the future. However, due to a reduction in staff numbers and the focus shifting to ensuring the completion of funded projects during this reporting period, the number of ANDS Community Events was not as active as in previous years. Events that did take place, to continue to building community efforts workshops in collaboration with our project partners at eResearch Australasia 2013, regular; monthly basis, 'Informals' for ANDS partners based in Victoria and Tasmania, roundtable events in conjunction with Intersect in NSW, and a Community event in Adelaide, in collaboration with the *Capabilities* team.

The team continued to work on background material in preparation for the institutional engagement phase of ANDS' activity and projects funded under NCRIS 2013. This includes internal procedures for managing the new MODC program funded under NCRIS 2013, the development of and identification of the status of each institution against the service package to inform ANDS prior to any engagement, and a re-evaluation of what skills and expertise can be offered to partners, following up on the assessment of internal skills and capability identified in the previous reporting period (2012-13) in order to deliver against the services identified in the package.

### 2.2.4 Program Highlights, Issues and Breakthroughs

The main highlights of the program were successful completion of the embedding of nationally coherent research data management infrastructure at most research institutions that is relevant to their needs. Perhaps more significantly the embedding of much greater research data infrastructure – technology, policies, data management capability, meant that there was a sophisticated response to the change in ARC funding rules.

Funds provided via the Seeding the Commons projects were an opportunity to bring together people from various areas from within the institution. It was a valuable exercise and the beginning of connecting various elements at the institution to think about data as an asset that needed to be collected and managed. The project encouraged 'learning by doing'. This resulted in Charles Darwin University extending their project to include a Metadata Store, the University of Melbourne, who completed their project in a previous reporting period,

have continued to extend on what started with the Seeding the Commons project. By increasing their capabilities at the institution regarding research data management, including employing a number of staff, including a Research Data Curator, developing training programs, and the implementation of the immersive Informatics pilot program aimed at building capacity and capability at the institution by providing professional staff with Research Data Management skills. The University of Southern Queensland, similar to other institutions, kept the governance in place that was set up during the Seeding the Commons project, which has enabled them to continue discussions at the institution on managing data and starting discussions on developing a policy.

In general the Seeding the Commons project included the creation of a number of research data management policies, the training of dedicated staff, a number of new ongoing positions have been created upon completed of the ANDS funded projects, various websites and other material on issues around research data management at institutions, and the creation and the input of a wide range of collection descriptions in to Research Data Australia.

To advertise and collate the existence of these outcomes the projects are listed on the ANDS Projects Registry at <http://projects.ands.org.au/getAllProjects.php?start=sc>.

At the completion of each project, a communication strategy was implemented to ensure that partners were informed when projects were completed via an email posting to the ANDS partners list and the announcement added on the front page (ANDS News) of the ANDS website. These announcements were appreciated by partners.

Another highlight is the continued growth in awareness of the importance of institutional support for data management as a result of these projects. A number of partners have formed high level committees or working groups with responsibility for oversight and development of work in this area. A number of eResearch or Research Services support units and/or teams have been set up to continue data management activities at institutions. Since the program began, a number of institutions: Queensland University of Technology, University of Adelaide, University of Western Australia, University of New England, Curtin, Edith Cowan, Flinders, Victoria, Griffith, Monash, and Melbourne universities have funded ongoing roles to continue their data management activities beyond the ANDS funded projects.

A challenge for the program has been recruiting adequate staff to vacated and newly created positions, following the reduction in staff during the previous reporting period.

#### PROGRAM LEARNINGS

The projects completed during this reporting cycle took longer to complete than originally expected. This was reported on last time, and the reasons for these delays remained the same; resourcing issues at the partners' institutions. These issues took a number of forms – appointing inappropriate or inexperienced staff, giving them insufficient support (reporting or metadata creation), expecting staff to absorb this new task within their existing role, and lack of understanding of the full implications of the terms outlined in the contract with ANDS. Internal restructuring at institutions, lack of research engagement, and staff turnover has also impacted on a number of the projects.

**Learning:** As reported last time, ANDS has a limited ability to influence an institutional partner with whom it contracts, or employs on the project, to deliver the required work within the allocated timeframes. The most effective response has been to pay close attention to the work being produced and to use our relationships with the institutions to detect potential issues. ANDS position has been to ensure good working relationships with our partners are maintained. As such, where possible ANDS staff have tried to work closely with partner staff to enable them to meet deadlines and deliverables, balancing the workload between what needs to be undertaken by our partners and how we can assist them; in some cases this has involved ANDS staff undertaking some of the work. The challenge is how best to balance being a true institutional partner with staying focussed on project compliance and delivery.

For new and future projects ANDS has become more involved in the planning process for each project, encouraging the creation of steering groups with high level buy-in, and discouraging projects that our experience have shown to be overly ambitious or under resourced. We have also been able to offer more direct support in the earlier stages of the projects, and re-use existing material from other projects to assist the new ones. This has been the same approach undertaken and reported on in previous years.

## 2.3 Frameworks and Capabilities

### 2.3.1 Overview of program

The Frameworks and Capabilities program addressed two of the systemic obstacles to the emergence of the ARDC: policy irregularity/absence and human capability constraints.

The common approach to addressing both of these generic issues was to partner with collaborators around specific solutions. The Frameworks and Capabilities program produced materials that addressed some of the fundamental shared issues in data intensive research. The key collaborators for the policy frameworks agenda were research leaders, funding agencies, and government agencies such as the Department of Innovation and the Office of the Australian Information Commissioner. Frameworks also increasingly contributed to government reviews, as for example, the Australian Law Reform Commission *Review of Copyright and the Digital Economy* and the Department of Innovation's Assessing the wider benefits arising from university-based research discussion paper.

Cohesive networks of research data are increasingly regarded as an important and enduring part of the collaborative research infrastructure. As a result ANDS Capabilities activities focussed in particular on building the capability of researchers and support staff to contribute to and better exploit national data infrastructure. ANDS worked with the sector to identify and document the fundamentals of working with research

data and the specifics of discipline-based data-intensive research. ANDS also worked with research communities and local e-Research support services to improve particular data-related competencies, as well as enhancing and adding national focus to institutionally based support, materials development, and training initiatives.

The policy Frameworks activities were focused primarily on the research community, the funders that fund their research, and the institutions in which they work, which include universities, museums, libraries, galleries and government agencies. ANDS worked towards harmonising and streamlining the overall policy framework within which a data commons can operate. An aspiration is to develop a shared vision of the opportunities, benefits and responsibilities of a data commons. It is important to acknowledge that the Frameworks activities worked through facilitating the goal of an effective research data commons rather than prescribing the specific policies required to achieve that end.

ANDS engaged with a number of initiatives including departmental activities around APS200 projects, RDIC, and the OAIC (the latter being primarily concerned with access to government data). ANDS' primary collaborators for 2013-2014 include:

- The former Department of Industry, Innovation, Science, Research and Tertiary Education (briefs and submissions; a highlight here is the Open Access Draft Policy Note and Open Access Principles,

which are being incorporated into a departmental public discussion paper on (open access to) publicly-funded research data.

- Research funding agencies such as the Australian Research Council (ARC), National Health and Medical Research Council (NHMRC).
- Office of the Australian Information Commissioner (briefs, submissions and working groups); a highlight here is Framework's input to three OIAC papers on principles and practice of open public sector data<sup>1</sup>.
- Institutional data holders (CSIRO, NCRIS Capabilities, National Library of Australia, Departments of Primary Industry, GeoScience Australia, Australian Bureau of Statistics, DAFF etc.).
- Cross-governmental groups such as Australian Government Information Management Office (AGIMO) and the Office of Spatial Policy (OSP).
- Government enquiries, consultations and initiatives such as "Opening up access to data produced from publicly-funded research in Australia" (ARCom/Innovation) and "Draft National Principles of Intellectual Property Management for Publicly Funded Research Conducted in the Public Sector" (Innovation).
- Peak bodies such as CAUL, CAUDIT, ARMS and UA.

1. <http://www.oaic.gov.au> (go to bottom of page Information Policy: reports 2-4; Open public sector information includes the ANDS submission)

## 2.3.2 Outline of projects

The Capabilities area of this program is a well-established program of ongoing activity, with a number of projects:

### COMMUNITY AND CAPABILITY BUILDING:

- Nurturing and extending the ANDS' Partners Community, through Google Groups, Community Bulletin Board, and ANDS Community Days.
- Conducting Public Events, principally through new webinar technology, providing targeted training, facilitating workshops, and extended intensives.

### GUIDES AND SUPPORT MATERIALS:

- ANDS Content Providers Guide (major project).
- Guides and Topic pages.

Ongoing projects have continued and new initiatives have been established in the three activity areas of Policy Frameworks:

### DATA COMMONS POLICY:

- National policy, which traditionally focused on reports and submissions to ARC, NHMRC and the Department of Innovation, has expanded to Universities Australia, CAUL, CAUDIT, OAIC. As reported in this update, Frameworks has been very active in drafting and co-drafting discussion papers and data commons policy documents with Innovation, as well as with several of its functionaries (RDIC, ARCom).
- Encouraging data licensing has progressed well during this reporting period, with AusGOAL increasingly recognised as the default licensing system for government, university and innovation sectors. This has been supported by ANDS webinars and new web pages on both the ANDS and AusGOAL websites.
- Ethics and data re-use policy has been very active and has produced a (web-based) guide, articles and submissions.



## 2.3.3 Activity/Deliverables for 2013-14

### COMMUNITY BUILDING

Capabilities have continued to further develop proven engagement strategies but also explored new ways of reaching out to, and improving professional networking opportunities for, the Australian research data management community.

New this year:

- Over the past year the Capabilities team has revisited basic concepts to inform and train the new cohort of institutional staff at the beginning of their data management journey e.g. Webinar *DOIs and Data Citation - Back to Basics* attracted 48 registrations.
- The addition of 3 new staff members (2.2fte) in March 2014 with particular connections and expertise has allowed for increased event activity with:
  - CAUL (Council of Australian University Librarians) - *Becoming a data librarian - everything you wanted to know ...* and *Data Citation: Griffith University's journey*.
  - Research Offices - ARC 1 *Perspectives on data management ...* and ARC 2 *Supporting the ARC Funding Rules Changes*.
  - NHMRC Fellow specialising in the publication and reuse of sensitive data - there will be a major roll out of an information campaign in Q4 2014.
- 1. Increased use of online media to promulgate information about new reports, events, updates: andsUP (fortnightly eNewsletter), Twitter, ANDS YouTube channel.

### Webinars and workshops

ANDS Webinars and workshops are interwoven to enable people from all over Australia to network, build community, engage, contribute and learn: both online and face-to-face.

Webinars allow participants to see and hear presentations, as well as interact with the presenters and other participants. The discussions are often lively and several webinars have run overtime when the questions and comments kept coming.

Where possible, webinars are recorded and made available through the ANDS website and the *andsdata* YouTube channel <https://www.youtube.com/user/andsdata>.

Webinar and workshop presenters are drawn from the community to ensure national and international perspectives and practices inform the development of the Australian community:

- National perspectives examples:
  - ANDS partners e.g. Almost every webinar and all ANDS workshops include ANDS partners as facilitators, discussion leaders, presenters, mentors. Whilst the titles and formats of a workshop series may appear the same, local presenters are always foremost on the programs and programs are contextualised to the needs of the specific community:

ANDS has created some excellent resources for Australian researchers showing exactly how to share data.

AMANDA NIXON, MANAGER  
ERESEARCH@FLINDERS  
FLINDERS UNIVERSITY

Data Citation workshop: ANU 9 April	Data Citation workshop: University of Melbourne 29 January
<ul style="list-style-type: none"> <li>• Overview: Gerry Ryder, ANDS</li> <li>• Data citation and data.gov.au: Pia Waugh, data.gov.au</li> <li>• Data citation at CSIRO: Cynthia Love</li> <li>• Data citation at GA: Sue Fyfe, Geoscience Australia</li> <li>• Is "Publish or Perish" now "Visible or Vanish"? Virginia Barbour, PLOS</li> <li>• The roads to data citation – many paths, one goal: Anne Lahey, ANU</li> </ul>	<ul style="list-style-type: none"> <li>• Overview: Gerry Ryder, ANDS</li> <li>• Do DOIs = Data Citation? Karen Visser, ANDS</li> <li>• Data citation at Deakin University: Christopher McAvaney</li> <li>• Discussion: Participant Round Robin: Richard Ferrers</li> <li>• Data citation at CSIRO: Cynthia Love</li> <li>• Data citation – where are the rewards? Gerry Ryder</li> <li>• Data citation at Deakin University: Allison Moloney</li> </ul>

- Topic experts e.g. Data Citation Catch-ups are a monthly online community engagement opportunities and is a mix of informal sessions and formal webinars. Presenters and facilitators included: Australian Antarctic Division, GeoScience Australia, CSIRO, Griffith University, ANDS, Public Library of Science.
- International perspectives:
  - Dr Ross Wilkinson (ANDS): *National and international trends and developments* webinar which focussed on the role of Research Data Alliance in developing global understandings of data management.
  - *Delivering Comprehensive Research Data Management Services at University College London* – presented by Dr Max Wilkinson.

### Matrix of activities: a Data Management Capability case study

This matrix approach combines a number of different activities, such as webinars, blogs, website upgrades, Guide releases, projects and/or workshops, in order to build community on a specific theme over an extended period of time.

Community interest in data management continues to gain momentum within the Australian and international research data communities as they seek to enable researchers and institutions to gain recognition for data reuse. The highly successful 2012-13 program has been extended and enhanced in 2013-14 to include:

- Expanded and simplified online resources to support both technical capability and education (launched June 2013). These are summarised in the Data Management Overview table at <http://ands.org.au/datamanagement/overview.html>.

- Technical enhancements:
  - The following activities support each Release of upgrades to Research Data Australia, the Registry and other ANDS data management software: Webinars (e.g. R12 A closer look), recordings, news releases, emails.
- Webinars:
  - ARC Funding Rules changes Feb 2014: Webinar #1 *Perspectives on data management*, followed by Webinar #2 *Supporting the ARC Funding Rules Changes*.
  - Delivering Comprehensive Research Data Management Services at University College London* – presented by Dr Max Wilkinson.
- Workshops and Informals:
  - ANDS/Intersect Research Data Management Roundtable: #5 Vocabularies – Technical.
  - Victoria (Tas.) eResearch / Data Management Informal get togethers held in various locations (e.g. Melbourne Brain Centre, CSIRO) each month.
- Training for Research Data Management (QUT, UQ)* – an online ANDS Community Data Management Clinic.
- Introductory *Data Management Roundtable* held in Perth in March and followed up with the more in-depth *Research Data: managing one of our most important assets* workshop in June.

Resources for these matrix activities can be found at the ANDS Presentations and Publication page at <http://ands.org.au/presentations/index.html>.

### ANDS YouTube channel: providing information in visual and aural formats.

ANDS Videos range from 3 minutes to over 60 minutes and may feature ANDS staff, ANDS partners reflecting on their projects, or subject experts supporting their communities.

The ANDS YouTube channel is being accessed predominately by Australian viewers, as is to be expected, but it is interesting to note that there is also significant interest from the UK, USA and Germany.

**Popular uploads**



**Data Management Training at Uni of Melbourne and Uni of ...**  
1,303 views 1 year ago



**Data Management Training at ANU and Swinburne Universi...**  
499 views 1 year ago



**Infrastructure for Research Data Management**  
476 views 2 years ago



**Data Citation Counts!**  
452 views 1 year ago



**ANDS and ORCID - Making the Connection**  
368 views 11 months ago

## COMMUNITY BUILDING – CONTINUING ACTIVITIES

ANDS continues to enrich the range of previously established and successful community building activities such as the ANDS Partners Google Group. In addition, 2 new community building activities have been added since June 2013: a Twitter account and a fortnightly eNewsletter called andsUP.

### andsUP

- Double the initial starting number of 512 subscribers which were migrated from the ANDS Google group lists in late May 2013
  - "Opens" indicate the messages are of interest globally
  - Number of "opens" continues to increase: from 24% in June 2013 to 35.8% in April 2014 (industry average is 18% opens)
  - Subscribe and back issues: <http://ands.org.au/subscribe.html>
- 



Ingrid Mason presenting at ANDS Workshop, Gold Coast, 2014



Attendees engaged at ANDS workshop, Gold Coast 2014

GUIDES AND TRAINING MATERIALS

Guides

The annual comprehensive review cycle of all existing guides was undertaken with a view to ensuring currency. Most

ANDS guides are written by ANDS staff in consultation with various partners and other subject experts.

	12 month views
<a href="/guides/persistent-identifiers-expert.html">/guides/persistent-identifiers-expert.html</a>	12,905
<a href="/guides/ethics-working-level.html">/guides/ethics-working-level.html</a>	10,500
<a href="/guides/what-is-research-data.html">/guides/what-is-research-data.html</a>	9,290
<a href="/guides/content-providers-guide.html">/guides/content-providers-guide.html</a>	8,175
<a href="/guides/dmf-it-infrastructure.html">/guides/dmf-it-infrastructure.html</a>	6,819
<a href="/guides/data-management-planning-awareness.html">/guides/data-management-planning-awareness.html</a>	7,045

Investment in ANDS Guides is well rewarded with several Guides continuing to attract attention over a very extended period of time.

*Research at Murdoch; Data Management at Newcastle; Data Management and Publishing (RMIT); Research Data Management Toolkit (UWA); Research Data Management (University of Ballarat).*

The Guides are also integral to research data management information at most Australian universities. For examples, ANDS Guides are referenced in *Supporting*

Internationally, ANDS Guides are referenced by several groups, including the Digital Curation Centre in the UK.

### 2.3.4 Program Highlights, Issues and Breakthroughs

Highlights for the "building capabilities" agenda include:

- A high degree of success in ensuring the skills base required to undertake ANDS projects resulting in an extremely high percentage of ANDS projects being completed, and on time.
- During the first half of 2014, ANDS has seen a dramatic upswing with the number of people registering for webinars and almost all workshops held between January and June 2014 were fully subscribed with wait lists.

Webinars held February-May 2014	Registrations
ARC 1 Perspectives on data management	216
ARC 2 Supporting the ARC Funding Rules Changes	78
DOIs and Data Citation - Back to Basics	48
Becoming a data librarian - everything you wanted to know ...	204
Data Citation: Griffith University's journey	78

### 2.3.5 Program Learning

The capability of researchers, support staff, and research organizations to take advantage of national research infrastructure remains a significant obstacle, and strong demand for any training, instructions, awareness raising support that ANDS can provide continues unabated. Two areas have been of particular focus in 2013-14:

- the needs of those using ANDS national services.
- the needs of our partners institutions as they build their part of the distributed national infrastructure in the context of the completion of their ANDS funded projects.

However, Capabilities is also focussed on capability-building for data-centric infrastructure at both research organisations and other NCRIS facilities.

## 2.4 National Collections

### 2.4.1 Overview of program

The value of establishing national collections of research data is considerable and addresses many nationally significant data needs. One is to have collections of data that enable big problems to be addressed, often by investigating "survey" questions. Another is to have a reference collection that enables researchers to compare their results with a standard. A further need is to have a locus of all research data of a particular form to be explored. By bringing together 'like' data and highlighting collections of national significance, ANDS supports research through the improvement of its discoverability and subsequent potential for reuse. This consolidation also has the potential for showcasing valuable holdings globally and increasing the potential attractiveness of Australian research environment.

Through the National Collections program, ANDS increased the focus on establishing a rich set of such collections, leveraging off its many institutional partnerships and proceeding in a manner complementary to RDSI and RDSI nodes activity in storing national collections. The National Collections program worked with other ANDS programs to effect these changes by supporting the research sector and the research data providers in the identification of nationally significant collections of research data and **ensuring the availability of these collections;** that they are well **managed, connected, discoverable** and able to be **used** as widely as possible for as many purposes as possible.

Two ANDS programs, Public Sector Data and National Collections, were combined to form the National Collections program. To date Public Sector Data has engaged with over nineteen agencies to raise the profile of data management and expose their collections. It has resulted in over thirty-four thousand collections and data collections being published through Research Data Australia (RDA). National Collections has promulgated awareness of the value of a collections approach to data and gained some traction in acceptance of distributed national collections. The merging of these programs is indicative of the importance public data collections play as an input to research. There is increased value in a collections approach when combined with the delivery of storage solutions and services in partnership with other organisations. The scope of this program includes collections of government, institutional, discipline and national facility data.

### 2.4.2 Outline of projects

In determining these national collections, this program worked with all parts of the research and public sector, through a variety of approaches that included:

- Engaging with selected data custodian on establishing rich national collections.
- Working around the initial themes of Urban Water, Climate Change Adaptation, Great Barrier Reef, TERN and ARC/NH&MRC grants, the collections team will engage with research and public sector institutions to identify data collections for publication within the context of the topic and the tools associated with them.
- In consultation with research and data providers, developed an inventory of collections that are both institutionally and nationally significant.
- identifying contributions to nationally significant distributed collections themes Liaise with RDSI nodes and data providers to establish appropriate storage and access.
- Engaging with service providers to provide access to tools associated with data collections.
- Richly describing national collections and publish in RDA and other relevant portals.
- Enhancing RDA to provide on the fly presentation of tools and services for national collections and click through access to those services.
- Working with NeCTAR and RDSI to ensure that their online services also utilise data discovery in RDA.

### 2.4.3 Activity/Deliverables for 2013-14

The following table describes the status of the projects in this area.

Agency or Institution or Project	Description and status
Urban Water	A flagship distributed collection in Research Data Australia. See: <a href="http://researchdata.ands.org.au/australian-urban-water-collection/267201">http://researchdata.ands.org.au/australian-urban-water-collection/267201</a> Engagement was facilitated through the Urban Water Research & Development Coalition and the Partnership Working Group. Work has continued on developing the Australian Urban Water Collection in RDA with a number of meetings held with potential contributors to the collection. A revised approach to our engagement strategy was developed based on leveraging existing relationships with ANDS Partners (data providers).
TERN	Work has continued with TERN to structure their metadata records so as to represent data publishing from each TERN Facility and bring this together to provide an overall TERN view in Research Data Australia. Assistance has been provided with describing entities and relationships, establishing metadata feeds, metadata management and metadata crosswalks.
Geoscience Australia	Work continued with Geoscience Australia (GA) to capture citation metadata for over 20K collection records. Building on this activity was assisting GA to implement a Digital Object Identifier minting service, currently in Production. GA also recently released a new metadata catalogue requiring ANDS to regression test the integration between the two technologies.
National Computational Infrastructure	There was significant activity with the National Computational Infrastructure including providing assistance with developing a data management policy, a metadata schema and the implementation of a Digital Object Identifier minting service.
Bureau of Meteorology	Meetings were initiated with the Bureau of Meteorology to identify additional collections for the Australian Urban Water Collection in RDA. While a number of potential collections were identified, there would be substantial additional work for the Bureau to provide these as a metadata feed to RDA. Advice was provided to the Bureau on best practice for metadata management going forward.
Population Health Research Network	Informal agreement was reached with PHRN for harvesting of metadata to RDA.
Commonwealth Department of the Environment	Input was provided to the Department's Data and Information Management Policy for the National Environmental Research Hub. The Department will use the policy to ensure researchers comply with data management guidelines. ANDS will continue to work with the Department to provide metadata creation, data management and data repository advice to its funded activities.
Public Records Office of Victoria and State Archive NSW	Public Records Office of Victoria and State Archive NSW were supported in their production and implementation of an ANDS-funded solution, resulting in over 30K Collection records. This unique collection spans from the mid 1830s to today, and includes a significant amount of digitised data. Discussions were held with QLD State Archives, who would like to implement the re-usable solution, enabling the QLD archival data to be discoverable via Research Data Australia.
NSW Office of Environment and Heritage	ANDS was approached by NSW Office of Environment and Heritage to assist with the development of a data management policy and procedures for their agency. ANDS facilitated meetings between them and other clients to share data management journeys and associated solutions.
Australian Institute of Marine Science	ANDS and AIMS worked together on several fronts to re-establish the metadata feed from AIMS to RDA. ANDS also worked on the AIMS style sheet to capture citation metadata. When AIMS implements the style sheet, citation metadata will be ingested and displayed on RDA.
e-Atlas	The National Collections program worked closely with e-Atlas to improve the quality of their metadata within the ANDS Registry; most notably, this has included the addition of citation metadata and improved representation of relationships between research objects. This work has been re-used to make similar improvements to the metadata from AIMS and IMOS.
IMOS	Through the Australian Ocean Data Network (AODN), IMOS have added an additional 2000 metadata records to RDA.
Data.gov.au portals	ANDS undertook analysis and technical activity to enable harvesting of metadata from CKAN. The CKAN technology is used by data.gov.au to host metadata and data. CKAN is also used by other government portals. Work started with data.gov.au When the data.gov.au ingest has been proven, work with the other portals will commence. Informal agreement was reached with NSW Dept. of Finance and Services (oversees <a href="http://data.nsw.gov.au/data/">http://data.nsw.gov.au/data/</a> ) for harvesting of metadata to RDA.
Thematic linkages to collections in RDA	Following the release of technical and software infrastructure to support the development of theme pages, 8 new pages representing thematic views of content, were released to RDA in June. The themes align with the <i>NCRIS 2011 Strategic Roadmap for Research Infrastructure</i> and include: Integrated Biological Diversity, Astronomy, Terrestrial Systems, Urban Settlements, Population Health Research Platforms, Solid Earth, Marine Environment, and Characterisation.

**Table 3:** Current status of the projects in National Collections

## 2.4.4 Program Highlights, Issues and Breakthroughs

This year saw ANDS entering new territory by supporting native metadata and transport protocols, thus making it much easier for partners to publish metadata to the ANDS portal.

Through technical assistance from ANDS, all 20K+ records from Geoscience Australia now include citation metadata and improved rights metadata. This lays the groundwork for Digital Object Identifier (DOI) minting and providing metadata to Thomson Reuters Data Citation Index. This work, undertaken with Geoscience Australia to significantly improve the transform from ANZLIC metadata profile to the metadata schema required by RDA, was reused to improve records from other clients who use the ANZLIC metadata profile or the (extended) Marine Community Profile. Such clients include: CSIRO; Australian Antarctic Division; Terrestrial Ecosystem Research Network (TERN); e-Atlas; the Australian Institute of Marine Science; NSW Office of Environment and Heritage; and Integrated Marine Observing System (IMOS). In doing so, ANDS has enabled the provision of citation metadata, enabling individual researchers to receive citation for their data, and the institution to use metrics to report to their funding bodies and stakeholders.

ANDS has pursued an opportunity with the data.gov.au federal government portal and has established a bi-directional relationship so that ANDS harvested metadata from them as well as having a presence within data.gov.au. ANDS undertook the work to interrogate and understand the CKAN portal technology and develop a crosswalk which could be reused with minimal customisation and a prototype for generating ANDS Registry records.

A priority project within the National Collections program was The Australian Urban Water Collection. During this period, the team initiated, or continued, engagements with partnering institutions including the Goyder Institute, the National Water Commission, the Australian Water Recycling Centre of Excellence, Water Research Australia, the CRC for Water Sensitive Cities and Smart Water Fund (Vic) to identify funded data collections that may form part of this collection. ANDS has also been involved in reviewing a Business Case for an Urban Water Research Knowledge Management portal led by the Partnership Working Group of the National Urban Water Research and Development Forum. A significant challenge for this program has been progressing the population of the Australian Urban Water Collection. There are a number of inhibitors including immature data management policies and the lack of data publication mandates by funders in this sector. In addition, many national institutions are constrained by the current fiscal environment.

The National Collections program continued to be involved in supporting feature development in RDA, including enabling increased user participation through the deployment of an annotation/tagging facility and supporting the cross-institutional linking of thematic collections of significance. At the conclusion of this reporting period, eight new pages representing thematic views of content, were released to RDA. The themes align with the *NCRIS 2011 Strategic Roadmap for Research Infrastructure* and include: Integrated Biological Diversity, Astronomy, Terrestrial Systems, Urban Settlements, Population Health Research Platforms, Solid Earth, Marine Environment, and Characterisation. This is a significant breakthrough in the presentation of content, and the highlighting of national collections, in RDA.

## 2.4.5 Program Learnings

There were a number of learnings that emerged for the program in this reporting period. Possibly the most significant one was readiness. The development of national collections is very dependent on the desire for both the creator to publish and the demand from the research area to drive the publication.

The Urban Water distributed collection also provided a number of learnings. One that was very evident was the lack of penetration of data management policy at the funder level, thereby removing a valuable imperative to the description and management of data by funded researchers and institutions. Addressing this deficit will be critical to the ongoing success of any collection activities that take the funding body as the starting point. The experience gained from the Goyder Institute project demonstrated that attaching publication as a condition of funding can be very effective.

Across the distribution collections the development of a community of practice of those data management staff from the various institutions has been very successful. It has enabled knowledge sharing, standards development and general moral support.

## 2.5 Project Office

### 2.5.1 Overview of program

This program is designed to ensure the effective and efficient delivery of the ANDS project with proper reporting of outcomes and compliance with all governance requirements. It has also adapted to support the delivery of all ANDS outcomes with the goal of delivering operational excellence. The workplace functions that sit under the project office are finance, communications, contract management, governance and reporting, internal IT systems, office administration and general operational support.

### 2.5.2 Outline of projects

Taking 'Excellence in operational management' as its goal, the program has been continually focused on streamlining processes and increasing the cohesion of the ANDS group for a more effective overall delivery of internal and external support. This has been achieved by reviewing existing processes and where relevant introducing new lightweight processes that allow for greater cross-organisational flexibility.

### 2.5.3 Activity/Deliverables for 2013-14

Over the last reporting period the activity of the Project Office team has been focused on supporting the ANDS team during what has been a challenging period. This period has seen an important transition in the life of the ANDS project as the remaining 60 projects funded under NCRIS and EIF concluded. Particular deliverables include:

- Providing a high level of administration and project support to ensure that all relevant programs funded under NCRIS and EIF were concluded with funds exhausted by the acquittal deadline.
- Continuing to drive a cohesive communications message to the sector using *share*, our e-newsletter andsUP and the social media platforms that we engage with.
- Providing ongoing training and support to non-administrative staff to reduce the amount of central administration effort required.



Image courtesy of Arthur John Picton (CC By 2.0)

## 2.6 Promotion

ANDS has undertaken a large number of promotional activities during the period July 2013 to the end of June 2014. These include the following.

### 2.6.1 Presentations/ attendance at Conferences & Workshops

ANDS staff have presented at and/or attended a range of international and local conferences to promote the service and to establish relations with other parties. These include:

- ODIN 1st year conference and code sprint, Geneva.
- eResearch Australasia, Brisbane.
- Research Profiling Conference, University of Melbourne.
- RD-A plenary 3, Dublin, Ireland.
- ODIN meeting, British Library, London.
- JISC/IDCC workshop, San Francisco.
- Open Repository Conference, Helsinki.
- Research Data Immersive Informatics course: University of Melbourne.
- Visualisation launch CAVE2, Monash University.
- Doing Data Better, Miniconference, University of Melbourne.
- RMIT Big Data Symposium, Good Data, Bad Data.
- Research Community Day, University of Melbourne.
- Copyright Council Training, Melbourne.
- Open Data Forum, NSW Government, Sydney.
- Apps4NSW, NSW Government, Sydney.
- RDA Plenary 2, Washington.
- Digital Curator Workshop, Sydney.
- International Digital Curation Conference, San Francisco.
- CSIRO eResearch Conference, Melbourne.
- Locate 14 Spatial Conference, Canberra.
- Committee on Publication Ethics.
- IEEE eScience'13, Beijing.
- CAUL Research Advisory Committee meetings, Brisbane.
- Thomson Reuters CONVERIS Roadshow, University of Queensland.
- PLOS and Open Access seminar with Dr Virginia Barbour, University of Queensland.
- DataCite Summer Meeting, Washington DC.
- Belmont Forum Steering Committee, Windsor.
- Third EU-AU Research Infrastructure Workshop, Canberra.
- Big and Complex Data Workshop, Melbourne.
- TERN and eResearch Infrastructure Workshop, Canberra.
- PARADISEC Conference, Melbourne.
- SKA Big Data Workshop.
- International Conference on Research Infrastructure, Athens.
- eResearch 2020 Workshop, Wellington, NZ.
- Belmont Forum Steering Committee, Vienna.
- TERN ACEAS Grand Synthesis Workshop, University of Queensland.
- TERN Flood Visualisation project stakeholders Workshop, University of Queensland.
- NeCTAR eResearch Projects Workshop, Melbourne.
- Australian Phenomics Workshop Strategic Planning Workshop, Adelaide.
- ARMS conference, Adelaide.
- Polar Data Forum, Tokyo.
- Open Access Conference, Brisbane.
- Research Profiles Conference 2014, Melbourne.
- 2nd Research Support Community day, Melbourne.
- AeRO National Forum, Perth.
- AuScope Symposium, Canberra.
- Urban Water R&D forum, Canberra.
- TERN IIDDG meeting, Canberra.
- Bioplatforms Australian and eResearch Coordination meeting, Canberra.
- QUT Open access and Research Conference, Brisbane.
- RD-A Council meeting, London.
- Meeting of the Data Infrastructures working group, London.
- Open Data Research Portal workshop, Gold Coast.
- Australian Government Science ICT Network Conference, Canberra.
- AusAID Symposium, Brisbane.
- RD-A Council meeting, Paris.
- AU Embassy visit, Brussels.
- 11th Australasian Data Mining Conference, Canberra.
- Data Management and Sharing in Government, Canberra.
- CSIRO eResearch Conference, Melbourne.

## 2.6.2 Forums

ANDS has hosted or presented a wide range of forums over the reporting period to build our communities, share knowledge and expertise, and provide support to our various audiences. It has changed focus from physical forums to more emphasis on virtual forums.

### VIRTUAL EVENTS

ANDS hosted a variety of free virtual events in the period to help our partners and communities learn, discuss and exchange ideas, and meet colleagues without even leaving their desk, and they have proven to be incredibly popular. Virtual events – including webinars, virtual meetings and 'how-to' sessions – remove any constraints associated with location, of either the presenter or participants. This has resulted in greater diversity of presenters and topics, which is a great benefit to our local and international audiences and also allows our communities to hear from and inform international and national perspectives on a wide range of topics and issues.

Increasingly we are noticing that the real value of the virtual events is how they are facilitating community building, and enabling the community to learn from each other.

Some sessions are recorded and available here: [ands.org.au/presentations/audio-video.html](http://ands.org.au/presentations/audio-video.html)

More detail around all events that have been run by ANDS between July 2013 and June 2014 can be found in section 2.3.3

## 2.6.3 Consultation meetings

ANDS staff have also continued to consult extensively with potential and current partners and stakeholders to discuss the services that ANDS offers and how they might be of interest to them.

## 2.6.4 Newsletter

The ANDS quarterly newsletter, *share*, continues to create awareness of ANDS, its activities and successes amongst the research community and stakeholders by providing updates on ANDS-funded projects, highlighting achievements and promoting ANDS events and objectives. The themed approach to each issue that was initiated in 2011 has been continued in all subsequent issues, resulting in more focused pieces that have been widely appreciated. This has resulted in an increase in the circulation of the newsletter, for both the digital and print versions. The three issues in the period have focused on Australia's role in global research data, celebrating the completion of ANDS funded projects and the value of research data.

## 2.6.5 Other activities

### TWITTER

In November 2011 the ANDS official Twitter account ([@andsdata](https://twitter.com/andsdata)) was launched at eResearch Australasia. Initially launched on a trial basis, the ANDS twitter account has primarily been used to communicate with our stakeholders about ANDS events, publications and news, as well as relevant stories/information from the sector. Twitter has subsequently become a useful communication channel that complements our two main communication channels: the ANDS website and share.

Interaction with [@andsdata](https://twitter.com/andsdata) from our stakeholders has grown steadily throughout the period. It is worth noting that we have a number of international stakeholders regularly interacting with [@andsdata](https://twitter.com/andsdata) including staff members from JISC (UK), Digital Curation Centre (UK), Australian Research Management Society and Public Library of Science. Tweets that have generated the highest number of click-throughs (meaning engagement with the content) are all promotional tweets about ANDS events, guides and share.

The number of [@andsdata](https://twitter.com/andsdata) followers has increased steadily over the last four years:

- **79** followers on 16 November 2011 (two weeks after the launch of [@andsdata](https://twitter.com/andsdata)).
- About **450** followers in April 2013.
- **864** followers on 12 May 2014.
- **978** followers on 14 August 2014.

[@andsdata](https://twitter.com/andsdata), on average, currently gains 1 new follower per day, from Australia or overseas.

[@andsdata](https://twitter.com/andsdata) followers are currently from the following countries/regions:

52%	Australia	2%	New Zealand
25%	Europe	1%	Africa
16%	US	1%	Middle East
2%	Asia	1%	South America

Those who have mentioned [@andsdata](https://twitter.com/andsdata) positively in their tweets include The Dataverse Project, Harvard University, and the British Library Science Team.



YOUTUBE

The last twelve months has seen an increased focus on utilising our Youtube channel. It has been a great way to share webinars as well as training videos and presentations that the capabilities team have presented. Below are some statistics that demonstrate the success that we have seen with Youtube over the last twelve months.

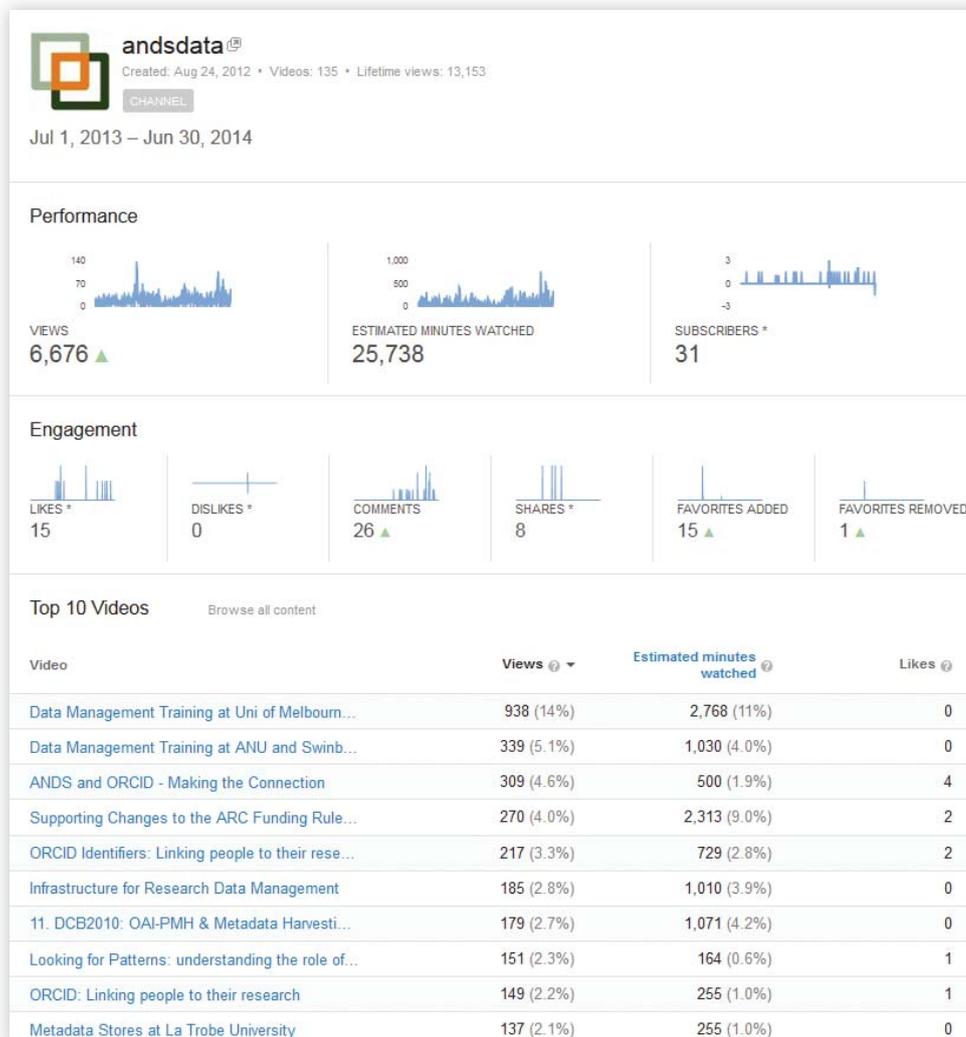


Figure 3: Youtube report outlining overview of performance and engagement

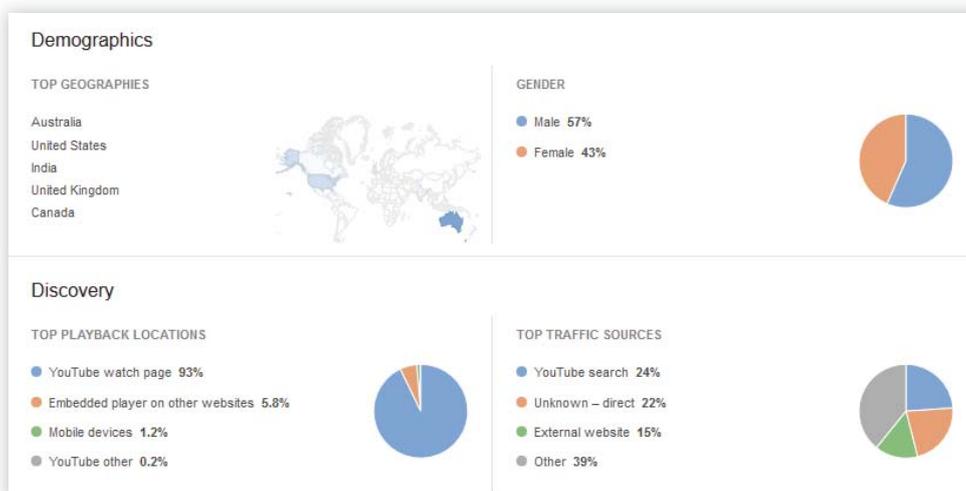


Figure 4: Youtube report outlining overview of demographics

## 2.7 Risk Management

ANDS maintains a Risk Register. The risk assessment methodology, adapted from the Australian Risk Management Standard AS/NZS 31000:2009, involves identifying and analysing each risk in terms of how likely it is to happen (Likelihood) and the possible impacts (Consequence). The risk score for each risk is calculated by combining Consequence score with the Likelihood score. This will give a risk score of between 2 and 10, which can then be mapped onto a Risk Scoring Matrix to give a risk rating of HIGH (8-10), SIGNIFICANT (7), MEDIUM (6) or LOW (2-5). Where there is more than one risk measurement area for scoring consequence, the highest combination of scores is taken as the final risk score.

The risk management register was reviewed in February 2014. ANDS assessed the residual risk level of all 11 existing risks in the register, taking into account the effect of the risk mitigation strategies that have been put in place.

At completion of this review it was concluded that there was only three changes from the previous review in June 2013.

Risk 1 had reduced from a High Risk to a Low risk due to recent changes around ARC funding. Risk 3 had reduced from a high risk to a medium risk as the number of outstanding has dramatically reduced with only 16 remaining. Also under the NCRIS 2013 plan we will have a much small amount of external contracts. A new Risk was added, Risk 12 - That ANDS is most effectively reporting under all the various programs.

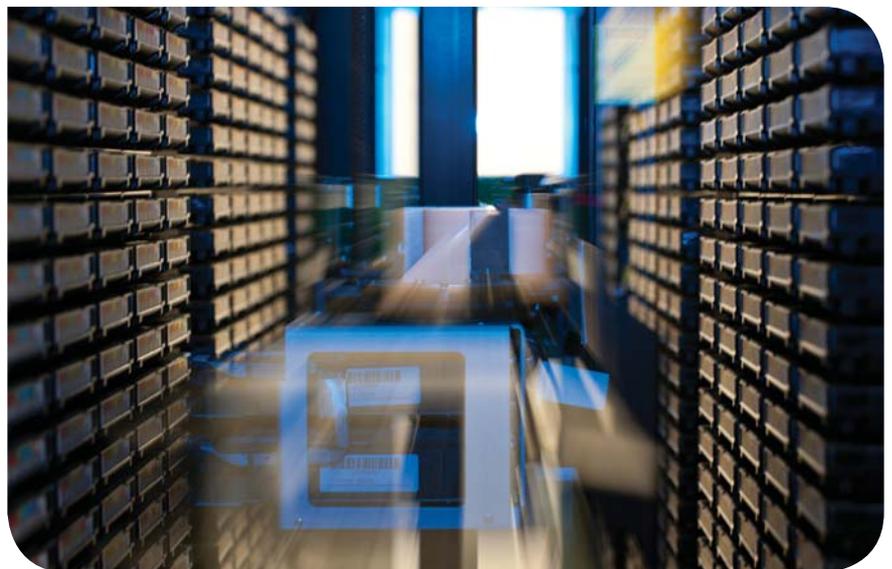


Image: LaRDS tape storage library courtesy of Monash Medicine Multimedia

# 3. Progress against milestones

## 3.1 Seeding the Commons

These are the milestones for the Institutional Engagement Program in the 2013-14 Annual Business Plan. Progress is reported against the NCRIS Seeding the Commons Program, the EIF Data Capture Program, the EIF Metadata Stores Program, and the ARDC Application Program.

Milestone Date	Milestone	Progress
13Q3	Five institutional agreements on work programs for their research data assets Successful conclusion of all Data Capture, Seeding the Commons, Metadata Stores and Applications projects (NOTE: most will have concluded by end of 13Q2)	Due to a number of factors at the partnering institutions which was outside the control of ANDS including; slow recruitment of lead roles, staff shortages, under-estimation of required resources and technology for researcher engagement, delays in starting projects etc. Data Capture, Seeding the Commons, Metadata Stores and Applications projects did not conclude in 13Q3 or 13Q2. These continued to be monitored and concluded in later quarters – refer to details in this table and corresponding table in EIF report 5 Seeding the Commons projects completed Community events held in conjunction with Capabilities, along with many local events being held
13Q4	Further five institutional agreements on work programs for their research data assets Identification of some institutional assets	3 Seeding the Commons projects completed 22 institutional engagements to identify institutional assets - Major Open Data Collections (MODC) – to form agreements funded under NCRIS 2013 Community events held in conjunction with Capabilities, along with many local events being held
14Q1	Further five institutional agreements on work programs for their research data assets Identification of some institutional assets Concluded agreed work on research data assets with initial five institutions	No further Seeding the Commons projects completed 2 institutional engagements to identify institutional assets - Major Open Data Collections (MODC) – to form agreements funded under NCRIS 2013 Community events held in conjunction with Capabilities, along with many local events being held
14Q2	Further five institutional agreements on work programs for their research data assets Identification of some institutional assets Concluded agreed work on research data assets with further five institutions	1 Seeding the Commons project completed. This concludes the Seeding the Commons project; as all 42 projects now completed 16 institutional engagements to identify institutional assets - Major Open Data Collections (MODC) and Open Data Collections (ODC)– to be funded under NCRIS 2013 Eight (out of 23 agreements) MODC contracts executed; three with the institution pending signatures; and two under review by institutional lawyers - ANDS in project discussion with remaining ten institutions Concluded agreed work with Eleven institutions - Open Data Collections contracts executed – projects to be undertaken under NCRIS 2013 Community events held in conjunction with Capabilities, along with many local events being held

## 3.2 Frameworks and Capabilities

Milestone Date	Milestone	Progress
13Q3	Training delivered for DOIs, ORCID, Metadata quality Materials developed: national collections, institutional engagements First ANDS post- projects national events held	Completed: one-to-one support sessions held as requested and identified by partners and ANDS Client Liaison staff  Not held: Changed format to short recordings and/or webinars in order to increase the reach of the project outcomes and limit cost.  See: <a href="http://www.youtube.com/user/andsdata">http://www.youtube.com/user/andsdata</a>
13Q4	Citation proof of concept projects: mid project reports received Community support webinars held: Geospatial, Data Citation, Identifiers, Vocabularies	Completed: all projects completed mid-project reports Ongoing: al full list of webinars are on the ANDS Events page: <a href="http://ands.org.au/events/index.html">http://ands.org.au/events/index.html</a>
14Q1	Community support events held: National Collections	Events were held in Melbourne, Perth, Adelaide, Sydney
14Q2	Citation proof of concepts: final reports received Research data management materials Phase 3 completed	Completed: all projects completed Ongoing: materials have been developed and are available on the ANDS website

---

Leaving it on the hard drive for your grandchild to take to pieces after you're dead is not a very good use for it. Doing something with it – putting it in a university repository for instance – is like building your own virtual mausoleum.

---

PROF RICHARD MALTBY, FLINDERS UNIVERSITY

## 3.3 National Collections

Milestone Date	Milestone	Achievement
13Q3	Commenced engagement with three institutions on establishment of collections of national significance with RDSI nodes or institutional infrastructure  Three topic pages and two themed distributed collections published	Engagement with institutions: <ul style="list-style-type: none"> <li>• TERN</li> <li>• Australian Water Recycling Centre of Excellence</li> <li>• Smart Water Fund</li> </ul>
13Q4	Commenced engagement with three institutions on establishment of collections of national significance with RDSI nodes or institutional infrastructure  Services associated with themed distributed collections available via RDA  Further two themes commenced for distributed collections  Further three topic pages published	<ul style="list-style-type: none"> <li>• National Water Commission</li> <li>• Water Research Australia</li> <li>• CRC for Water Sensitive Cities</li> <li>• Bureau of Meteorology</li> <li>• Water Services Association of Australia</li> <li>• NSW Office of Environment &amp; Heritage</li> <li>• National Computational Infrastructure (NCI)</li> <li>• Data.gov portals</li> <li>• Geoscience Australia</li> <li>• Commonwealth Department of the Environment</li> <li>• e-Atlas</li> <li>• Australian Institute of Marine Science</li> <li>• Population Health Research Network</li> <li>• Public Records Office of Victoria</li> </ul>
14Q1	Commenced engagement with three institutions on establishment of collections of national significance with RDSI nodes or institutional infrastructure  Collections published with initial three institutions  Services associated with themed distributed collections available via RDA  Further three topic pages published	<ul style="list-style-type: none"> <li>• Geoscience Australia Contributor Page showcasing two national collections: Australian Reflectance Grid (ARG25) and ASTER Geoscience Map of Australia</li> <li>• Australian Terrestrial Ecosystem Research Collection (TERN)</li> <li>• Australian Urban Water Collection</li> </ul> <p>Topic (theme) pages published (8) with associated services available via RDA:</p> <ul style="list-style-type: none"> <li>• Marine Environment</li> <li>• Terrestrial Systems</li> <li>• Solid Earth</li> <li>• Astronomy</li> <li>• Integrated Biological Discovery</li> <li>• Characterisation</li> <li>• Urban Settlements</li> <li>• Population Health Research Platforms</li> </ul>
14Q2	Commenced engagement with three institutions on establishment of collections of national significance with RDSI nodes or institutional infrastructure  Continued engagement with next three institutions  Collections published with next three institutions  Further three topic pages published	

## 4. Deviations from the Project Plan

The ANDS Business Plan for 2013-14 had to take into account NCRIS, EIF, CRIS, DWF, and NCRIS 2013 funding, and this report describes activity against the NCRIS and EIF components of this funding. It has thus had to separate activity in an integrated set of activities based on the funding source, and being mindful of the differing conditions of grant.

There have been no substantial changes or deviations from the business plan that was submitted, but the focus on institutional engagement during the year moved from completion of funded institutional research data infrastructure projects, to

broader institutional engagement, and the establishment of NCRIS 2013 funded major open data collections. These collections diluted to some extent focus on more general collection development to focus on open data collections. This became particularly important given the growing strength of open data expectations in Australia and internationally.

ANDS also crafted a considerable response to the ARC funding rule changes that were not anticipated in the business plan but were consistent with the business plan.

---

[Previously] There was no system for centrally documenting and recording the existence of research datasets at the University [of Western Australia], or for publicising their existence nationally and internationally. Through the implementation of the VIVO metadata hub, the University now has a platform for recording and publicising datasets, and for contributing descriptions of datasets to Research Data Australia. A total of 83 dataset descriptions were made available through Research Data Australia as a result of this project.

---

BUILDING A RESEARCH DATA REGISTRY FOR THE UNIVERSITY OF WESTERN AUSTRALIA

# 5. Financial and Human Resources

The following table indicates ANDS expenditure by program for July 2013 to June 2014. Income is shown in the audit statement in section 9.

## ANDS Expenditure for the Financial Year ended 30<sup>th</sup> June 2014

Item	2013-2014
<b>SALARIES</b>	
Frameworks & Capabilities	16,247
Seeding the Commons	80,064
Monash University Seeding the Commons Project	3
Seeding the Commons – Outreach	6,514
National Collections	76,531
Project Office	263,258
<b>Total Salaries</b>	<b>442,617</b>
<b>OPERATING EXPENSES</b>	
Frameworks & Capabilities Internal Program Expenses	15,330
Frameworks & Capabilities Contracts	62,501
Frameworks & Capabilities ANU Program Funding Payment	666,980
Seeding the Commons External Infrastructure Contracts	727,500
Seeding the Commons External Outreach Contracts	312,180
Seeding the Commons Internal Program Expenses	8,541
National Collections Internal Program Expenses	20,546
Project Office Internal Expenses	301,690
<b>Total Operating Expenses</b>	<b>2,115,268</b>
<b>NCRIS Total Expenditure</b>	<b>2,557,885</b>

Table 4: ANDS Expenditure for 2013-14

Staffing and recruitment has been extensive since the establishment of the project. The organisation chart at the end of June 2014 is shown in Figure 5.

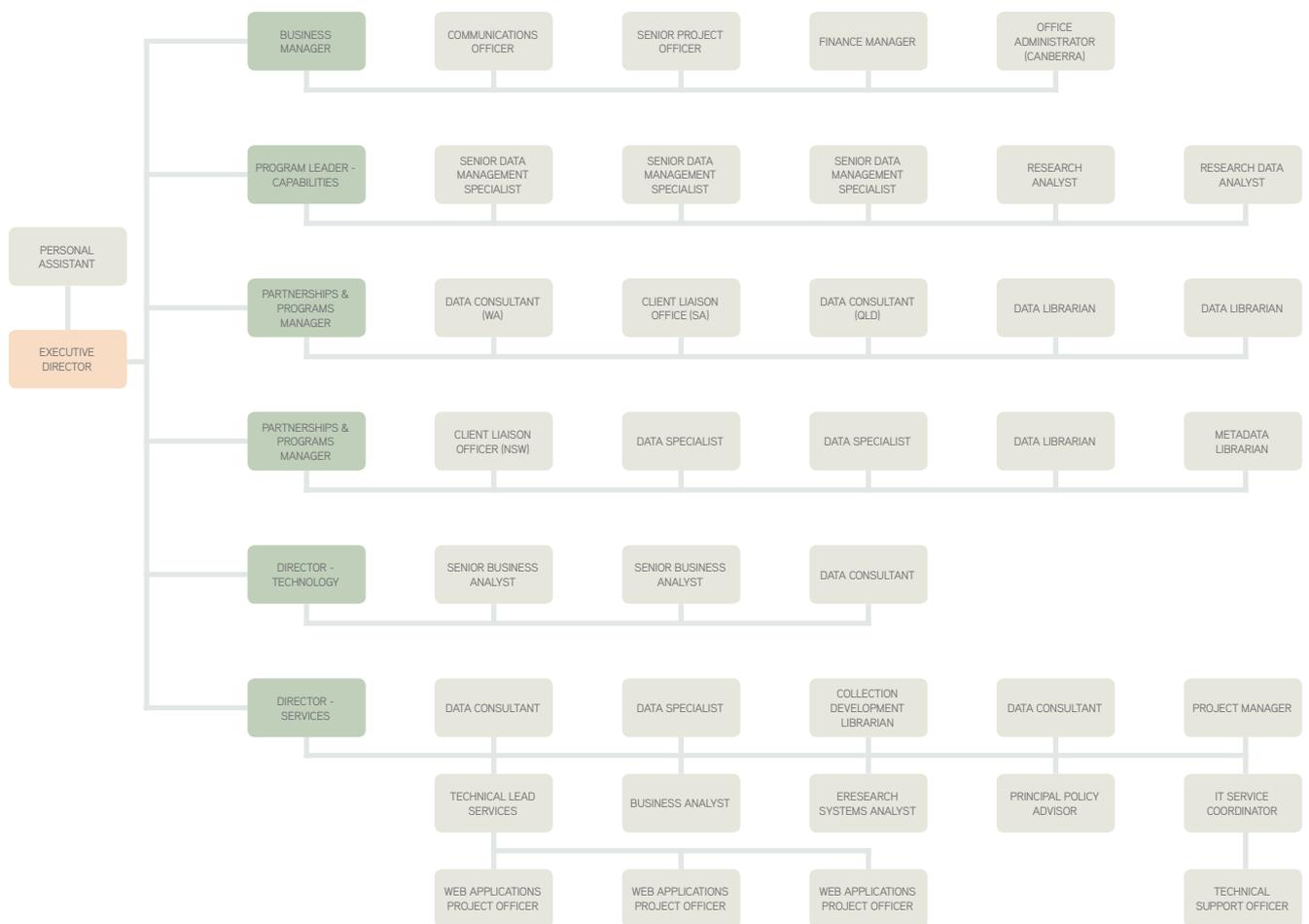


Figure 5: ANDS Organisational Chart

# 6. Co-Investment

## 6.1 Access and Pricing

All ANDS services are available free of charge to all Australian researchers at publicly funded research institutions and discovery is available to all. As such no cash has been received for the provision of ANDS services, which is in line with expectations.

## 6.2 Project Co-Investment

As a result of the original intended timeframe for the project of 2 years, it was agreed that it would be inappropriate to require co-investment in ANDS projects. The 2009-10 and 2010-11 Business Plans identified that ANDS would place a wholly-funded ANDS staff member within an institution to achieve the aims of that institution and of ANDS. This was changed to a more flexible arrangement where ANDS contracted with partners to allow for a mix of staffing needs.

ANDS investments at institutions in Seeding the Commons and Data Capture activities have triggered substantial co-investment and post-investment with a number of ongoing or fixed term positions being created, and over \$3.8M of institutional investment made to date, and over \$3M of post-project investment identified to date. This indicates the extent to which institutions are embedding research data infrastructure into standard operations. It is pleasing to note that this additional effort is being continued beyond the life of the projects, showing the importance being placed on research data, and the extent to which institutions are establishing coherent research data infrastructure as part of "business as usual" operations.

As a part of the agreement to fund metadata stores at institutions ANDS required that the institution indicate what co-investment it would make to demonstrate a whole-of-institution commitment to their metadata infrastructure. This amount varied across the institutions, but in total ANDS had commitments of around \$3.8 million in co-investment, beyond the \$5 million that ANDS has provided.

To provide a perspective over the life of the project to date:

- Total ANDS Funds \$75,431,120
- Total Partner co-investment \$13,993,316
- Total Post project co-investment is estimated to be 18.55% of ANDS investment

Looking just at the financial year 2013-14:

- Total ANDS Funds \$18,972,670
- Total Partner co-investment \$3,367,866
- Total Post project co-investment is estimated to be 17.75% of ANDS investment

# 7. Performance Indicators

## 7.1 KPI Report

The following are the KPIs agreed in the 2010-11 Business Plan, with actual results and commentary included (some of these results have been achieved through EIF ARDC project activity, but cannot be usefully separated):

1. The number and coverage of data repositories providing metadata feeds to the national registry compared to the number of data repositories ANDS intends to build at least 80 automatic plus 100 manual metadata feeds. This will cover at least 35 out of the approximately 50 research data-holding institutions that we know about.

**Result:** 71 institutions fed collection descriptions to RDA along with 46 individual collections. From these institutions, 157 data source feeds were setup (101 automatic and 56 manual feeds).

2. The number and coverage of institutions and number of research groups with which ANDS has engaged: ANDS will continue to engage with all Australian universities, PFRO's, and 4 major Government data providers this year, and through them at least 50 research groups.

**Result:** ANDS engaged with the following:

- 39 universities.
- Publicly Funded Research Organisations: ANSTO, CSIRO and AIMS.
- Government data providers: ANDS has engaged with over 30 government agencies apart from the PFROs. These include GeoScience Australia (GA), Australian Institute of Health & Welfare (AIHW), Australian Antarctic Division (AAD), Murray Darling Basin Authority, Bureau of

Meteorology, Queensland Dept of Employment Economic Development & Innovation, Public Records office of Victoria (PROV) and Australian Bureau of Statistics (ABS) directly; and Royal Australian Navy through the engagement with AODN; National Archives of Australia, State Records NSW and State Archives of Queensland through the engagement with PROV; and 18 museums through the Museum Metadata Exchange project including Powerhouse Museum, Australian Museum and state museums. Through the AustLII project we have exposed public data from Attorney General's Department and various courts around Australia including the High Court. Through the project with AuScope, engagement has been with Bureau of Meteorology and various state Departments of Primary Industry and Sustainability & Environment.

- National facilities: Australian Animal Health Laboratory (AAHL), Australia Telescope National Facility (ATNF), Australian Synchrotron and research vessels: Southern Surveyor and Aurora Australis.
- National capabilities: Astronomy Australia, Integrated Marine Observing System (IMOS), Terrestrial Ecosystem Research Network (TERN), AuScope, Atlas of Living Australia (ALA), Population Health Research Network (PHRN), Australian Urban Research Infrastructure Network (AURIN), Bioplatforms Australia, Australian Phenomics Network (APN), Australian Plant Phenomics Facility (APPF), National Imaging Facility (NIF), Australian Microscopy & Microanalysis Research Facility (AMMRF), EMBL Australia,

Australian National Fabrication Facility, AusBiotech, Research Infrastructure Support Services, Australian Biosecurity Intelligence Network, Australian Social Sciences Data Archive, Australian Access Federation, National Research Network, Pawsey Supercomputing Centre, National eResearch Collaboration Tools and Resources (NeCTAR), Research Data Storage Infrastructure (RDSI).

3. The number of institutions with research data management policies and practices consistent with ANDS recommendations: 25.

**Result:** 12 – CSIRO, Monash University, University of Melbourne, Queensland University of Technology, Griffith University, University of Wollongong, University of Newcastle, Edith Cowan University, La Trobe University, Australian Catholic University, University of New England and University of Canberra. ANDS engaged with many more institutions on the development of research data management policies and practices this year, however comparatively few of those engagements have been finalized at this stage.

4. The number of times a search is initiated with an ANDS discovery service: There was no target for this year; this is the first year that we have taken these measurements.

**Result:** In November 2013, ANDS started using Google analytics for logging searches in Research Data Australia. From then on, there were a total of **23,769** searches in Research Data Australia, with an average of around **2,900** searches per month.

5. The number of times an ANDS data page is accessed: 300,000 in this year, up from the KPI of 100,000 for last year.

**Result:** 310, 697 page views and 200,760 unique page views (from Google Analytics tool). Unique page views increased 75% from last year, increased more than 4 times since June 2011.

6. The satisfaction of researchers and partners (see below) with ANDS services as measured by an annual survey - no number can be given here, but a report will be provided.

**Result:** A Survey was completed by 101 respondents: including partners and researchers. Over 84% of the Universities that ANDS works with responded to the Survey.

I get the most satisfaction out of knowing that our data is valuable to other people.

EUAN SANGSTER, CSIRO

Key Outcomes:	
1. Satisfaction with ANDS services overall	<ul style="list-style-type: none"> <li>94% were satisfied or very satisfied</li> <li>5% dissatisfied</li> </ul>
2. Awareness of the value of data	<ul style="list-style-type: none"> <li>81% thought data valuable or very valuable</li> <li>4% thought research data of little value</li> </ul>
3. Change in research data management (RDM) practices over last two years	<ul style="list-style-type: none"> <li>75% noticed a change (up from 63% last year)</li> <li>13% partner) noticed no change (down from 18% last year)</li> <li>13% were unsure (down from 20% last year)</li> </ul>
4. Satisfied with Research Data Australia (RDA)	<ul style="list-style-type: none"> <li>75% satisfied or very satisfied</li> </ul>
5. Research Data routinely deposited into data repositories	<ul style="list-style-type: none"> <li>20% usually or always</li> <li>58% seldom</li> </ul>

Compared to the last survey ANDS has kept the level of satisfaction stable at 82%, a slight rise (3%) from prior year.

**Summary of comments:**

ANDS impact though overall positive (per above), still attracts significant criticism, particularly from researchers. Over 200 comments were gathered in the Survey. These ranged from strongly positive to strongly negative and were overall mixed rather than predominantly positive or negative. The negative comments indicate there is much work still to be done. The researcher criticism is a result of ANDS engagement primarily at Institutional rather than researcher level.

While it would be useful to summarise the Survey comments into a shopping list of 'must fixes' for ANDS, what emerges

from the comments is the voice of the stakeholders and the dichotomy of opinion. While most stakeholders are overall satisfied there is much tension in the community reflecting that the sector is not yet mature. Key themes that emerged from the comments were the tension between tradition and the new emerging activity, the need for support, data as core to research, and a theme of making progress. Events, project support and the website received strongest support while Research Data Australia attracted the most (but still only a few) negative comments.

Examples of Positive Comments	Examples of Negative Comments
<p>ANDS has been an <u>invaluable catalyst</u> for better eResearch and research data management generally.</p> <p><b>Awareness:</b> More staff are aware of the function and the importance of what ANDS is striving to achieve.</p> <p><b>RDA:</b> The new RDA portal looks fantastic. Great update to the interface. Well done.</p>	<p>I think in general ANDS <u>lacks visibility</u>.</p> <p><b>Awareness:</b> We are aware of ANDS service but do not use them.</p> <p><b>RDA</b> is a very clunky database, improved by the latest release but still not fantastic...</p>

7. The number of data access and sharing agreements with stakeholders – principally research institutions, government data agencies, government research agencies: ANDS aims to strike at least 30 agreements to make data available.

**Result:** ANDS had agreements with **71 organisations** to publish data collection descriptions in either the public or draft systems, 3 new agreements added since June last year.

There are two measures that ANDS will not have full control over, but that are important and will measure our success in influencing others' behaviour:

8. The number of research data sets in the ARDC: more than 10,000 collections.

**Result:** 97,672 collections were made available in Research Data Australia as at 30 June 2014. 89% of the total collection records came from the top 5 top contributors: 22% from Geoscience Australia (GA), 19% from Queensland Facility for Advanced Bioinformatics (QFAB), 17% from State Records NSW, 16% from Public Records Office Victoria (PROV) and 15% from Australian Ocean Data Network (AODN). The remaining 11% of the total collections are from 66 contributors with a total of 85 corresponding groups appearing in Research Data Australia.

The number of ANDS research data sets increased **45%** from June last year.

9. The number of research data sets with persistent identifiers: 10,000.

**Result:** **12,082** persistent identifiers (9,677 PIDS handles, 2,405 DOI handles), 17% higher than June last year. These identifiers were minted by 17 Identify My Data (PIDS) clients and 13 out of 29 registered Cite My Data (DOI) clients.

There is a final measure that ANDS aspires to – it will be measured but is unlikely to be a useful short-term KPI.

10. The number of times a data set is reused and referenced – the ultimate long-term measure. At present ANDS is unable to report on this. However, the DOI-based data citation service will encourage publication of data collections with persistent identifiers in a citable form, and existing citation tracking services are expanding their coverage to include such data citations.

**Notes:**

An ANDS data page is a page generated from the ANDS collections registry that describes a data set, a collection, a research group, a research project, or an institution.

ANDS will focus on monitoring Institutions that are research data producing organisations, such as the Bureau of Meteorology, Landsat, the Australian Synchrotron, the Cultural Collections sector, and the research data using organisations, such as the Universities, the PFROs, and affiliates. Many organisations have both roles.

Researchers have many partners in carrying out research and ANDS needs to satisfy their needs as well – this includes funders, assessors, institutional representatives, such as DVC-Rs, eResearch Directors, Information providers such as libraries, IT providers such as University ITS Departments, partner service providers, such as ARCS and NCI, as well as umbrella organisations such as disciplinary bodies such as the Academies, international research bodies, etc.

The qualitative measures are intended to capture not only usage figures, but also attitudinal attributes – ANDS only succeeds with cultural change, so this will be measured as well. The first survey has set benchmarks, but also help inform future surveys.

---

The genomics revolution has come to fruition over the last five years, and we took that opportunity to prepare data that would facilitate collaboration across research communities.

---

ANDREW GILBERT, BIOPLATFORMS  
AUSTRALIA GENERAL MANAGER

## 7.2 Progress over the Life of the ANDS Project

Having been in progress for 51 months it is possible now to indicate life over the project, along with expectations based on current activity, and when appropriate an indication of what might be possible.

Based on activity taking this year, it is possible to determine an estimate for 2013-14, and when appropriate give a figure indicating total coverage. The measures that are described augment

the KPI information with additional measures that help understand the Australian Research Data Commons.

Measure	09-10	10-11	11-12	12-13	13-14	Maximum <sup>6</sup>
Research Institutions Participation in RDA <sup>1</sup>	NA <sup>5</sup>	21	35	40	41	43
Research Institutions with Data Management Policy and Practice	3	4	10	12	25	43
Institutional context capture tools	0	6	43	53	71	
Institutional Research Metadata Store	0	9	25	25	31	43
Research Data Provider Participation <sup>2</sup>	NA <sup>5</sup>	9	30 <sup>7</sup>	28	35	
Research Data Infrastructure Partners <sup>3</sup>		4	23	23	23	26
Research Data Collections	1,173	26,746	40,811	56,599	97,672	
Research Data Exploitation Tools		9	26	30	30	
Research Fields of Research Coverage <sup>4</sup>	5	21	22	22	22	22

**Table 5:** Progress over the life of the project

Note 1. Research Institutions denotes all 39 Universities and 4 Publicly Funded Research Organisations

Note 2. Research Data Provider participation is measured by those organisations that are not Universities and Publicly Funded Research Organisations that have an agreement with ANDS to publish research data collections descriptions – this may be indirect through Infrastructure provider partners

Note 3. Research Data Infrastructure partners refer to those NCRIS and EIF infrastructure providers that could exchange research data collections descriptions – this measures how effectively ANDS is partnering with other problem specific data investments

Note 4. Research Fields of Research based on ANZSRC FOR codes – all, including DIVISION 22 PHILOSOPHY AND RELIGIOUS STUDIES are now covered in Research Data Australia (RDA)

Note 5. Research Data Australia had not been launched as at 30th June 2010

Note 6. This denotes the total number of Research Institutions: all 39 Universities and 4 Publicly Funded Research Organisations

Note 7. This number included organisations that we are now considering as under the auspices of a Research Institution

## 7.3 Overall Progress

In previous Annual Reports we were able to report on some overarching achievements:

- A populated Australian Research Data Commons (ARDC) has been established.
- ANDS has driven a change in the research data management uptake in Australia. ANDS is engaged with all major research institutions, and importantly they are engaged with and learning from each other's approaches.
- Research data infrastructure and research data management practices have both been established at a significant number of research institutions.
- Data is overwhelmingly on the agenda in research and research infrastructure, and ANDS has helped position Australia internationally.
- The Australian Research Data Commons has matured and grown substantially.

There were a number of notable changes in the research data environment that are worth noting as they had an influence on ANDS activities:

- ARC change in Policy to require data planning as part of the grant submission process.

- The European Horizon 2020 funding program allocating E1.4b to open data.
- Publishers increased involvement with research data.
- The increased level of commercial data services such as Figshare and Zenodo.
- The advent of RDSI data storage availability - at a level and scale that was very different a year ago.
- Substantial acceptance of the need for data licencing such as through AusGOAL.

The major activities for ANDS as a whole in the 2013-14 calendar year were:

- Making more than 97,000, research data collection pages discoverable through Research Data Australia, Google, and other search engines harvested by ANDS at over 70 research data providing institutions.
- Adoption of the data citation services by 25 research institutions.
- Maturing research data management community strongly engaged with ANDS with substantial increase participation in, and demand for, ANDS events.
- Greater contribution to the national policy environment as exemplified by contributions to the Chief Scientist's

Office, National Environment Research Program, Research Data Infrastructure Committee.

- Demonstrating the value of reusing data with a suite of applications that has enabled high profile researchers to provide compelling demonstrations of new ways of conducting data intensive research.
- Completed a program of installation of research data infrastructure at institutions through the Seeding the Commons, Data Capture, Metadata Stores, and Data Applications programs.
- Building the RD-A and subsequent growth (Plenaries: 230, 365, 500, 700) 8 founding members to 300 to now over 1000 member organisations.
- Engagement with the research institutions in general and the research offices in particular around the rule changes by the ARC.

# 8. Audit Statement

The signed copy of the following document will be delivered separately.



**ands**  
AUSTRALIAN NATIONAL DATA SERVICE

**Australian National Data Service Project - NCRIS Funding**  
Statement of Income and Expenditure for the Financial Year Ended 30 June 2014

Item	2013-2014
Income	
Grant	-
Interest	29,599
<b>Total Income (a)</b>	<b>29,599</b>
Expenses	
Salaries	442,617
Non-Salary	2,115,243
<b>Total Expenses (b)</b>	<b>2,557,860</b>
<b>Surplus/(Deficit) for the reporting period (a - b)</b>	<b>(2,528,260)</b>
Brought forward surplus/(deficit) from 30 June 2013 (c)	2,528,260
<b>Balance Carried Forward to next Reporting Period (a - b) + (c)</b>	<b>0</b>

We, Ian Smith and Joel Chibert, hereby confirm the following:

- (i) The detailed statement of income and expenditure for the ANDS Establishment Project (shown above) represents a correct view of the financial performance for the period ended 30 June 2014
- (ii) The Funding was expended for the Project and was used in accordance with the agreement with the Department of Innovation, Industry, Science, Research and Tertiary Education.



**PROFESSOR IAN SMITH**  
Vice-Provost, Research & Research Infrastructure  
Office of the Provost and Senior Vice-President, Monash University



**MR JOEL CHIBERT**  
Director, Research & Revenue Accounting Services  
Office of the Chief Financial Officer and Senior Vice-President, Monash University

# 9. Summary of the Conduct of the Project

The ANDS project commenced on January 1<sup>st</sup> 2009, as a result of an extensive consultation that led to the "Towards the Australian Data Commons Report" that set the directions of the project, and an extensive process in the ANDS establishment process that determined governance, management and processes for the project, and determined the ANDS project plan. The project commenced with offices in Melbourne and Canberra, and staff from Monash, ANU, and CSIRO, working through four programs:

- Seeding the Commons.
- Developing Frameworks.
- Building Capabilities.
- Providing Utilities.

The first major engagement was a consultation around the country which led to an important decision: the long term custodians of research data are largely long lived institutions, so ANDS should focus its efforts on support of research institutions, and research data providing institutions, and support institutions in their direct engagement with researchers.

By June 30<sup>th</sup> 2009, partnerships were being established with research institutions, local and regional eResearch providers, research data providers, and NCRIS partners, particularly ARCS as we jointly ran the NeAT projects. By that date:

- An operational discovery service has been created.
- Initial research collections from IMOS, TARDIS and iVEC were discoverable.
- Initial material was developed to support research institutions developing their own research data capability.

- Guidelines on how to comply with the data management requirements in the Australian Code were prepared.
- NeAT projects were established and managed.

As a result of the Super Science program being announced in May 2009, the work of ANDS needed to rapidly change in size (\$24M to \$72M); in nature (from an organization that worked more directly on the changes to Australia's research data environment to one that funded partner organisations to do the bulk of the work); and in time (from two years to four).

The major activities for ANDS as a whole in 2009-10 were:

- Establishing relationships with all of our key partners in making the ARDC a reality.
- Establishing processes to fund programs effectively and efficiently.
- Initiating a \$10M set of activities in ARDC called Fast Start activities that enabled early progress on a 2 year program.
- Establishing infrastructure to identify, register and publish collections descriptions through Research Data Australia, the ANDS portal into the Australian Research Data Commons.
- Commencing a \$16.15M EOI process for both Data Capture and Seeding the Commons based on research output, with appropriate checks and balances to enable good outcomes and accountability.
- Continuing the NeAT program of discipline enhancing tools for improved collaboration and exploitation of research data.

- Establishing a set of landmark public sector data activities to make public data available to researchers.

Whilst there was inevitable disruption with the changes described and this has slowed some delivery, it has been very helpful in deepening engagement with our partners, and increasing the scope of data capture and sharing that had previously been envisaged in the Seeding the Commons program.

The breakthrough that occurred in 2009-10 was that research data management and ANDS has moved from a "good idea" but with little institutional engagement, to a situation where 33 Universities and many other organisations are actively engaged on data transformation projects at their institutions, and where institutional investment in managing their research data is occurring at the policy, procedural and practical infrastructure. ANDS has been a catalyst for much of this activity and research data collections are recognized now as key to many institutions' research ambitions.

By June 30<sup>th</sup>, 2011, the Australian Research Data Commons (ARDC) had been established. The ARDC is a combination of the set of shareable Australian research collections, the descriptions of those collections including the information required to support their re-use, the relationships between the various elements involved (the data, the researchers who produced it, the instruments that collected it and the institutions where they work), and the infrastructure needed to enable, populate and support the commons. This included:

- Enabling research data collections to be described and harvested by ANDS at 26 research institutions.

- Enabling research data collections to be described and harvested by ANDS at 27 research data providers including other research infrastructure providers, public sector providers, cultural institutions and research consortia.
- Making 26,746 research data collection from 21 providers discoverable.

By June 30<sup>th</sup> 2012, there had been a greatly increased research data management capacity at our research institutions. ANDS estimates that there were approximately 300 people working on data management within research institutions, which is probably a ten-fold increase compared to January 2009. Research institutions are seeing substantial value in this infrastructure. ANDS investments at institutions have triggered substantial co-investment and post-investment with over \$2M of institutional investment made to date, and over \$3M of post-project investment. This indicates the extent to which institutions were embedding a research data infrastructure into standard operations.

Significant international activity took place focused on interacting with the European Commission, national activity and EC funded projects. Through engagement initiated by the Department leading to a meeting on research infrastructure, ANDS participated in discussions leading to establishing Research Data Alliance, where Australia is partnering with the US and the EU.

The major activities for ANDS conducted by June 30<sup>th</sup>, 2013 were:

- Made more than 70,000 research data collections discoverable through Research Data Australia, Google, and other search engines harvested by ANDS at over 70 research data providing institutions.
- Substantially updated the national data services enabling better publication and discovery of research data.
- Over 30,000 archival collections from major state archives became discoverable through Research Data Australia.
- All major research institutions had installed substantial research data infrastructure, notably metadata stores, that enabled them to participate in the Australian Research Data Commons.
- Demonstrated the value of reusing data with a suite of applications that enabled high profile researchers to provide compelling demonstrations of new ways of conducting data intensive research.
- Established the Research Data Alliance, in partnership with the EU and the US which is committed to "data sharing without barriers" with an initial Plenary in Gothenburg, Sweden, and attracted a very substantial level of international commitment.

Taken together, two important changes occurred during the final year of the project – firstly there was clearly an increased desire to publish and get credit for research data. The great interest in

data citation is evidence of this, so the ability of Australian researchers to publish their data will be increasingly important. Secondly, Australia's response to the open data agenda internationally has been made easier by having a locus of thinking about research data and having conversations, developing policy responses, and involving all relevant agencies, though ANDS policy and practice capabilities.

By June 30<sup>th</sup>, 2014 – the final year of the project, all components of the project were completed successfully. A summary of the year's activity is described in the previous chapter. Perhaps most notably, the Australian Research Council made a major change to its funding rules by requiring statements about research data management to be made at the time of funding submissions, and the research sector was ready. ANDS immediately ran a comprehensive engagement supporting research institutions in their engagement with their researchers. ANDS had been engaging at many levels on research data policy, enabling a substantial change to the research data landscape in Australia.

In summary, Australia has a coherent approach to research data, has an international profile enabling Australia's researchers and research institutions to form partnerships internationally to tackle the most demanding data intensive research. Much needs to be done – while this advantage has been established, the data collections themselves had not been established to the extent necessary or desirable, and partnership advantages have not been exploited to the extent possible. Nevertheless, Australia has established an international competitive advantage in research data.

# 10. Expected Future Usage of the Infrastructure

Research data, and research data infrastructure will be vital for the best research to be conducted, so will continue to be used. Increasingly there is an expectation that publicly funded research will make its outputs, including research data publicly available at no cost.

The value to the nation of its research data assets being available publicly are very high, as is the value to research institutions. It is thus highly desirable that research data infrastructure continues to be available at no cost, to ensure that researchers do not see barriers to publicly sharing data generated through their projects. Equally important, the need for freely accessible means of publishing research data, enabling researchers to reap the citation benefits, builds the value of open data to researchers.

---

Staff from the Science Partnerships Section of the Department of the Environment have worked closely with ANDS throughout 2014 on the development of data and information management guidelines for the National Environmental Research Program. The department's primary objective for working with ANDS was to address the need for clarity and discoverability in our data and information management guidelines. We obtained assistance from ANDS to develop a clear understanding of project data resulting from funded research, regardless of where it was derived and/or stored. ANDS' expertise and existing partnerships with research institutions provided valuable insights and guidance that has helped us develop a document that meets the needs of the Department and is clear and understandable for funding recipients.

---

NAOMI DWYER, ASSISTANT DIRECTOR  
SCIENCE PARTNERSHIPS SECTION (NERP)  
DEPARTMENT OF THE ENVIRONMENT

# 11. Expected Future Trends

Internationally the momentum to conduct data intensive research is increasing, and international effort to establish improved access to data, through open data policy, and improved data infrastructure, is also increasing. The rapid growth of the Research Data Alliance, and the G8 Science Ministers' declarations are strong evidence of these trends. As well, within domains, the staggering growth of publicly available data through such initiatives as EMBL and PDB, or data being generated by CERN, demonstrate just how important it is to be a leader in data intensive research. The development of the SKA with its staggering levels of data is both an Australian opportunity and a challenge. Many areas of research that used to have relatively little digital information are being transformed by new data capture approaches. Archaeology is using three-dimensional laser scanning of megalithic rock art, LiDAR data, and underwater video to capture data, leading to substantial infrastructure challenges.

Australia will see the same challenges, and will need to assemble the data resources to tackle its own biggest challenges, seeking the best international partners, with infrastructure that enables partnership over data that cannot move due to its size.

Research data must become the means that enables research, industry, education, government and public to engage – over policy, over evidence, to support training, to enable innovation, and to solve the national challenges. Research data must thus be available with rich context beyond the purposes for which it was collected.

Annually, Australia may spend up to \$1B on new research data – it must get the best value from that data – initially with the best researchers partnering with a data toolset, and beyond as data is assembled for new purposes.

# 12. Appendices

## 12.1 Confidential Information

There is no confidential information.

## 12.2 Progress against activities

### 12.2.1 Seeding the Commons

All Seeding the Commons, Data Capture, Applications and Metadata Stores ANDS funded projects are now complete and involved running over 170 discrete projects with most of the 39 research institutions based in all states and territories around

Australia. Collectively these projects have delivered a number of successful outcomes, including a cultural change within the institutions on how data is managed; resulting in a number of ongoing positions being established at

the institutions within the Library and/or eResearch Office dedicated to supporting researchers at the institution in managing data assets.

Activity	Progress
<p>To grow the data commons, ANDS will emphasise the recruitment of existing content into repositories, identifying existing repositories of useful content, and making all that content discoverable through the ARDC.</p>	<p>ANDS has encouraged this primarily through the EoI projects. The projects have generally taken one of three forms:</p> <p><b>Broad:</b> a wide audit of available data, and of policies currently in place, with a view to describing data and creating wider policy (21 projects take this form).</p> <p><b>Exemplar:</b> working with exemplar data collections within the institution, with a view to applying the lessons learnt to other areas, and to the broader data management policy framework within the institution (10 projects take this form).</p> <p><b>Combined:</b> working with the Data Capture projects funded by ANDS, to apply the lessons learnt and help create a broader data management policy framework (2 projects take this form).</p> <p>The funding of the National Linguistics Corpus project is also designed to make more useful content available through the ARDC.</p>
<p>Where institutions with valuable existing content do not have the required systems, ANDS will work with them to improve their ability to store, describe, persistently identify and register their research data assets, in collaboration with the Metadata Stores program.</p>	<p>ANDS staff are closely involved in this through the funded projects. Common elements of these projects include:</p> <ul style="list-style-type: none"> <li>• Identification of existing data collections within an institution, with a view to enabling the exposure and description of them where possible.</li> <li>• Examining existing processes, policies and workflows, in data management with the intention of using what is learnt to improve data management capabilities within the institution.</li> <li>• Creating descriptions of data collections to increase understanding of how these descriptions can best be created.</li> <li>• Building tools to enable and simplify metadata capture related to data collections.</li> <li>• Regular interaction with the Metadata Stores program occurs to provide advice.</li> </ul>
<p>Where repositories (or federations) already exist, ANDS will assist with their integration into the ARDC.</p>	<p>Being undertaken as part of the funding program, at an institutional level.</p> <p>The funding of the National Linguistics Corpus project is also designed to make more useful content available through the ARDC.</p>

Activity	Progress
<p>To enable ANDS to undertake this across the country, partnerships have been established with various eResearch support organisations. In these ANDS funds staff at the partners, whose role is to work with ANDS, the partners and other regional or discipline specific organisations to achieve the program's goals. It is ANDS' intention to co-fund these positions in 2012/13 with selected partners. Support for part of the program will also come from within the ANDS team. This includes both work undertaken through partners, and provision of specialist advice.</p>	<p>ANDS has funded staff at iVEC; now based at Edith Cowan University since Q14, Intersect, QCIF and eRSA, to work with ANDS partners in their states, and to connect ANDS to their organisations.</p>
<p>To focus funding effectively a large number of Expressions of Interest have been sought to understand better the needs and aspirations of institutions in this area. This will provide funding for institutions to better understand their data management needs, their current data holdings, and put in place systems to manage and exploit this data.</p>	<p>At June 30, 2012, ANDS had either entered into contracts (or had substantially agreed on project descriptions) for Seeding the Commons projects at 40 institutions. A breakdown of the progress made in relation to this is provided in section 2.2.3. Detailed descriptions of the contracted and agreed projects can be found in Section 10.2.</p>
<p>ANDS will also aim to share lessons learned and examples of best practice across the sector to create a community of data managers. This will involve bringing together data managers either directly employed by ANDS, or funded by ANDS, from across the country, to effectively share knowledge and experiences.</p>	<p>ANDS has continued to provide support to partners by offering advice and expertise on a one-on-one basis; facilitating exchange within the community; by participating in events hosted by institutions, the <i>eResearch Australasia</i> conference, facilitating informal sessions, roundtables, and distribution via the ANDS partner's list.</p> <p>Outcomes of ANDS projects have been disseminated through the ANDS mailing lists, news items and the website.</p>
<p>To achieve this ANDS will need to continue to develop and update an effective set of internal resources to centralise the necessary information and then to disseminate it widely. It is expected that members of the community will both add to and learn from these resources. ANDS will also work with other programs, key researchers, local bodies and overseas institutions to identify tools and infrastructure that could be co-developed to improve the quantity and quality of the data that is managed, and increase the richness of the contextual information around the data that is available. Close coordination with the Frameworks and Capabilities program will also play a key role.</p>	<p>Work has been underway in conjunction with the Frameworks and Capabilities program, focussing on selected areas within data management that have been most in demand. This includes information and analysis on data management policy, researcher interviews, data management planning, organisational structures and the data management framework. This has been disseminated through direct engagement with partners, and the ANDS website.</p> <p>Outcomes of ANDS projects have been disseminated through the ANDS mailing lists, news items and the website.</p>
<p>As the projects under this and other ANDS programs are completed it is expected that a wide range of services, documents and software will be made available. ANDS intends to examine these outputs with a view to adapting, redeploying or further developing them for use in other institutions or projects. ANDS staff will work with the institutions to achieve this. The process for undertaking this is described in the next section.</p>	<p>ANDS intends to divide access to these services and outputs in two ways: the ANDS Project Registry will list all funded projects and link to their available outputs. This is now available at <a href="http://projects.ands.org.au">http://projects.ands.org.au</a>.</p>
<p>Given the limited amount of remaining funding, decisions on future external funding will be limited to small scale projects, such as the capture of specific collections, targeted partners or the deployment of software and other outputs of the underway projects. This process will be further refined during the year. However, selection of these projects or on redeployment will be based on:</p> <ul style="list-style-type: none"> <li>• Coverage of existing discipline areas (where those areas with less coverage will need to be targeted)</li> <li>• The number of new collections that are likely to be created</li> <li>• Availability of relevant resources within ANDS</li> <li>• The ability of the potential partner to commit similar resources and to support the services developed or deployed without ANDS funding</li> <li>• The development of a catalogue of ANDS tools, derived from work funded by ANDS and elsewhere, and their suitability at potential partner institutions</li> </ul> <p>These decisions will be made within ANDS, based on the outputs of the projects listed above. As such, final decisions will be made in second half of 2011.</p>	<p>In late 2011 it was decided to direct the remaining funds in the program to Seeding the Commons projects at seven smaller universities that had not previously been funded by ANDS. Six of the seven responded (Australian Catholic University, Charles Darwin University, Southern Cross University, University of Ballarat; now Federation University, Bond University, and University of the Sunshine Coast). These projects have been completed.</p>

## 12.2.2 Frameworks and Capabilities

Activity	Progress
<b>Data citation</b> DataCite collaboration and contribution Proof of concept projects with mature archives and disciplines Documentation and events to support ANDS DOI services, events	Community of Practice was well established with early adopters showing willingness to share their experiences and support others. Solid examples of the benefits of Data Citation emerged. Documentation is strong and downloaded frequently.
<b>Institutional research data management infrastructure development</b> Contribution to pan-ANDS research data management infrastructure service offering for research organisations (policy, plans, frameworks...)	Significant progress was made with almost all universities now providing data management policies, plans and support mechanisms (see <a href="https://projects.ands.org.au/policy.php">https://projects.ands.org.au/policy.php</a> ).
<b>Data management plans</b> Working with partners International liaison Monitoring Australian policy	Significant progress was made with almost all universities now providing data management policies, plans and support mechanisms (see <a href="https://projects.ands.org.au/policy.php">https://projects.ands.org.au/policy.php</a> ).
<b>ANDS Infrastructure Usage Support</b> Support for Australian researchers and research organisations building ARDC infrastructure or using ANDS National Services	Five (5) <i>Raising your Research Profile using Research Data</i> workshops was held (Australian Digital Humanities, UniAdelaide, UWA, AIMS, JCU) with 3 more planned (UTas, IMAS, NERP). These workshops model strategies for data support staff to roll out these workshops across their institutions.
Providing approaches that support best data management practice and re-use by researchers and groups	Each workshop was tailored to the group or institution, whilst maintaining standard clear messaging (see <a href="http://www.ands.org.au/presentations/index.html">http://www.ands.org.au/presentations/index.html</a> ).
Develop a shared base of resources used by the community of practice research support staff	Ongoing: andsUP and the Data Snippets Blog (see <a href="http://andscentral.blogspot.com.au/">http://andscentral.blogspot.com.au/</a> ) augmented the website during the year.
Engage with the ANDS Community to support the development and implementation of institutional research data management policies	Ongoing: 3 workshops and webinar series were held: Data Citation; Raising your research profile using research data; and Data Management informals.
Train staff working on ANDS projects in the fundamentals of data and information management	Ongoing: All ANDS Days intensives held each year in March are aimed at building capability within ANDS staff. These were augmented by frequent Brown Bag Lunches and monthly All ANDS virtual meetings.
Develop targeted documentation to support the improvement of data management, citation, analysis, curation and preservation	Continuing: the ANDS website was continually improved.
Provide access to documentation through various media and through direct contact to support widespread access and use and reuse of research data	Increasing use was made of ANDS Google groups, Twitter and andsUP – resulting in significant increases in attendances at events and website downloads.
Contribute to the organisation of workshops and seminars sponsored by other ANDS Programs	An extremely active program in the past 12 months attracted increasing numbers of attendees (see <a href="http://www.ands.org.au/events/index.html">http://www.ands.org.au/events/index.html</a> ).
Enabling improved data provision to research through determining the economic benefits of sharing data	A major report has been commissioned for Q4 2014 to support understanding of the economic benefits of data repositories. This will continue on from Horton's 2012 Cost Benefit Analysis (see <a href="http://www.ands.org.au/resource/cost-benefit.html">http://www.ands.org.au/resource/cost-benefit.html</a> ).
Provide select organisations with resources to undertake capability building activities	Significant work was undertaken with Western Australian institutions to build a state community of practice. Two community workshops (March, June) were reinforced with individual support events for UWA, Curtin, Edith Cowan and ongoing virtual support.
Develop capability through workshops, seminars and other forums. The process for doing this involves identifying appropriate opportunities and then seeking the approval of ANDS management	ANDS worked with 25 institutions and groups during the 2013-14 reporting year (universities, ARC/NHMRC, CSIRO, AAD, ARMS, CAUL, Council of Publishing Ethics, Public Library of Science, AusGOAL, TERN, ORCID etc.) to develop capability both for their staff and their researchers.
Provision of submissions:	Australian National Data Service (ANDS) submission to "Assessing the wider benefits arising from university-based research - discussion paper" (Former Dept. Innovation). Major (25 page) input into Stream 2 ARCom Open Research Policy (Former Dept. Innovation). Multiple invited submissions to ARC and NHMRC (mainly on increasing, rather than 'open access' to the outputs of publicly-funded research). ARC: "Pathways to research data in research excellence evaluation: a report prepared by the Australian National Data Service" (focusing on how data publications might work within the ERA framework).

Activity	Progress
Contribution to committees and working groups:	<p>Attended AusGOAL Steering Committee meetings.</p> <p>Regular attendance at AusGOAL Practitioner Working Group.</p> <p>Regular attendance to Spatial Data Management Committee.</p> <p>Attended at ad hoc Office of Spatial Policy meetings.</p> <p>Contributed to several focus groups within the Office of the Australian Information Commissioner.</p>
Support for institutional engagements:	<p>New resource was developed for ANDS website, see <a href="http://www.ands.org.au/guides/ethics-working-level.html">http://www.ands.org.au/guides/ethics-working-level.html</a>.</p> <p>Published in the Research Whisperer an article on ethics and data publishing (<a href="http://theresearchwhisperer.wordpress.com/2013/07/23/data-ethics/">http://theresearchwhisperer.wordpress.com/2013/07/23/data-ethics/</a>).</p> <p>Article in Share, that forecast an upcoming detailed guide to handling sensitive data (<a href="http://ands.org.au/newsletters/newsletter-2014-07.pdf">http://ands.org.au/newsletters/newsletter-2014-07.pdf</a>).</p> <p>Worked with UA on an idealised approach to (open access) university IP policies.</p> <p>(With the new ARC Funding Rules) Reworked a major guide which outlines the roles and responsibilities of individuals and institutions (see <a href="http://www.ands.org.au/datamanagement/data-management-practice-guide.pdf">http://www.ands.org.au/datamanagement/data-management-practice-guide.pdf</a>).</p>

## 12.2.3 National Collections

Agency or Institution or Project	Description and status
Urban Water	A flagship distributed collection in Research Data Australia. See: <a href="http://researchdata.ands.org.au/australian-urban-water-collection/267201">http://researchdata.ands.org.au/australian-urban-water-collection/267201</a> Engagement was facilitated through the Urban Water Research & Development Coalition and the Partnership Working Group. Work has continued on developing the Australian Urban Water Collection in RDA with a number of meetings held with potential contributors to the collection. A revised approach to our engagement strategy was developed based on leveraging existing relationships with ANDS Partners (data providers).
TERN	Work continued with TERN to structure their metadata records so as to represent data publishing from each TERN Facility and bring this together to provide an overall TERN view in Research Data Australia. Assistance has been provided with describing entities and relationships, establishing metadata feeds, metadata management and metadata crosswalks.
Geoscience Australia	Work continued with Geoscience Australia (GA) to capture citation metadata for over 20K collection records. Building on this activity was assisting GA to implement a Digital Object Identifier minting service, currently in Production. GA also recently released a new metadata catalogue requiring ANDS to regression test the integration between the two technologies.
National Computational Infrastructure	There was significant activity with the National Computational Infrastructure including providing assistance with developing a data management policy, a metadata schema and the implementation of a Digital Object Identifier minting service.
Bureau of Meteorology	Meetings were initiated with the Bureau of Meteorology to identify additional collections for the Australian Urban Water Collection in RDA. While a number of potential collections were identified, there would be substantial additional work for the Bureau to provide these as a metadata feed to RDA. Advice was provided to the Bureau on best practice for metadata management going forward.
Population Health Research Network	Informal agreement was reached with PHRN for harvesting of metadata to RDA.
Commonwealth Department of the Environment	Input was provided to the Department's Data and Information Management Policy for the National Environmental Research Hub. The Department will use the policy to ensure researchers comply with data management guidelines. ANDS will continue to work with the Department to provide metadata creation, data management and data repository advice to its funded activities.
Public Records Office of Victoria and State Archive NSW	Public Records Office of Victoria and State Archive NSW were supported in their production and implementation of an ANDS-funded solution, resulting in over 30K Collection records. This unique collection spans from the mid 1830s to today, and includes a significant amount of digitised data. Discussions were held with QLD State Archives, who would like to implement the re-usable solution, enabling the QLD archival data to be discoverable via Research Data Australia.
NSW Office of Environment and Heritage	ANDS was approached by NSW Office of Environment and Heritage to assist with the development of a data management policy and procedures for their agency. ANDS facilitated meetings between them and other clients to share data management journeys and associated solutions. This will be an ongoing engagement.
Australian Institute of Marine Science	ANDS and AIMS worked together on several fronts to re-establish the metadata feed from AIMS to RDA. ANDS also worked on the AIMS style sheet to capture citation metadata. When AIMS implements the style sheet, citation metadata will be ingested and displayed on RDA.
e-Atlas	The National Collections program has been working closely with e-Atlas to improve the quality of their metadata within the ANDS Registry; most notably, this has included the addition of citation metadata and improved representation of relationships between research objects. This work has been re-used to make similar improvements to the metadata from AIMS and IMOS.
IMOS	Through the Australian Ocean Data Network (AODN), IMOS added an additional 2000 metadata records to RDA.
Data.gov.au portals	ANDS undertook analysis and technical activity to enable harvesting of metadata from CKAN. The CKAN technology is used by data.gov.au to host metadata and data. CKAN is also used by other government portals. Work started with data.gov.au When the data.gov.au ingest has been proven, work with the other portals will commence. Informal agreement was reached with NSW Dept. of Finance and Services (oversees <a href="http://data.nsw.gov.au/data/">http://data.nsw.gov.au/data/</a> ) for harvesting of metadata to RDA.
Thematic linkages to collections in RDA	Following the release of technical and software infrastructure to support the development of theme pages, 8 new pages representing thematic views of content, were released to RDA in June. The themes align with the <i>NCRIS 2011 Strategic Roadmap for Research Infrastructure</i> and include: Integrated Biological Diversity, Astronomy, Terrestrial Systems, Urban Settlements, Population Health Research Platforms, Solid Earth, Marine Environment, and Characterisation.

## 12.3 Risk Register

ANDS maintains a Risk Register. The risk assessment methodology, adapted from the Australian Risk Management Standard AS/NZS 31000:2009, involves identifying and analysing each risk in terms of how likely it is to happen (Likelihood) and the possible impacts (Consequence).

The key risks for ANDS in executing the Projects and the risk management strategies to be employed can be grouped into four major categories.

### 12.3.1 Political and Governance

**Risk 1 – That there are persistent negative perceptions of the Project among funding agencies and influential groups leading to a lack of buy-in**

Risk Factors:

A particular project does not have the confidence of a subsection of a community.

Lack of confidence in governance, management, or Project delivery.

Perceptions of slow engagement with areas of the sector.

Change of emphasis with regard to the policies around publicly funded research data.

Lack of certainty of the funding of the function of ANDS.

International engagement is halted as a result of limited support of ANDS.

Risk Mitigations:

The communications plans have been updated to ensure that the specific research communities have input into specific projects and their outcomes before, during and after the projects are undertaken.

Diagnostic strategies have been implemented to mitigate against failure.

Use a central point where progress of the ARDC is being tracked by metrics such as number of collections available, and numbers of data collections accessed, and the status of every project is tracked.

Clearly articulate the Project's message and brand.

Engage actively with communities to avoid perception (or reality) of not meeting its needs.

Ensure that the Project reflects the Government's expectations through constant dialogue.

Maintain close contact with key DIISCRTE officers to ensure they provide input to decision making, including having an observer on the Steering Committee.

ANDS communicates the message about the longer term vision of the function of ANDS in the sector.

Working with funding agencies on future plans for investment in the function of ANDS.

**Risk 2 – That the ANDS Project is not managed effectively**

Risk Factors:

Lack of effective mechanisms for planning, leadership and management.

The structure of ANDS has a negative impact on coordinated delivery of required activities.

Collaboration between the Project and across locations is not effective.

EIF funding guidelines do not allow for sufficient Project staff to administer funded programs of work.

State based staff have competing priorities and insufficient oversight.

Projects have insufficient time to complete.

Managers departing.

Risk Mitigations:

Management and planning processes have been put in place that include formal reporting and regular reviews to ensure the efficient conduct of the Project.

Regular meetings of Project staff are held to build a team approach. Communication structures in place to facilitate working together.

Staffing levels are monitored and adjusted as required.

Contracts and partnerships with state based organisations that host Project staff have been put in place that ensure that staff are clear about their role. Ensure that ANDS-funded staff based in organisations who are ANDS sub-contractors are not placed in a position of conflict of interest.

Ensure timely projects commencement.

Ensure all late starting projects are closely managed.

**Risk 3 – That the continued emphasis on external contracted engagements represents too big a burden on the lead agent**

Risk Factors:

University processes, focused on student and supplier engagement, are not a good fit for sector wide activities. ANDS' role as a sector wide agent in many of its programs has imposed additional requirements on the lead agent causing pressure on its staff to assist ANDS.

ANDS EOI approach generates clusters of work with tight timelines that impact on specific university functions such as the Solicitors' Office and Finance.

Risk Mitigations:

Approval has been obtained for streamlined approaches at Monash University to enable ANDS to work more effectively.

Fund additional staff or specific work at Monash University to enable ANDS to work more effectively.

ANDS reduces the number of projects that are externally funded.

Ensure ongoing, strong engagement with the Research Sector, including research infrastructure capabilities.

All activity plans were developed after consultation with relevant stakeholders.

Membership of the Steering Committee includes key stakeholders.

Performance measurement for the Project should include effective stakeholder engagement.

Effective communication of benefits to stakeholders.

Provide a clear rationale behind the decision process for project funding.

Communications activities have been increased to create awareness of the value of ANDS' activities.

ANDS effort has been increased in creating partnerships as compared to contracting.

Formal procurement processes have been implemented to ensure that the requirements are understood and that potential suppliers meet the set criteria.

Provide ongoing contract management to ensure the delivery of required outcomes to the contracted service levels.

Effective vendor and partner engagement approaches have been put in place.

### 12.3.2 Relationships

#### Risk 4 – That the Project's external stakeholders are not effectively engaged

Risk Factors:

Stakeholders are not prepared to undertake the changes within their own organisations that are necessary for the realisation of the ARDC.

Stakeholders do not see their interests in data management and those of the Project as being aligned.

Risk Mitigations:

Maximise the effectiveness of connections between the Project and related eResearch and other initiatives, including involvement of groups outside ANDS in the ANDS Policy Forum, the ANDS Technical Forum, and the ANDS Content Forum.

Ensure that ANDS' engagement with stakeholders meet their research data ambitions as well as ANDS' requirements.

#### Risk 5 – That the Project's partners do not appropriately contribute to the Project

Risk Factors:

Partner produces outcomes of low quality or does not meet the requirements of the contract.

Partner expends funds in a way that is not consistent with the EIF guidelines.

Lack of effective arrangements in place to ensure the contracted services are provided to an agreed service level.

Service providers see themselves as disconnected from the Project's decision-making or strategic planning.

Risk Mitigations:

Collaboration Agreement is in place to manage output and management of joint venture partners.

#### Risk 6 – That ANDS is not perceived as a long-term partner and hence the services are not taken up

Risk Factors:

The impending end of ANDS NCRIS and EIF funding together with the different purposes of CRIS and NCRIS 2 funding, causes a perception that ANDS initiated services will not continue.

Risk Mitigations:

ANDS gained approval to expend existing funding over longer timelines (consistent with other Super Science funded activities).

ANDS creates reliable sustainable services that are offered over the longer term by other long term service providers.

Securing CRIS funding and mapping ANDS services through this new funding regime will preserve long-term services.

Strong contribution to DIISCCRTE Roadmap, RDIC and NRIP processes will be a mitigating factor.

**Risk 7 – That there is confusion about role of ANDS versus other related service providers in the e-Research sector which impedes effective service delivery**

Risk Factors:

ANDS and eResearch infrastructure partners' offerings are confused by possible users.

Relationship between ANDS and state-based eResearch providers (such as Intersect) is not clear to users.

Greater expectation of collaboration between eResearch infrastructure partners based on Research Data Infrastructure Committee (RDIC) report.

Risk Mitigations:

Ensure that ANDS' communications to a range of stakeholders provide greater clarity about ANDS services.

Ensure that ANDS' offerings are clearly targeted and that this is clearly stated.

Seek greater clarity from other eResearch service providers about their offerings, avoiding either actual or perceived overlap with ANDS' offerings.

Increased coordination of offerings by eResearch service providers through eResearch Infrastructure.

Discussion with NCI, NeCTAR and RDSI taking place to ensure clarity of eResearch service offerings.

Ensure RDIC provides guidance for improved communication.

### 12.3.3 Impact

**Risk 8 – That data providers/federators do not make their data available**

Risk Factors:

The storage needs of researchers are not met, so will not consider sharing their data.

Researchers do not wish to share their research data.

Confidentiality agreements prevent researchers from making their data available.

Existing data federations see insufficient value in making their data available.

Risk Mitigations:

Strategically promote incentives and rewards in the research system for data publishers.

ANDS will co-ordinate with RDSI and Institutional stores to mitigate this risk.

Enable data citation so that researchers get recognised for the publication of their research data.

Encourage the use of access controlled data stores.

Ensure that ethics agreements balance confidentiality with openness.

Recommend that funding be linked to the provision of data via the ARDC as it becomes available.

Provide targeted assistance to data federations to assist with integration into the ARDC.

**Risk 9 – That re-users of research data do not use ANDS Services to discover, access and exploit data**

Risk Factors:

The various strategies for exposing data in the ARDC do not result in the data being easily discoverable.

Access control mechanisms are too restrictive or complex.

Other sources of data for re-use are more attractive or easier to use.

Risk Mitigations:

Ensure a nuanced and multi-faceted approach to exposing the Project's accessible data.

Work with AusGOAL and the Australian Access Federation to identify a simple set of licensing and standard access control policies.

Ensure that it is easy to re-purpose ARDC accessible data.

**Risk 10 – That the standards and technologies that ANDS adopts are not adopted more widely**

Risk Factors:

ANDS is the only user and maintainer of actual or *de facto* standards, leading to inability to share maintenance and development costs.

ANDS is the only source of development activity on particular technologies (RIF-CS, ORCA, ANDS Handle code).

Risk Mitigations:

Promote community ownership of standards such as RIF-CS, for example through community-led advisory boards.

Seek international engagements and partnerships to take up standards and technologies favoured by ANDS and share development load.

Ensure enough people are trained on the standards and technologies that ANDS is adopting to support wide adoption.

Make implementation decisions such that ANDS is not dependent on particular standards and technologies, but on general approaches that can be transferred across technologies.

Encourage the use of ANDS-developed technologies by other data aggregators such as Terrestrial Ecosystem Research Network (TERN).

## 12.3.4 Resourcing

### **Risk 11 – That high quality staff are hard to recruit and retain**

Risk Factors:

Limited availability of skilled staff (both within ANDS and in ANDS-funded projects) impacts ability to perform tasks funded by ANDS.

Funding uncertainty leads to potential for staff departures.

Risk Mitigations:

Build a vision for the function of the ANDS for the longer term and communicate this to staff.

Provide as much certainty to staff as is possible, and involve them in navigating the future.

### **Risk 12 – That ANDS is most effectively reporting under all the various programs**

Risk Factors:

Staff needing to report on similar activities multiple times throughout the year as per funding agreements that have been signed with the Commonwealth Government.

Risk Mitigations:

Develop a reporting calendar and circulate amongst relevant staff and provide reminders.

This page is intentionally left blank





[ands.org.au](https://ands.org.au)