Australian National Data Service (ANDS)

BUSINESS PLAN 2014-15
# INDEX

1. Executive Summary  
2. ANDS Context and Approach  
3. Status of Project  
   3.1 Establishment  
   3.2 NCRIS funding  
   3.3 Super Science Funding  
   3.4 Additional Funding  
   3.5 Time Extension and Reporting  
   3.6 NCRIS 2013 Funding  
   3.7 The Function of ANDS  
   3.8 ANDS Programs and their Transformation  
   3.9 ANDS Principles  
   3.10 Scope  
   3.11 ANDS’ Impact  
4. Research Infrastructure  
   4.1 National Collections  
   4.2 National Services  
   4.3 Institutional Engagement  
   4.4 International Collaboration  
   4.5 Overall 2014-15 Expected Outcomes  
5. Confidential Information  
6. Access and Pricing  
7. Summary Financial Information  
   7.1 Expenditure by Program  
   7.2 Expenditure of funds over the life of ANDS  
   7.3 Co-investment  
   7.4 Expenditure of funds by type  
8. Governance  
   8.1 Steering Committee  
   8.2 Management structure  
9. Risk Management  
   9.1 Political and Governance  
   9.2 Relationships  
   9.3 Impact
9.4 Resourcing

10 Key Performance Indicators
   10.1 Key Performance Indicator Series
   10.2 Estimates against Key Performance Indicators for 2014-15
1 Executive Summary

The Australian National Data Service (ANDS) was established by the Australian Government in January 2009 following the 2007-08 ANDS Establishment Project. ANDS was originally created as part of the National Collaborative Research Infrastructure Strategy (NCRIS) initiative to ensure that research data is used as effectively as possible by Australian researchers. Additional funding from the Education Investment Fund (EIF) was received in May 2009 to establish the Australian Research Data Commons (ARDC). The EIF was a component of the Super Science initiative. The combined NCRIS and EIF funds provided an opportunity to build a co-ordinated set of programs through to June 2013. As a result of additional Collaborative Research Investment Scheme (CRIS) funding, and then NCRIS 2013 funding, ANDS has been funded to operate until June 2015. Expenditure against original NCRIS and EIF funding will have concluded by June 2014, so this plan describes activity funded by the CRIS and NCRIS 2013 initiatives.

ANDS exists to transform Australia’s research data environment by making Australian research data collections more valuable though managing, connecting, enabling discovery and supporting the multiple use of this data. The purpose of this activity is to enable richer research, more accountable research, more efficient use of research data, and improved provision of data to support policy development. The outcome of this activity will be that Australia’s research data as a whole becomes a nationally strategic resource.

In this document ANDS’ plans for the 2014-15 financial year are described in detail. This plan is in accordance with ANDS’ extended strategic direction and describes an increased focus on data. Rather than institutions just focussing on meeting ANDS’ goals, the team will encourage organisations, research groups and researchers to realise their own research data ambitions.

The work of ANDS will continue to be carried out under four programs:

- **National Collections** of government, institutional, discipline and national facility data.
- **National Services** including registration, publication, discovery, and advisory services.
- **Institutional Engagement** with all of Australia’s major research organisations.
- **International Collaboration** with data infrastructure providers to ensure that Australian research data infrastructure is compatible with international approaches.

Significant progress has been made to date, and by June 2014 ANDS will have:

- Established several national services: a data collections registration service, a dataset identification service through the DataCite consortium, a data collection description publication service, a researcher identification service in partnership with the National Library of Australia, and Research Data Australia—a data collections discovery service.
- Populated the ARDC with data collections that have been managed and connected, providing over 90,000 collections descriptions discoverable through Research Data Australia, Google and other mechanisms.
• Helped establish coherent institutional research data infrastructure at 32 Universities, CSIRO, the Australian Synchrotron and ANSTO, including automated tools for capturing rich metadata from instruments, metadata stores, and pipes connecting to institutional systems and national services.

• Improved the ability of the Australian research system to exploit the ARDC with guidance on research data management, including responding to the *Australian Code for the Responsible Conduct of Research*.

• Improved data management around the country with the ability to support institution responsiveness to ARC funding requirements.

• Developed new tools that enable more effective reuse of research data, with this use demonstrated through a range of high profile examples across a number of disciplines.

During 2014-15, ANDS will:

• Ensure collections that are formed, whether by NCRIS capability, public sector or research institutions, are able to be published appropriately though integrated services.

• Maintain access to institutionally and nationally significant collections as openly as possible.

• Continue to deliver valued and reliable national data technical and advisory services.

• Continue to populate the ARDC with more data collections that have been managed and connected, ensuring over 90,000 collections descriptions discoverable through Research Data Australia, Google and other mechanisms.

• Continue to support research institutional capability and capacity to manage research data.

• Maintain ANDS as a trusted partner of research institutions.

• Maintain Australia’s leading international role in research data infrastructure.

By the end of 2014-15 researchers across every discipline and at nearly every research institution will be represented in the ARDC, and nearly all research institutions will have improved their research data management, leading to routine publication of their data with ANDS persistent identifiers into a data store that feeds information to the ANDS collections registry. In addition, researchers will be able to find and use a wide variety of data sets using the ANDS data pages through a variety of discovery paths, and more institutions will be successfully engaged in meeting their responsibilities described in the *Australian Code for the Responsible Conduct of Research*. They will support researchers with good data management reflecting the ARC mandated data management plans. ANDS will have helped Australia maintain a leading international role in research data infrastructure. Importantly ANDS will continue to be a trusted partner of the research community in transforming research data practice.
2 ANDS Context and Approach

Research is becoming more data intensive, and the data is becoming more complex. Moreover, the problems being tackled are increasingly large scale and can span multiple disciplines. Consequently, and as a result of high level Government reviews, the Government has invested in improving the Australia’s research sector’s capability to use and reuse research data. This was guided first by the NCRIS roadmaps and then by the document entitled Towards the Australian Data Commons \(^1\) (TADC).

In support of this goal, Towards the Australian Data Commons identified a range of objectives for ANDS. These objectives were based on the belief that “ANDS can contribute most effectively by developing services and activities that enable stewardship within multiple federations of data management and data user communities” (p. 6). TADC identified a number of longer-term objectives for data management:

a) A national data management environment exists in which Australia’s research data reside in a cohesive network of research repositories within an Australian ‘data commons’.

b) Australian researchers and research data managers are ‘best of breed’ in creating, managing, and sharing research data under well-formed and maintained data management policies.

c) Significantly more Australian research data is routinely deposited into stable, accessible and sustainable data management and preservation environments.

d) Significantly more people have relevant expertise in data management across research communities and research managing institutions.

e) Researchers can find and access any relevant data in the Australian ‘data commons’.

f) Australian researchers are able to discover, exchange, reuse and combine data from other researchers and other domains within their own research in new ways.

g) Australia is able to share data easily and seamlessly to support international and nationally distributed multidisciplinary research teams. (p. 6)

As a result of these goals, initial activity, and consultations, ANDS role is to enable Australia’s research data to be transformed, as suggested by the following table:

<table>
<thead>
<tr>
<th>From Data that are:</th>
<th>To Structured Collections that are:</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Unmanaged</td>
<td>• Managed</td>
</tr>
<tr>
<td>• Disconnected</td>
<td>• Connected</td>
</tr>
<tr>
<td>• Invisible</td>
<td>• Findable</td>
</tr>
<tr>
<td>• Single use</td>
<td>• Reusable</td>
</tr>
</tbody>
</table>

This will deliver a nationally significant resource so that Australian researchers can easily publish, discover, access and use Australian research data.

ANDS is doing this by first creating the Australian Research Data Commons (ARDC), the focus of the Super Science project. 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. ANDS does not hold the actual data, but points to the location where the data can be accessed. The ARDC can be envisaged below, where ANDS is contributing to the green pipes and boxes:

![Figure 1: Australian Research Data Commons](image)

ANDS is thus creating a combination of national services and coherent institutional research data infrastructure, combined with the ability to exploit that infrastructure with tools, policy and capability. To deliver against these objectives, ANDS has four inter-related programs of activity (Institutional Engagement, National Services, National Collections, and International Collaboration).
The ANDS activities (whether CRIS funded or NCRIS 2013 funded) are being conducted under the same management structure which can thus maximise opportunities for cost savings. However, as required by contract, separate reports will be provided to the Department for the CRIS project and NCRIS 2013 project. ANDS has also entered into two other Funding Agreements with the Department. The first agreement was to progress outcomes arising from the First Joint Australia-EU Research Infrastructure (RI) Workshop in June 2011. Subsequently, in 2012, two more agreements were entered into, for the Second Joint RI Workshop, and the DataWeb Forum. All of these activities will have concluded by June 2014. This business plan provides a combined view of how ANDS intends to execute CRIS and NCRIS 2013 activities in 2014-15.

The programs, or components, are:

- **National Collections**, ensuring that the formation and curation of research data collections, together with associated services, is maintained.
- **National Services**, ensuring that the research data services, both technical and advisory, that have been established under the ANDS programs to date are maintained.
- **Institutional Engagement**, ensuring that institutional research data infrastructure is maintained, and continues to develop, using institutional funds, in a cost effective and nationally coherent manner.
- **International Collaboration**, ensuring that the research data collaboration internationally is maintained, principally through the established international partnerships, notably the Research Data Alliance.

These four programs, taken together, will enable ANDS to continue its data partnerships with:

- Our research institutional partners – focused on the Universities and the PFROs.
- Our research data providing partners – this includes NCRIS capabilities, national data generating facilities, public and private sector data providers, international data providers, and the public.
- Our research data infrastructure partners – this includes RDSI, NeCTAR, NCI and Pawsey, regional data service providers, and institutional eResearch partners.
- Our international partners, notably through RESEARCH DATA ALLIANCE, but including the US, the EC, New Zealand, U.K, the Dutch, amongst others.

The particular focus of this phase of ANDS activity is research data. We will ensure that the infrastructure that has been set up through the Australian Research Data Commons is maintained and operated, but with an increased emphasis on making data visible, available and reusable. This will be carried out using both the internal resources of ANDS as well as partnerships funded by three engagements:

- **Major Open Data Collections engagements** – partnering with research institutions to promote collections of enduring significance to institutions that support their research strategy and their research partnerships strategy. The result of these engagements will be at least twenty five internationally significant open data collections which are supported by the institutions and which will exploit Australia’s research data infrastructure.
- **The Access to Data for Research Infrastructure Capability Areas engagement** will enable ANDS to partner and fund the six capability areas that are data intensive and which already have a strong ANDS partnership. This will provide continued support for data collection publication and will
enable services that are being generated through the Nectar Virtual Laboratories to be connected, provide long term access to data, reliable access to collections via rich descriptions, and connections to international collections through Research Data Australia.

- The Maintaining Connections engagements will ensure that connections between institutional metadata stores and RDSI node data registries, will enable institutions to have a comprehensive view of their research data assets, as well as ensuring connections between the collections held on RDSI nodes, their descriptions, in RESEARCH DATA AUSTRALIA, and the developing data services, such as ones being developed through the Nectar Virtual Laboratories are connected.

These outcomes, in concert with previous year’s activities, will produce the following desirable consequences.

**Better Data**

- A national resource of managed, connected, findable and re-usable research data now exists that previously did not.
- More nationally-significant open research data collections will be available.
- More data are being automatically captured from a wide range of instruments, ranging from small sensors to major national facilities.
- Data providers are increasingly seeing the value in providing their data through richer shared environments.
- Data are now visible to researchers where previously they were not.
- Researchers can more easily find and access to public sector data.

**Better Research**

- A cohort of leading Australian researchers has demonstrated the value of a richer data environment.
- Australian researchers can publish and cite their research data outputs.
- Larger questions and grander challenges can be addressed by a data commons that spans facilities, institutions, disciplines and sectors.
- New types of research are now possible that previously were not.
- Richer data environments can facilitate evidence-based policy.

**Better Institutional Capacity**

- Data can be captured, managed, and shared more easily at most research organisations in Australia.
- Research organisations and public sector agencies now provide policy support and invest more in data management.
- Australian research institutions can more easily comply with the requirements of the Australian Code for the Responsible Conduct of Research and support ARC requirements for data management planning.
- There is an established community of data management professionals.
- Australian researchers and research organisations have better capability to operate and exploit the new research data infrastructure.

**Better Investment**

- More research data is openly available for more researchers.
- There is more cost-effective reuse of research data software in Australia.
- Australia has an improved ability to measure the quality and impact of all of its research outputs.
- Research Data Australia reduces duplication of effort and cost of unnecessary data acquisition by facilitating reuse.
- Common approaches, standards, and services have increased efficiency and coherence across the national data infrastructure investment.
- Established data management infrastructure now allows for stronger national policy and easier institutional compliance.

**Better International Engagement**

- Australia is seen to be an international leader in research data infrastructure.
- Australian researchers can more easily engage with international partners on data intensive research.

As a consequence of these individual outcomes there is an overarching outcome; Australian researchers now have access to infrastructure that enable them to:

- Systematically, reliably and authoritatively connect their research data to project, institutional and disciplinary descriptions.
- Simultaneously publish citable research data collections through institutional, disciplinary and national services.

This will ensure that Australia has a mature, globally-leading capability in research data, making it a key locus for data intensive research. This capability will be demonstrated by leading researchers in a variety of disciplines to show the power of this infrastructure to enhance their research.
3 Status of Project

The status of the Australian National Data Service can best be understood in terms of its four sequential stages: establishment, initial NCRIS funding, Super Science funding, additional funding, and time extension.

3.1 Establishment

The Australian National Data Service can trace its beginnings back to the Platforms for Collaboration (PfC) capability as part of the development of the National Collaborative Research Infrastructure Strategy. During the course of the PfC facilitation process, a number of workshops were held to determine the activities that might be included in the investment plan to assist research data management. Following the approval of the overall PfC investment plan by NCRIS, an implementation workshop with wide representation was held to confirm the proposal to establish the Australian National Data Service (ANDS). This workshop took place on May 29, 2007. It endorsed the ANDS concept and proposed that a technical working group (the ANDS TWG) should be formed to draft a more detailed statement on the purpose and goals for ANDS, moving beyond the conceptual definition provided in the PfC investment plan. This working group met both physically and virtually over the course of 2007, and in October produced Towards the Australian Data Commons: A proposal for an Australian National Data Service.

In late 2007, the former Department of Education, Science and Training (DEST) asked Monash (as the lead agency) to work with ANU and CSIRO on a project to take the next step and establish the Australian National Data Service (ANDS). The ANDS establishment project concluded in December 2008.

3.2 NCRIS funding

The ANDS Draft Interim Business Plan was submitted in September 2008, and the Interim Business Plan was submitted in December 2008. ANDS commenced officially in January 1 2009. By March 1, 2009, ANDS had 16 staff. NeAT Round 1 projects were underway and a second round of NeAT projects was identified.

3.3 Super Science Funding

In the May 2009 budget, the Commonwealth government announced a series of initiatives collectively labelled as Super Science. The ANDS 2009-10 business plan was submitted in March 2009 (prior to this announcement) and accepted in July 2009 (post this announcement). The substance and execution of this plan was substantially affected by the ARDC project (announced under the

\[http://ands.org.au/towardstheaustraliandatacommons.pdf\]
Super Science program and funded from EIF). Consequently considerable effort was expended on creating a project plan for the ARDC that was complementary to the NCRIS-funded activities.

The ANDS Steering Committee decided in mid-2009 to recommend to the (former) Department of Innovation, Industry, Science and Research (DIISR) that ANDS manage the NCRIS-funded and EIF-funded activities as an integrated project. The Steering Committee also decided to reshape the portfolio of ANDS programs to better reflect the implications of, and constraints on, the added funding. As a consequence, the existing separate Frameworks and Capabilities programs were merged, and the Utilities program was moved from NCRIS-funded to EIF-funded and renamed ARDC Core. Four new EIF-funded programs were instated: Data Capture, Metadata Stores, Public Sector Data, and Applications. In the period July 2009-March 2010, ANDS consulted widely on these changed plans, and after some fine-tuning to respond to consultation feedback commenced executing against them.

3.4 Additional Funding

As previously described, ANDS has also entered into three other Funding Agreements worth a total of $462K with the Department. The first agreement was to progress outcomes arising from the First Joint Australia-EU Research Infrastructure (RI) Workshop in June 2011. Subsequently, in 2012, two more agreements were entered into, for the Second Joint RI Workshop, and the DataWeb Forum. The DataWeb Forum has provided partial funding for Australia’s participation in the establishment and development of the Research Data Alliance (previously called the DataWeb Forum).

The most recent funding agreement was to maintain this research infrastructure under the Collaborative Research Infrastructure Scheme (CRIS). This business plan provides a combined view of how ANDS intends to execute all of these activities in 2013-14.

3.5 Time Extension and Reporting

In April 2010, the Department was asked whether a short extension might be possible for ANDS, and they advised that there was an opportunity to extend for a further period of two years to harmonise with other NCRIS and EIF investments. ANDS staff and the Steering Committee managed to identify shifts of funding and timing across reasonably permeable boundaries that still delivered a viable ANDS, one able to continue to deliver on behalf of the Australian research community through to June 2013. This extension of time required a re-allocation of funds between programs, as well as a change to the funding profile within programs. It also required the funding of the project office over a much longer period. Consequently, an additional $0.5M was provided by the Department under NCRIS funding to support the operation of ANDS over a longer period. A three-year high-level project plan was developed, so that ANDS:

- Honoured all existing commitments.
- Continued with existing partnerships - this means continuing to actively engage with the research institutions.
• Retained the capacity to work with data champions.
• Maintained an ongoing capability of engaging with the sector.

A further extension was sought in order to most effectively use the CRIS funds, as they were sufficient for only 7 months of the 18 month period from July 2013 to December 2014. The project plan shows an uneven level of expenditure (as ANDS had already made substantial commitments) but does balance the need to engage with the sector over a longer period of time, and to demonstrate value early. The chart in the finance section shows the intended expenditure pattern for the various programs. This chart shows the expenditure against the newly constituted programs, mapping the previous structure to the new structure.

Another important variation to the ARDC project contract, agreed in March 2011, replaced quarterly milestone reports being delivered individually with annual reporting that incorporates reports on ARDC progress as well as NCRIS progress.

During this time ANDS reorganised its programs into four merged programs: National Collections, National Services, Institutional Engagement, and International Collaboration.

### 3.6 NCRIS 2013 Funding

ANDS was offered $13.05M additional funding to maintain Australia’s research data and research data infrastructure. This funding is not to develop new capability but to protect Australia’s research data infrastructure advantage to ensure Australia remains a research data partner of choice and to ensure infrastructure developed to data is maintained. The decision was taken to emphasise maintenance of national services, and access to research data collections, notably open data collections, over research data infrastructure that has been established at institutions.

### 3.7 The Function of ANDS

ANDS exists to transform Australia’s research data environment by making Australian research data collections more valuable though improved management, more richly connecting data, enabling discovery and supporting the multiple use of this data. The purpose of this activity is to enable richer research, more accountable research, more efficient use of research data, and improved provision of data to support policy development.

The ANDS project now runs to June 2015 after starting in January 2009. The function of ANDS to help transform Australia’s research data environment will not be completed by June 2015. As a consequence ANDS will develop a map of the function of ANDS over 10 years informed by *Towards the Australian Data Commons*. This will determine future directions necessary to ensure the effective delivery of the broader 2011 Strategic Roadmap for Australian Research Infrastructure.
3.8 ANDS Programs and their Transformation

Originally there were nine programs established to conduct the work of ANDS and these have been modified, concluded and reshaped over time into the current four programs:

- **National Collections** of government, institutional, discipline and national facility data.
- **National Services** including registration, publication, discovery, and advisory services.
- **Institutional Engagement** with all of Australia’s major research organisations to ensure effective research data technologies and infrastructure.
- **International Collaboration** with data infrastructure providers to ensure that Australian research data infrastructure is compatible with international approaches.

The following table shows the intended size and focus of the programs over the whole funding period. It takes into account interest earned as well as project funding (excluding CRIS funding).

<table>
<thead>
<tr>
<th>Programs</th>
<th>NCRIS 2013 ($M)</th>
<th>CRIS ($M)</th>
<th>Other ($M)</th>
<th>%</th>
<th>Focus</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Collections</td>
<td>1718.11</td>
<td>391.62</td>
<td></td>
<td>12.93</td>
<td>ANDS</td>
</tr>
<tr>
<td>National Services</td>
<td>2238.79</td>
<td>1031.77</td>
<td></td>
<td>20.04</td>
<td>ANDS</td>
</tr>
<tr>
<td>Institutional Engagement</td>
<td>7541.32</td>
<td>1211.29</td>
<td></td>
<td>53.63</td>
<td>Institutions</td>
</tr>
<tr>
<td>International Collaboration</td>
<td>838.17</td>
<td>86.68</td>
<td>174.76</td>
<td>6.74</td>
<td>ANDS</td>
</tr>
<tr>
<td>Governance and Management</td>
<td>744.03</td>
<td>344.68</td>
<td></td>
<td>6.67</td>
<td>ANDS</td>
</tr>
<tr>
<td><strong>Total Allocated Funds</strong></td>
<td><strong>13080.41</strong></td>
<td><strong>3066.03</strong></td>
<td><strong>174.76</strong></td>
<td><strong>100.00</strong></td>
<td></td>
</tr>
</tbody>
</table>

**Notes:**

- National Collections distributions are through a combination of CSIRO and Monash with Monash disbursing contract payments directly.
- National Services distributions are through a combination of ANU and Monash with Monash disbursing some contract payments directly.
3.9 ANDS Principles

In responding to the new objectives and program requirements, ANDS continues to follow its foundational principles:

**Commons Framework**: ANDS has started in a way that anticipates the need to scale up and adapt over time via an extensible framework of data stores, federations and services that enable better data creation, capture, management and sharing.

**Focus**: ANDS will continue to identify and work with those who are ready, willing and able to contribute significantly to the ARDC vision, and who provide the most strategic return to the ARDC for the effort expended. However, ANDS will endeavour to support all of the larger research institutions directly, in order to rapidly achieve critical mass.

**Content**: ANDS is initially focussing on content recruitment into stores and federation across stores so as to achieve a wide coverage of data quickly at an agreed level of quality; in later years the emphasis will shift towards quality improvement.

**Service Provision**: ANDS is focussed on service provision and infrastructure development, not research and exploration; its programs will develop, integrate, and continually improve production-level systems in support of well-understood services.

**Strategic Partners**: ANDS recognises the need to be open to, and engage appropriately with, innovations and external institutions relevant to the ARDC, including the Australian Access Federation (AAF), the National Computational Facility (NCI), the National eResearch Collaboration Tools and Resources (NeCTAR) and the Research Data Stores Initiative (RDSI).

**Stores**: ANDS assumes an environment where storage and long-term curation occur in nationally or institutionally-supported stores, either existing or brought into being over the life of ANDS. These stores will preferably hold objects described by various discipline-specific and documented metadata schemas. ANDS will work with whatever repositories exist, national, institutional or disciplinary.

**Sustainability**: Research data management requires a long-term commitment. ANDS has developed its plans on the assumption that the current funding does not represent a one-off investment in data. The enduring changes forecast in this document within each program are also intended to be sustainable beyond the end of the ANDS planning period.

3.10 Scope

**Constituency**: ANDS works with a variety of publicly funded institutions that produce, manage or consume research inputs and outputs to achieve its aims. The scope includes:

- All Higher Education Providers in Australia
- All research organisations that are publicly funded, including CSIRO, GeoScience Australia (GA), Bureau of Meteorology, Australian Bureau of Statistics (ABS), Australian Institute of Marine Science (AIMS), the Australian Antarctic Division, Departments of Primary Industry
• Members of the cultural collections sector (galleries, libraries, archives and museums).

ANDS is funded to work with all research disciplines in Australia, not just the NCRIS capabilities. This means that the specific concerns of the Humanities and Social Sciences are also taken into account.

**ANDS Community:** The ANDS Community consists of providers of research data and ANDS services, consumers of research data and those services, and managers of research inputs/outputs. This includes key stakeholder aggregations such as CAUDIT (The Council of Australian University Directors of Information Technology) and CAUL (The Committee of Australian University Librarians). The ANDS Community includes the general public only to the extent that they will be able to use some ANDS services to discover and access publicly available data.

**Data:** ANDS is concerned with the digital data that is produced by researchers as well as data that is used by and made accessible to them. Data is the information that researchers study, that is transformed by researchers and produced by researchers. Research publications are not included within the scope of ANDS but files, images, tables, databases, models, computer outputs, and similar digital representations are included. ANDS will support the ability to create links between data, publications, software code and visualisations, where these may appear as either research inputs or research outputs.

### 3.11 ANDS’ Impact

#### 3.11.1 Research Data Collecting Institutions Overview

**Collections Approach**

As the concept of data publication and sharing gains traction—and the volume of data available in the Australian Research Data Commons increases—the value of taking a collections approach is becoming apparent. This approach has long been used in libraries to introduce focus and applicability into the presentation of material. Collectors can track content across time series, locations or related subject. The heart of this approach is to partner with institutions to manage their data assets.

In ANDS this approach allows the highlighting of data collections especially those of national significance. It will work with collecting institutions to bring this material together, describe it in a way that enables flexibility in presentation, helping to ensure the data is as open as is possible, and highlight it for discovery. Enabling this approach to the management of data provides the following benefits:

- Enables increased discoverability and ‘browsability’ of data collections.
- Provides a more complete picture of data on any given topic.
- Shines a spotlight on important and significant data.
- Enables new and more complex research questions to be addressed.
By working with collecting institutions, ANDS can leverage their understanding of the data and research environment in which it is created and used to develop the data descriptions that enable it to be presented in meaningful contexts.

**Collection Types**

There are a number of ways in which this collecting may occur.

1. **Single Significant National Collections**
   
   These collections are collected and maintained by a single institution or unit. They can be definitive, comprise specimens or observations and may be seminal in nature. Their function may be likened to reference material in libraries and be of broad uptake outside a specific discipline. However there may also be collections in this category that are largely confined to a specific discipline, but may have unexpected uses. Curation of these collections lies with a single institution. They can be seen as a significant institutional asset.

2. **Collections brought together for a specific collaboration**

   These are often gathered from a range of sources and they could also cross a range of disciplines. They are brought together to address a specific research activity and may be transitory in nature depending on the size and longevity of the research activity. The question of the continuing maintenance of the collection, once the collaboration is complete, will arise. Datasets may be repurposed. Curation of these data collections will be as varied as the number of collaborators unless data curation is identified as a specific activity within the collaboration. This may then become an issue when the collaboration is complete and funding for curation ceases.

3. **Major Open Data Collections**

   Data collections can enhance the strategic aims of a research institution, and strengthen partnerships. An institutionally focused major open data collection leverages significant institutional and national research data infrastructure to create one or more signature collections that are supported by the institution, using their data management capacity.

4. **Distributed Collections**

   One of the values of a national metadata store is that it provides the ability to pull together ‘like’ data. There are two ways this can come about. One is a coordinated approach to collecting from institutions (collecting with intent) in which relevant institutions describe and publish collections with a view to their use. Examples would include the collections that are in discipline portals. Another is to do this on the fly within Research Data Australia using the metadata. This method allows a ‘slice’ to be taken that crosses disciplines and formats and profiles a collection based around a topic or location. Thus a collection could be formed of the data within RDA on water that includes observations, models, photographs, historical anecdotes, commentary on water use, socio-economic impact statements etc. It has the potential to provide wider and deeper insight into a given issue. It also delivers value in the serendipitous exploration of data for innovative research and strongly supports the addressing of the ‘big’ questions: complex
research issues. Additionally it presents data in a way that supports e-research methodologies. A characteristic of this type of collection is its fluidity. A dataset can belong to more than one collection and be presented in many contexts.

5. Institutional profiles

While not necessarily a national collection, a ‘collections approach’ has the value of being able to showcase the data produced by an institution, either as research output or as support for Australia’s research effort. The value of Australia being able to showcase the data outputs of its institutions will contribute to attracting research and funding internationally.

Collections Engagement

Through previous ANDS Institutional Engagement and National Collections program ANDS has engaged with institutions to support their data collections and to support the role of data collecting. Our focus in the National Collections program will be on those collections of national significance, particularly those collections that may be made available through an RDSI node. In all cases the engagement has been structured to enhance the value of those collections through improved management, connectedness, discoverability, and reuse.

3.11.2 Research Institutional Engagement Overview - Coherence

As many institutions are now reaching significant milestones in their data management maturity, two particular aspects of ANDS’ support for institutional research data management are emerging:

- The continuing relevance of ANDS’ engagement with the institution as a whole.
- The importance of institutions undertaking an enterprise approach in a range of areas in order to deliver the ANDS “four transformations” effectively (viz. table on p.5).

At the same time, ANDS is ramping up the provision of services to allow for consistent ways of publishing, discovering, accessing and using data.

In order to bring together these services and the institutional approach, ANDS is encouraging a coherent approach to research data infrastructure; without this, Australian research will be unable to reach its full potential.

ANDS plans to engage with its research institutional partners through a coordinated approach that will assist partners to achieve this coherence and realise the following benefits:

- Develop knowledge of the full extent of the institution’s data holdings and the subsequent ability to present it as their body of work.
- Secure valuable data assets for validation and reuse.
- Improve the institution’s ability to collaborate through a consistent and supported data management infrastructure.
Position the institution to be fully acknowledged in the reuse of its data by others and the consequent reputational benefits.

Improve the institution’s ability to demonstrate research excellence.

The development of coherent institutional research data management across the sector ultimately supports the building of a sound ARDC infrastructure.

This approach is a way of delivering a combination of services, software, policy and resourcing that will enable a research-producing institution to develop and provide the necessary infrastructure so that its researchers can effectively manage their data to be published, discovered, accessed and used.

The aim of the approach should be to present a unified, consistent picture of the infrastructure, policy and procedures that ANDS thinks an institution needs to have in place for effective research data management, publishing and research support. The approach should actively support the four transformations as seamlessly and effectively as possible.

An institution with a coherent approach is one whose data infrastructure can be accessed and used across the organisation, as well as being consistent with infrastructure approaches across the sector. ANDS will continue to establish services that enable and support this form of coherence across the sector. Solutions should not limit the ability of an institution to interact with the broader environment in the sector. The aim of the approach should be to enable every data collection that can be made available to be made available in a timely fashion.

The way the approach is implemented, and the focus areas that the institution chooses to address, should be in line with the larger data aspirations of that institution. That is, they should enable the institution to participate in collections of particular interest or value to them.

ANDS’ role will be in encouraging, supporting and/or funding institutions to have the following coherent capabilities in place:

1. **An institution-wide data management policy** that includes relevant planning options for researchers where appropriate.

2. **IT infrastructure** for the institution that enables this policy to be implemented. At the most basic level, if there is nowhere to store data, and/or no easy way to collect it into that storage, then there is no point in having a management policy for it. Equally this infrastructure should be capable of connection to IT tools that will allow for sharing, reuse and reconfiguring. These tools may have been developed locally, through ANDS funded programs such as Data Capture or Applications, through other Government funded projects such as NeCTAR, or outside Australia.

3. **Co-ordination of the policy and IT infrastructure** to encourage and facilitate the creation and storage of data in a fashion that will enable sharing and reuse at an appropriate future time.

4. **Storage and collection mechanisms** that support the capture and creation of metadata about the data, so that the data can be shared and reused. Where possible this should be drawn
from existing sources (to avoid re-keying) and be as widely shared as possible. From an ANDS perspective, these metadata should be made available in the Australian Research Data Commons.

5. **Dedicated services** in place to support all of the above that draw from the relevant areas of expertise across the institution.

Institutions will implement research data support in a wide variety of ways that will best suit their needs. By seeking common outcomes from similar initiatives at institutions across the sector in Australia, ANDS is seeking to enable the Australian Research Data Commons to be a nationally strategic resource. ANDS will work with institutions in this way through all of its programs but with the engagement being focused through its Institutional Engagement program.

### 3.11.3 National Research Data Infrastructure Partner Overview

ANDS is part of the Australian Government’s investment in research data infrastructure. This investment can be seen very broadly as comprising all of the investments in research data-generating activity—indeed many research grants have a substantial data acquisition component. However ANDS has a particular responsibility and opportunity to engage in partnership with other major national data infrastructure investments. ANDS has to date particularly engaged with:

- Some of the major data generating instrument investments including the Australian Synchrotron, the Australian Nuclear Science and Technology Organisation’s neutron beam instruments, the Australian Telescope National Facility, as well as investments in a larger number of smaller data generating instruments through investments like BioPlatforms Australia, the Australian Microscopy and Microanalysis Research Facility, and the National Imaging Facility.

- The major discipline focused data investments, including the Terrestrial Eco-Research Network (TERN), the Integrated Marine Observation System (IMOS), the Atlas of Living Australia, AuScope, the Australian Urban Research Information Network (AURIN), the Australian Data Archive, the Australian Biosecurity Information Network, and the Population Health Research Network (PHRN).

- The major eResearch enabling data investments including NeCTAR, with investments in data tools and data collaboration, RDSI for data storage, and the high end computation facilities to enable data analysis and generation.

ANDS’ engagements to date, as well as planned engagements are occurring across all of the ANDS programs. ANDS continues to engage with a number of capabilities on data licencing issues through our work with AusGOAL and an ANDS sponsored working group. A number of institutionally-based projects in our Seeding the Commons and Data Capture projects have supported the richer capture and management of research data from NCRIS or Super Science funded instruments, including for example, flux tower data accessed using a Monash University metadata access tool into a Metadata Store as part of the TERN OzFlux program. Collections from many of the facilities are discoverable though Research Data Australia and collections can be persistently identified. TERN has launched its
own portal based on the ANDS registry tool to expose its own collections. Public Sector Data is funding work to make Geological Data more available to exploit using the AuScope portal and more discoverable using Research Data Australia. ANDS funded the automation of real-time publication of data from Research Vessel *Southern Surveyor* and Research Vessel *Aurora Australis*, which is available through the IMOS portal and with collections descriptions discoverable through Research Data Australia.

ANDS has partnered with a number of the National Data Investments to fund demonstrations of the value of new approaches to research data through its Applications program. ANDS has partnered with the Terrestrial Eco-Research Network, the Atlas of Living Australia, BioPlatforms Australia, the Integrated Marine Observation System, the National Imaging Facility, the Population Health Research Network and the Australian Urban Infrastructure Research Information Network to together demonstrate these advantages, in particular demonstrating the value of integrating data from different disciplines. Examples include the bringing together of satellite, genetic and field observations of Australian soil and the ability to visualise and combine biological data though many dimensions from whole-of-organism through microscope level to genetic level. A particularly powerful example is the demonstration of the value of examining very wide scale effects of climate change, response and adaptation though built environment, natural environment and health though sharing climate data downscaled to local conditions. This is achieved using National Computational Infrastructure computation, though a portal at Griffith University, utilising the network of researchers around Australia organised through the National Climate Change Adaptation Research Facility and with the data stored on an RDSI funded node.

Finally ANDS is focusing on nationally significant research collections though partnering with a range of collection institutions with a view to enabling Australian researchers to get better access to this data, often stored and made available through RDSI funding.

ANDS thus has engaged with a very wide range of activities in collaboration with other research data capabilities, but also participates in efforts to ensure that ANDS strong institutional focus complements other more specific investments.

The net effect of all of this activity is that the many data investments that Australia makes are enhanced by having data collectors, research institutions, and research infrastructure providers combine to give Australian researchers a significant research data advantage.
4 Research Infrastructure

For researchers to work in the world of data-intensive research, they will need:

- Policies that support a new way of working.
- A technical data fabric that enables storing and moving data.
- A metadata infrastructure to manage rich information about their data.
- A referencing mechanism that enables input data, modelling outputs (such as visualisations), software code and documents to be cross referenced.
- The ability to search across all the collections that have been registered.
- Training and training materials that enable the infrastructure to be used well.

To deliver that infrastructure there is a need for substantial data collections, and a combination of national services and coherent institutional services. ANDS has created many of those services, and is helping to seed institutional services that are optimised to be part of the Australian Research Data Commons, where reuse, sharing and commonality of approach is possible because of the coherence achievable with national investments. These services can also be integrated with other national services—whether they be data storage services, high intensity data analysis, or discipline or problem specific data services.

ANDS has instituted four programs to establish and maintain this infrastructure:

- National Collections.
- National Services.
- Institutional Engagement.
- International Collaboration.

These programs will be delivered in a co-ordinated manner that will ensure that ANDS’ partners engage with ANDS as a whole, not with the individual programs. This has the consequence that ANDS needs to have a greater emphasis on customer relationship management so that partners do not have to navigate their way through different parts of ANDS.
4.1 National Collections

The National Collections program represents ANDS’ focus on data collections that are contributed by national data providers such as the NCRIS data intensive facilities, the NCRIS eResearch facilities, and Commonwealth science agencies and data collecting and aggregating departments. This national focus complements the institutional focus of the institutional engagements program which caters for Australian research organisations, typically universities.

In undertaking this activity, the National Collections program is trying to move beyond a focus on data collections that are managed, connected and findable, to include also the critical fourth transformation: collections that are reusable. In this context, National Collections is also responsible for advanced collections functionality: custom presentation of particular collections, themes to aid with discovery and reuse, data reuse tools, and connections to available services over data. These features will be progressively embedded within Research Data Australia as part of our commitment to maintaining the ‘reusefulness’ of data.

In order to achieve these objectives, the program draws on two functions: national engagements and national projects, which together reflect the spectrum of interaction ANDS has with these national providers.

Through national engagements, ANDS maintains long term relationships with nationally-focused data providers, and, through national projects, delivers particular NCRIS 2013 funded projects with NCRIS eResearch services and NCRIS data intensive facilities. The resources of the entire program are pooled to provide a coherent experience for key stakeholders, sufficient resource scaling to provide value, effective sharing of corporate memory and specific skills.

4.1.1 Expected Highlights

Highlights of this program include:

- Partnerships with NCRIS data intensive facilities
- Collaborative arrangements with NCRIS eResearch facilities in support of the above
- Maintenance of ongoing relationships with key public sector data providers

4.1.2 Expected Difficulties

Coordinating business planning and resource allocation with other NCRIS facilities will be a challenge for this highly collaborative area of activity. Uncertainty of resources at many Commonwealth agencies and departments will result in unpredictable project and engagement timelines in these areas.

The timelines for planning and delivery of a set of engagements where e-Research infrastructure activity is aligned with the needs of data-intensive NCRIS capabilities are very tight, leaving little slack time in the schedule.
4.1.3 Expected Breakthroughs

Maintaining access to research data through coordination of storage (either institutional or on RDSI nodes), services (delivered either by NeCTAR or other data service providers, as well as by ANDS) and ANDS collection descriptions in support of NCRIS facilities will be a breakthrough that underlines the pivotal and cross domain nature of ANDS services within the NCRIS model.

4.1.4 2014-15 Expected Activities

The National Collections program will be engaged in three activity areas for 2014-5:

1. National Engagements.
3. Advanced Collections.

National Engagements

ANDS maintains existing data publication support arrangements for

- Data Intensive NCRIS facilities.
  - IMOS, TERN, BPA, ALA APPN APN, AuScope, AURIN, PHRN, NIF and MMRF, NCI
- Priority Commonwealth data collectors and aggregators.
  - GA, BOM, ABS, Environment, AIHW.

National Projects

These projects will maintain data access through both funding for, and strong ANDS engagement with, areas that have been identified in the 2011 Strategic Roadmap for Australian Research Infrastructure:

- Marine Environment.
- Terrestrial Systems.
- Solid Earth.
- Astronomy.
- Climate.
- Climate Adaptation.
- Integrated Biological Discovery.
- Characterisation.
- Urban Settlements.
- Population Health Research Platforms.

This will enable ANDS to continue its engagement that has already informed the thinking of the Research Data Infrastructure Committee, and will leverage existing engagements with the data intensive NCRIS facilities. These projects provide continued support for data collection publication, to enable services that are being generated through the NeCTAR Virtual Laboratories to be connected, provide long term access to data, reliable access to collections via rich descriptions, and connections to international collections through Research Data Australia.
These projects are an opportunity to extend ANDS collaboration with other eResearch infrastructure providers where they are also supporting these same capabilities. They will ensure connections between the collections held on RDSI nodes or other storage solutions, developing data services, such as ones being developed through the NeCTAR Virtual Laboratories or other providers, via descriptions made available through Research Data Australia. These engagements will enable ANDS, RDSI, and NeCTAR to fund Nodes to provide an integrated e-research infrastructure experience for data-intensive capabilities. As a result, there will be connectivity between the data and the services, and the data will be maintained and visible. The sum of $2,400,000 is allocated to this engagement with the specific detail of the way of optimising ANDS, RDSI, and NeCTAR investments to be determined jointly between the projects, the service providers, and the service users.

These projects will be informed and supported by the eResearch Coordinating project lead by Dr Rhys Francis.

**Advanced Collections**

This will include:

- Services over data associated with specific collections held at research institutions.
- A wider set of thematic views of RDA content.
- Signature collection presentation.
- Continued input into content strategy and business solutions for RDA (including information architecture review for a renewed user focus).

### 4.2 National Services

Through the National Services program, ANDS provides data publication infrastructure as well as support for those using ANDS infrastructure. The support extends more broadly to research organisations and public sector data providers building their own institutional capacity and capability as part of the Australian Research Data Commons.

#### 4.2.1 Expected Highlights

Continuity of provision of a stable and reliable national data publication platform is the aim of ANDS online services and national infrastructure. Orchestration of existing data with existing data services to facilitate data reuse is an expected highlight. With a steady increase in awareness of research data in the sector, a continued and increased demand for ANDS education and outreach services in 2014-5 is expected including more workshops, greater support for communities of practice, and more specialised support materials.

#### 4.2.2 Expected Difficulties

Coordinating and balancing ANDS infrastructure and capability building activity with that of our institutional and NCRIS facility partners requires ongoing attention.
4.2.3 Expected Breakthroughs

During 2013-14 ANDS expects to expand its support to include the Research Office/Research Management community more specifically.

Policy movement at major funding agencies will focus the attention of eligible organisations to view data infrastructure and capability as an enterprise asset.

The technical community of software developers who integrate systems with ARDC data publication services will grow into a self-supporting community.

4.2.4 2014-15 Expected Activities

The National Services Program has two major activity centres:

1. Infrastructure Maintenance and Online Service Provision.
2. Capability Building (Skills, Resources, Policy).

1. Infrastructure Maintenance and Online Service Provision

This activity is intended to:

- Maintain and optimise 111 existing national services.
- Maintain access to existing data collections.
- Maintain data connections (among ANDS, institutions, eResearch facilities, and data-intensive NCRIS facilities).
- Support ANDS project portfolio (such as MODC, Data-intensive NCRIS facilities, NCRIS eResearch Connectivity).
- Support ANDS strategic business directions (such as Data Value, Data Reuse, Data Citation, Data Connections, Data-publications-grants, International).

The key infrastructure activity areas therefore include:

- Showcase Collections (support for the needs of the MODC and NCRIS projects as well as National Collections).
  - custom presentations; custom harvests.
  - content highlighting mechanisms.
  - collections and services orchestration (see below).
- Services over Data (systems support for the Data Reuse Taskforce).
  - systems development to support trial.
  - schema change support.
  - Consultation.
  - support for NCRIS projects collection-service orchestration(see above).
- Linking Data with Publications and Grants (Research Link).
- DOI Service Optimisation.
  - DOI custody transfer.
  - management interface/ reporting.
Quality of Service (service uptime monitoring and notifications).

- Thomson Reuters DCI export.
  - business process optimisation; admin interfaces.
  - Filtering.
- ORCID.
  - technical and community implementation support.
  - integration optimisation.
- RDA optimisation (Strategic Directions for RDA).
  - Information Architecture Reconfiguration (Project Fallingwater).
- International (RD-A working group support).

2. Capability Building (Skills, Resources, Policy)

The Capability Building activity contributes to ANDS endurable outcomes by building capability in strategic areas with target stakeholders through activities such as virtual events, workshops, product development, awareness raising, publications and communications. These are design to ensure a sustainable future for data management and access.

The key activities for 2014-15 therefore include:

- Webinar Series.
  - Data management.
  - Data Citation.
  - Joint CAUL-ANDS series.
  - Data planning.
- Workshops.
  - MODC support workshop.
  - Data Access and Data Connections project support.
  - Research managers.
- Events: Open Data Collections Jamboree.
- Publications.
  - Web site materials optimisation and extension.
  - User stories gallery (reuse, collaboration, citation, profile, research impact, research value).

4.3 Institutional Engagement

The aim of the program is to ensure that institutional research data infrastructure is developed, maintained and operated at all of Australia’s major research organisations. This will ensure a continuing coherent Australian Research Data Commons, and that the research data assets of the institutions provide as much value as possible. The program will maintain activity funded in previous years under the Seeding the Commons, Data Capture, Metadata Stores and Applications programs. These programs built a fabric for data management with a view to increasing the scope of the data commons; improving data capture and management across the research sector; funding the
establishment of infrastructure in institutions for depositing data and metadata into well-managed stores; and funding the creation of a number of demonstrations of the value of bringing data together to answer new questions.

### 4.3.1 Expected Highlights

ANDS will engage with institutions through projects, with separate institutions and through community engagement activities. ANDS will meet with senior representatives of 26 institutions to establish their desire to establish a Major Open Data Collection (MODC) which is relevant to their institution and has significance at an institutional, national, and international level. This $5.2 million dollar engagement strategy is intended to ensure that Australian research institutional data assets are uncovered to attract international attention to 26 major international open data collections. These collections will exploit Australia’s research data infrastructure and Australian research institutional data capabilities to enable the best possible international profile and partnerships.

The Major Open Data Collections engagement will be formalised by a contract, outlining an agreed division of responsibilities and activities between the institution and ANDS, which will seek to deliver significant collections by June 2015. A progress report by June 2014 will be requested from the institution.

ANDS will also engage with the head of libraries at a further 14 smaller and regional universities to determine their desire to establish an Open Data Collection. This $490K dollar engagement strategy is aimed at supporting these universities and their libraries to contribute to the successful establishment of data collections, support the institution’s strategic outcomes and demonstrate its preparedness to embrace, in a timely way, the change which will progressively occur in the national policy framework for research data.

The Open Data Collections engagement will be enable the institution can employ the national infrastructure involve ANDS expertise in the process, and enhance the institution’s open data collections, ensuring a broad involvement in the national data infrastructure.

### 4.3.2 Expected Difficulties

Due to delays in the NCRIS 2013 funding announcements, ANDS was unable to formalise the MODC engagement with institutions until late in 2013. As a result, high level discussions with the relevant institutions occurred prior to the universities closing down at the end of the year. Therefore internal discussions at the institution regarding its significant collections could not occur until the end of January. This has resulted in contracting arrangements occurring later than anticipated.

By learning from its past experience, ANDS has attempted to minimise the time required for contract reviewing by the University Solicitor’s Office by using existing contract arrangement formed under past ANDS projects. However, as it cannot predict whether further contract reviews will be undertaken by individual institutions, delays in commencing projects at the institution may occur.
These delays may result in institutions not having as much time as expected to deliver the open data collections outlined in their MODC Project contracts.

It is important that ANDS remains connected with the interests of the institutions in order to offer meaningful and valuable activities and resources to the institutions. If this is not the case, institutions might well not be interested in engaging in discussions with ANDS and community activities. This becomes especially relevant if there is not a perspective of project funding for the institutions through ANDS.

The ODC funding is provided without the requirement that the data collection is identified up front, thereby offering the university a chance to explore options within their university for potential collections. Finding a suitable collection that can be made openly available is a valuable experience in the process for the institutions involved, but can take some time and ANDS will offer help in this process to identify potential collections.

4.3.3 Expected Breakthroughs

Through the Major Open Data Collection engagement selected Australian research institutions will showcase their collections to the world by June 2015. These collections may not have previously had the same level of access or visibility to researchers, as they may have been restricted to selected researchers and/or disciplines. Increasing visible to the data will not only promote the institutional brand, but will enable Australia to have its world class research data open, visible, and citable for the extended research community and partners around the world.

It is anticipated that a number of the MODC projects will have synergies with each other, other ANDS activities in National Collections Programs, and contribute to themes of research importance in Australia.

An additional benefit of the MODC engagement will be an increased awareness by institutions in the value of their institutional data assets and the ongoing need for sustainability and management of these assets at an institutional level. Economies of scale will arise by maintaining ongoing access to research data assets and reduce the duplication of important existing datasets.

In addition to promoting and assisting institutions in making their research data open, accessible, available, and citable through its Institutional Engagement Program, ANDS will continue to support the research community meet the challenge of ARC funding requirements by supporting institutions on how to address the management of research data.

Through the Open Data Collection engagement, Australian research institutions will showcase their collections to the world in June 2015. These are not required to be of the same international significance as the MODC projects as the provided amount of funding is significantly lower.

The ODC projects also offer the smaller and regional universities the opportunity to showcase their research and tie this in to national collections and open pathways to international collaboration. It prepares them for a national policy framework which supports data-intensive research and excellence in research data management. Through the method of delivery it will provide university
libraries with the opportunity and associated support to gain experience in managing, describing and sharing research data and specifically the challenges around making data openly available.

### 4.3.4 2014-15 Expected Activities

The continuation of a series of institutional engagements, through projects and activities, will enable new data focused partnerships in order to assist organisations involved achieve their research data ambitions. Engagements will be supported by a service offering that includes advice, tools and training, to be offered within a set of categories such as management, publication citation, reuse.

The Institutional Engagement Program will be responsible for the Major Open Data Collections (MODC) engagement. These projects will enable ANDS to partner with and provide $200,000 to each of 26 institutions that have installed data capture and/or institutional metadata stores.

The aim of this engagement is to establish long-term open access to one or more collections of strategic significance to the institution. Engagements will be formalised with the following institutions:

- Monash University
- La Trobe University
- University of Melbourne
- Deakin University
- RMIT University
- Swinburne University
- Flinders University
- University of Adelaide
- University of South Australia
- University of Newcastle
- University of NSW
- University of Sydney
- University of Technology
- Macquarie University
- University of Wollongong
- University of Western Sydney
- Australian National University
- CSIRO
- University of Tasmania
- James Cook University
- Queensland University of Technology
- Griffith University
- University of Queensland
- University of Western Australia
- Curtin University
- Regional University Network

The aim of the Major Open Data Collection (MODC) engagement is to deliver collections that are openly available and accessible which have the following characteristics:

- Substantial collections where significant data already exists.
- Has a demonstrated research need for the data beyond the scope of the institution.
- Data that is well described and well linked; having richly connected data collections and sub-collections.
- Are open and accessible, although allowing for the possible requirement for registration.
- Data collection is discoverable through appropriate means including Research Data Australia as well as institutional, international and discipline specific portals.
- The institution can demonstrate its enduring commitment to its data collection.
The Institutional Engagement program will also be responsible for the Open Data Collections (ODC) engagement providing $35,000 to each of the 14 institutions listed below. This funding opens the route to collaborate with the universities that prepared an institutional infrastructure using the Seeding the Commons funding.

- Australian Catholic University
- Bond University
- Central Queensland University
- Charles Darwin University
- Charles Sturt University
- Edith Cowan University
- Murdoch University
- Southern Cross University
- University of Canberra
- University of Southern Queensland
- Federation University
- University of the Sunshine Coast
- University of New England
- Victoria University

The aim of the Open Data Collection (ODC) engagement is to deliver collections that are openly available and accessible which have the following characteristics:

- Substantial collections where significant data already exists.
- Are open and accessible, although allowing for the possible requirement for registration.
- Data collection is discoverable through appropriate means including Research Data Australia as well as institutional, international and discipline specific portals.
- The institution can demonstrate its enduring commitment to its data collection.

The aim of Institutional and Community Engagements is to regularly engage with stakeholders at the central level within universities, research institutes, and with selected state government bodies on their progress on managing, sharing and re-using research data. ANDS will seek out opportunities in which to assist institutions in contributing to the Australian Data Commons and to address their challenges within the institution. The Institutional Engagement team will facilitate contact with experts within ANDS that provide specialist knowledge to institutions to address their questions regarding better data management, data sharing and reuse approaches. Through these engagement ANDS will be closely connected to the real life challenges that the universities, research institutes and state government bodies face, and adapt targets accordingly. Leading to more datasets being continuously openly available and findable through Research Data Australia. An example of such a request is supporting the institutions in responding to the Research Data requirement in ARC proposals.

By engaging with the research data community ANDS will develop a community of experts on research data across institutional boundaries that can share challenges, learn from each other and identify shared needs. This will provide a benefit of scale, as there are often only a few experts focused on managing research data within a single institution, yet a wealth of experience across Australia. Universities exploring research data can learn from how others have addressed a challenge. ANDS will facilitate these communities through regional face to face meetings, virtual exchange and capability building activities (see capabilities). The topics addressed will be based on the institutions’ needs. Leading experts from these communities can also be brought into international exchanges to ensure that Australia develops experts that can contribute to the international debate and ensure that the Australian voice is also heard (see International).
4.4 International Collaboration

The main vehicle for International Collaboration will be the processes and opportunities provided by the Research Data Alliance, although a number of existing activities will continue in parallel. The purpose of the Research Data Alliance is to accelerate international data-driven innovation and discovery by facilitating research data sharing and exchange, use and reuse, standards harmonization, and discoverability. This will be achieved through the development and adoption of infrastructure, policy, practice, standards, and other deliverables. The Research Data Alliance has been funded by The Australian Commonwealth Government through the Australian National Data Service, the European Commission through the RDA/Europe project funded under the 7th Framework Program, and the United States of America through the RDA/US activity funded by the National Science Foundation.

The work of the Research Data Alliance is primarily being undertaken through its working groups and interest groups. The development of solutions will occur through working groups. These are intended to come into existence, work on a problem and deliver a solution that is both adopted and that improves data exchange at the end of a 12-18 month period. Interest groups are for people with an interest in a particular data technology or research discipline. Interest groups may identify data interchange problems that need to be solved. Participation in working groups and interest groups, starting new working groups, and attendance at the twice-yearly plenary meetings is open to all.

Australian representation on the RDA is currently through:

- Dr Ross Wilkinson (ANDS) as one of the members of Council.
- Dr Andrew Treloar (ANDS) as one of the co-chairs of the Technical Advisory Board.
- Dr Stefanie Kethers (ANDS) as one of the Secretariat.
- Dr Simon Cox (CSIRO) as one of the members of the Technical Advisory Board.

4.4.1 Expected Highlights

Expected highlights over the planning period will be:

- Continued momentum of the RDA.
- Successful holding of Plenary 4 (Amsterdam, September 2014) and Plenary 5 (Venue TBD, March 2015).
- Approval and commencement of the RDA Research Data Registry Interoperability Working Group, lead by staff from ANDS.
- Greater involvement of Australian participants in RDA Interest Groups and Working Groups.
- Successful completion of the ODIN project, funded by the EU, in which ANDS is an active participant.
- Commissioning of data flows from Research Data Australia into the Thompson-Reuters Data Citation Index.
- Ongoing expansion of the DataCite organization, of which ANDS is a founding member.
4.4.2 Expected Difficulties

The major difficulty anticipated is the challenge of greater Australian involvement in the RDA when the bulk of the members come from Western Europe and the USA. This manifests itself in distances to be travelled to physical meetings, and the time zones chosen for virtual meetings. This should improve over time as the membership of RDA becomes more geographically diverse.

4.4.3 Expected Breakthroughs

No specific breakthroughs are anticipated during the planning period, which will be characterised by steady positive progress.

4.4.4 2014-15 Expected Activities

Expected activities will include:

- Actively seeking and working with Australian representatives of discipline communities for whom an RDA working group is an appropriate vehicle to develop and implement international data exchange technologies.
- The Research Data Registry Interoperability Working Group working towards a successful conclusion.
- Ongoing contribution to RDA Governance bodies as outlined above, including contributions to the planning committees for the RDA Fourth Plenary.
- Contributing to the ongoing success of Datacite, including attending DataCite Board Meetings and Annual General Meetings.
- Involvement in a follow-on project to ODIN (if their grant application is successful)
- Ongoing engagement with ORCID (author identifiers) and Thompson Reuters (Data Citation Index).

4.5 Overall 2014-15 Expected Outcomes

- Ensure collections that are formed, whether they are by NCRIS capability, public sector or research institutional partners are published appropriately through integrated services.
- Maintain access to institutionally and nationally significant collections as openly as possible.
- Continue to deliver valued and reliable national data technical and advisory services.
- Continue to populate the ARDC with more data collections that have been managed and connected, ensuring over 90,000 collections descriptions discoverable through Research Data Australia, Google and other mechanisms.
- Continue to support research institutional capability and capacity to manage research data
- Maintain ANDS as a trusted partner of research institutions.
- Maintain Australia’s leading role in research data infrastructure internationally.

5 Confidential Information

There is no confidential information.
6 Access and Pricing

The mechanisms for deciding access and pricing will be consistent across the ANDS services. Generally speaking, ANDS will provide services for research purposes and aims to ensure the legitimate research use of those services will be free and access to the services will be open.

Software developed under the programs will be released as Open Source code, with the choice of licence and licensing conditions varying on a case-by-case basis. Documents produced or funded by ANDS will be made available as public documents, on a no warranty, royalty free basis using the CC-BY license. ANDS will maintain a register of software that is produced through its funded projects.

However, content access and charging regimes belong in the hands of content providers, so that the access and pricing issue in ANDS relates to the rules under which content may be provided into the ARDC and therefore supported by ANDS utilities and other support activities.

ANDS services will be restricted to users who are non-commercial, and engaged in research, with the exception of Research Data Australia, which will be searchable by any interested party. Research Data Australia will not be used for the advertisement of paid services. Equally, the Identify my Data service will only be accessible for non-commercial use.

7 Summary Financial Information

ANDS works in two main ways with a range of partners: either by way of teams drawn from within ANDS and the partner organisation, or external funding of activities to achieve ANDS’ objectives. In some cases there may be a combination of such work – for instance training at an institution might be provided by ANDS staff to assist the work of the project staff employed at that institution. In a typical partner engagement ANDS would either fund a project or activity at an institution or place a fully-funded ANDS staff member within an institution to achieve the institution’s and ANDS’ aims.

Table 2 below provides a view of each program’s budget and breakdown of costs by internal costs (i.e., staff costs and operating expenses) and external expenditure.
7.1 Expenditure by Program

Table 2: Expenditure by Program

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Actuals Jan 2013 - Dec 2013 $'000</td>
<td>Forecast Jan 2014 - Jun 2014 $'000</td>
<td>$'000</td>
<td>$'000</td>
</tr>
<tr>
<td>Opening Cash Balance</td>
<td>-</td>
<td>4,875.78</td>
<td>4,973.08</td>
<td>-</td>
</tr>
<tr>
<td>NCeRS 2013 Funding</td>
<td>2,871.00</td>
<td>2,871.00</td>
<td>7,308.00</td>
<td>13,050.00</td>
</tr>
<tr>
<td>CRIS Funding</td>
<td>2,000.00</td>
<td>1,000.00</td>
<td>-</td>
<td>3,000.00</td>
</tr>
<tr>
<td>Interest Earned</td>
<td>26.65</td>
<td>69.40</td>
<td>72.48</td>
<td>168.52</td>
</tr>
<tr>
<td>Total Income</td>
<td>4,897.65</td>
<td>8,816.18</td>
<td>12,353.55</td>
<td>16,218.52</td>
</tr>
<tr>
<td>Total Expenditure</td>
<td>21.87</td>
<td>3,943.10</td>
<td>12,353.55</td>
<td>16,218.52</td>
</tr>
<tr>
<td>Closing Cash Balance</td>
<td>4,875.78</td>
<td>4,973.08</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Expenditure</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>National Collections</td>
<td>0.13</td>
<td>585.61</td>
<td>1,697.55</td>
<td>2,283.20</td>
</tr>
<tr>
<td>Management and Governance Costs</td>
<td>0.13</td>
<td>15.90</td>
<td>147.60</td>
<td>163.53</td>
</tr>
<tr>
<td>Salaries and On-Costs for technical staff</td>
<td>9.71</td>
<td>605.42</td>
<td>615.12</td>
<td></td>
</tr>
<tr>
<td>Infrastructure Maintenance</td>
<td>549.94</td>
<td>919.93</td>
<td>1,469.87</td>
<td></td>
</tr>
<tr>
<td>Total Administration</td>
<td>10.06</td>
<td>24.60</td>
<td>34.66</td>
<td></td>
</tr>
<tr>
<td>Utilities</td>
<td>2.01</td>
<td>4.92</td>
<td>6.93</td>
<td></td>
</tr>
<tr>
<td>Rent</td>
<td>6.64</td>
<td>16.23</td>
<td>22.88</td>
<td></td>
</tr>
<tr>
<td>Consumables</td>
<td>1.41</td>
<td>3.44</td>
<td>4.85</td>
<td></td>
</tr>
<tr>
<td>International Collaboration</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Others</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>National Services</td>
<td>0.28</td>
<td>436.01</td>
<td>3,161.09</td>
<td>3,597.38</td>
</tr>
<tr>
<td>Management and Governance Costs</td>
<td>0.28</td>
<td>26.76</td>
<td>286.74</td>
<td>313.78</td>
</tr>
<tr>
<td>Salaries and On-Costs for technical staff</td>
<td>49.26</td>
<td>1,977.30</td>
<td>2,026.56</td>
<td></td>
</tr>
<tr>
<td>Infrastructure Maintenance</td>
<td>346.67</td>
<td>859.92</td>
<td>1,206.59</td>
<td></td>
</tr>
<tr>
<td>Total Administration</td>
<td>13.33</td>
<td>37.12</td>
<td>50.45</td>
<td></td>
</tr>
<tr>
<td>Utilities</td>
<td>2.67</td>
<td>7.42</td>
<td>10.09</td>
<td></td>
</tr>
<tr>
<td>Rent</td>
<td>8.80</td>
<td>24.50</td>
<td>33.30</td>
<td></td>
</tr>
<tr>
<td>Consumables</td>
<td>1.87</td>
<td>5.20</td>
<td>7.06</td>
<td></td>
</tr>
<tr>
<td>International Collaboration</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Others</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Institutional Engagement</td>
<td>21.35</td>
<td>2,544.41</td>
<td>6,779.82</td>
<td>9,345.59</td>
</tr>
<tr>
<td>Management and Governance Costs</td>
<td>0.99</td>
<td>82.79</td>
<td>466.36</td>
<td>550.14</td>
</tr>
<tr>
<td>Salaries and On-Costs for technical staff</td>
<td>0.04</td>
<td>92.21</td>
<td>1,779.97</td>
<td>1,872.22</td>
</tr>
<tr>
<td>Infrastructure Maintenance</td>
<td>20.33</td>
<td>2,325.57</td>
<td>4,436.09</td>
<td>6,781.98</td>
</tr>
<tr>
<td>Total Administration</td>
<td>43.85</td>
<td>97.40</td>
<td>141.26</td>
<td></td>
</tr>
<tr>
<td>Utilities</td>
<td>8.77</td>
<td>19.48</td>
<td>28.25</td>
<td></td>
</tr>
<tr>
<td>Rent</td>
<td>28.94</td>
<td>64.29</td>
<td>93.23</td>
<td></td>
</tr>
<tr>
<td>Consumables</td>
<td>6.14</td>
<td>13.64</td>
<td>19.78</td>
<td></td>
</tr>
<tr>
<td>International Collaboration</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Others</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>International Collaboration</td>
<td>0.10</td>
<td>277.06</td>
<td>715.10</td>
<td>992.26</td>
</tr>
<tr>
<td>Management and Governance Costs</td>
<td>0.10</td>
<td>3.89</td>
<td>58.68</td>
<td>62.58</td>
</tr>
<tr>
<td>Salaries and On-Costs for technical staff</td>
<td>-</td>
<td>188.60</td>
<td>188.60</td>
<td></td>
</tr>
<tr>
<td>Infrastructure Maintenance</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Total Administration</td>
<td>4.86</td>
<td>11.33</td>
<td>16.18</td>
<td></td>
</tr>
<tr>
<td>Utilities</td>
<td>0.97</td>
<td>2.27</td>
<td>3.24</td>
<td></td>
</tr>
<tr>
<td>Rent</td>
<td>3.20</td>
<td>7.48</td>
<td>10.68</td>
<td></td>
</tr>
<tr>
<td>Consumables</td>
<td>0.68</td>
<td>1.59</td>
<td>2.27</td>
<td></td>
</tr>
<tr>
<td>International Collaboration</td>
<td>268.32</td>
<td>456.48</td>
<td>724.80</td>
<td></td>
</tr>
<tr>
<td>Others</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Total ANDS Expenditure</td>
<td>21.87</td>
<td>3,943.10</td>
<td>12,353.55</td>
<td>16,218.52</td>
</tr>
</tbody>
</table>

Notes:
1. An additional $3.0M was received at the ANDS establishment phase. The balance of funds after expenditure during that phase was rolled over at the beginning of 2009 ($0.679M).
We now provide a programmatic view of our revenue and expenditure of our CRIS funds, NCRIS 2013 funds and Other Funding (Data Web Forum & 2nd Joint RI Workshop).

Table 3: ANDS expected budget allocation of CRIS funds

<table>
<thead>
<tr>
<th></th>
<th>2014-2 $'000</th>
<th>2015-1 $'000</th>
<th>Total $'000</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Collections</td>
<td>450.30</td>
<td></td>
<td>450.30</td>
</tr>
<tr>
<td>National Services</td>
<td>1,169.08</td>
<td></td>
<td>1,169.08</td>
</tr>
<tr>
<td>Institutional Engagement</td>
<td>1,338.75</td>
<td></td>
<td>1,338.75</td>
</tr>
<tr>
<td>International Collaborations</td>
<td>107.91</td>
<td></td>
<td>107.91</td>
</tr>
<tr>
<td>Total Budgeted Expenditure</td>
<td>3,066.03</td>
<td>-</td>
<td>3,066.03</td>
</tr>
<tr>
<td>Interest Income</td>
<td>17.91</td>
<td></td>
<td>66.03</td>
</tr>
</tbody>
</table>

Table 4: ANDS expected budget allocation of NCRIS 2013 funds

<table>
<thead>
<tr>
<th></th>
<th>2014-2 $'000</th>
<th>2015-1 $'000</th>
<th>Total $'000</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Collections</td>
<td>589.64</td>
<td>647.55</td>
<td>1,822.93</td>
</tr>
<tr>
<td>National Services</td>
<td>460.39</td>
<td>1,518.29</td>
<td>2,414.98</td>
</tr>
<tr>
<td>Institutional Engagement</td>
<td>2,900.52</td>
<td>2,496.72</td>
<td>7,963.01</td>
</tr>
<tr>
<td>International Collaborations</td>
<td>277.06</td>
<td>325.28</td>
<td>879.51</td>
</tr>
<tr>
<td>Total Budgeted Expenditure</td>
<td>4,227.61</td>
<td>4,987.84</td>
<td>13,080.42</td>
</tr>
<tr>
<td>Interest Income</td>
<td>40.90</td>
<td>13.67</td>
<td>102.49</td>
</tr>
</tbody>
</table>

Notes:
1. Budgeted Expenditure for 2013-2 and 2014-1 totals $3,864.97M

Table 5: ANDS expected budget allocation of Other (DWF, 2nd Joint RI Workshop) funding

<table>
<thead>
<tr>
<th></th>
<th>2013-2 $'000</th>
<th>2014-1 $'000</th>
<th>2014-2 $'000</th>
<th>2015-1 $'000</th>
<th>Total $'000</th>
</tr>
</thead>
<tbody>
<tr>
<td>International Collaborations</td>
<td>193.82</td>
<td>138.71</td>
<td>25.00</td>
<td></td>
<td>357.53</td>
</tr>
<tr>
<td>Total Budgeted Expenditure</td>
<td>193.82</td>
<td>138.71</td>
<td>25.00</td>
<td>-</td>
<td>357.53</td>
</tr>
<tr>
<td>Surplus/Contingency</td>
<td></td>
<td></td>
<td></td>
<td>79.92</td>
<td>79.92</td>
</tr>
<tr>
<td>Interest Income</td>
<td>11.53</td>
<td>2.17</td>
<td>1.31</td>
<td>1.19</td>
<td>16.20</td>
</tr>
</tbody>
</table>

Based on the cash flow for FY 2014-15, and assuming current interest rates, the estimated total interest income over FY 2014-15 should be c. $168k.
7.2 Expenditure of funds over the life of ANDS

In order to provide a context for this expenditure, the following table shows the expenditure by program over the life of ANDS, by half year, with a corresponding chart indicating the uneven nature of that expenditure. This is generated by considering expenditure from previous programs as part of the expenditure of their successor programs.

Table 6: Expenditure by program over the life of ANDS

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>National Collections</td>
<td>0.04</td>
<td>2.52</td>
<td>0.93</td>
<td>0.91</td>
<td>0.47</td>
<td>1.01</td>
</tr>
<tr>
<td>National Services</td>
<td>3.95</td>
<td>1.95</td>
<td>1.04</td>
<td>1.47</td>
<td>1.40</td>
<td>1.52</td>
</tr>
<tr>
<td>Institutional Engagement</td>
<td>2.29</td>
<td>9.52</td>
<td>5.15</td>
<td>6.34</td>
<td>6.35</td>
<td>6.43</td>
</tr>
<tr>
<td>International Collaboration</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.00</td>
<td>0.19</td>
<td>0.06</td>
</tr>
<tr>
<td>Governance &amp; Management</td>
<td>0.91</td>
<td>1.37</td>
<td>0.76</td>
<td>0.60</td>
<td>0.86</td>
<td>0.52</td>
</tr>
<tr>
<td>Total Planned Expenditure</td>
<td>5.19</td>
<td>15.35</td>
<td>7.87</td>
<td>9.32</td>
<td>9.08</td>
<td>9.61</td>
</tr>
<tr>
<td>Internal Total</td>
<td>3.42</td>
<td>5.74</td>
<td>3.06</td>
<td>3.41</td>
<td>3.70</td>
<td>3.61</td>
</tr>
<tr>
<td>External Total</td>
<td>1.77</td>
<td>9.61</td>
<td>4.81</td>
<td>5.91</td>
<td>5.38</td>
<td>5.99</td>
</tr>
<tr>
<td>TOTAL</td>
<td>5.19</td>
<td>15.35</td>
<td>7.87</td>
<td>9.32</td>
<td>9.08</td>
<td>9.61</td>
</tr>
</tbody>
</table>

The peak in expenditure in period 2012-13 was attributable mainly to:

- Applications and Metadata Stores programs.
- Delay in payments in Data Capture.
7.3 Co-investment

ANDS engagement with institutions through contracted work has not required that institutions commit specific cash or in-kind investment. However the NCRIS-funded Seeding the Commons Program and the EIF-funded Data Capture and Metadata Stores Program were intended to enable institutions to accelerate their research data infrastructure investments by developing specific solutions and re-using solutions developed at other institutions. ANDS has observed that institutions that have completed ANDS funded infrastructure have continued to invest in these from internal resources. In the Major Open Data Collection discussions to date there has been significant co-investment planned but not yet agreed. Thus, although there is no requirement for investment, there is a pattern of institutional investment that follows national investment. The advantage of this approach is that there is greater coherence in national institutional infrastructure as a result of initial national funding. This further institutional investment will be tracked by ANDS and detail provided to the Department upon request.

7.4 Expenditure of funds by type

External expenditure is intended to cover staffing administration and other costs related to the projects that ANDS has agreed to fund. This is difficult for ANDS to quantify accurately. The types of staffing positions that will be funded include the following:

- events/community coordinators
- policy specialists
- technical analysts
- business analysts
- project managers
- software developers
- business analysts
- project managers
- software developers

8 Governance

The Governance and Management arrangements for ANDS are described in the contracts for the NCRIS project and the EIF project, as well as in a separate Collaboration Agreement. These arrangements have been deliberately designed to ensure that the governance is as open as possible, consistent with the acceptance and management of risk by the lead agency. The Governance arrangements were established for the NCRIS contract, but the Department, Monash University, the lead agent and the Steering Committee have agreed to use the same approach to governance and management for the EIF ARDC contract and other contracts as well.

Monash University entered into an agreement with the Department to implement the Projects, receive NCRIS and EIF Funds and be accountable to the Department for execution and performance of both Projects. Monash University has established a Collaboration agreement with the Australian National University and CSIRO as partners in the projects.

Monash hosts and operates one of the ANDS Offices, which is used to manage the Project. ANU hosts the other office that houses both ANU and CSIRO staff.
Monash appointed the independent Chair of the Steering Committee after consultation with the Department and the ANDS partners and formally includes the independent Chair in the performance management arrangements of the Executive Director of ANDS. The Executive Director of ANDS is Dr Ross Wilkinson.

8.1 Steering Committee

The current ANDS Steering Committee comprises a minimum of four (4) and a maximum of eight (8) voting members, including:

(a) An independent chair appointed by Monash;

(b) One representative appointed by each of the ANDS Members; and

(c) Such additional persons as the ANDS Steering Committee may agree, such as data provider, data policy and other specialist representatives.

The Department has nominated a non-voting observer, Cheryl Kut or her nominee.

The processes of the ANDS Steering Committee will be as transparent as possible.

As at March 2014 the current ANDS Steering Committee Members are:

- Independent Chair: Dr Ron Sandland;
- Ms Cathrine Harboe-Ree (Monash University);
- Mr David Toll (CSIRO);
- Ms Roxanne Missingham (The Australian National University);
- Prof Mark Ragan (University of Queensland);
- Mr Paul Sherlock (University of South Australia);
- Dr Siu Ming Tam (Australian Bureau of Statistics);
- Prof Craig Johnson (University of Tasmania); and
- Executive Director (ex-officio): Dr Ross Wilkinson (Australian National Data Service).

It is anticipated that the ANDS Steering Committee membership can be expanded over time to incorporate any additional requirements of the Project.

8.2 Management structure

ANDS is currently managed by a full time executive staff comprising an Executive Director (located at Monash), and two Directors (a Monash Director, and an ANU Director) as currently agreed under the ANDS Collaboration Agreement, as well as Program Managers.

Directors and Managers report to the Executive Director with regard to ANDS activities and to a nominated person in the host institution for administrative purposes (the Supervisor). The Supervisor is normally the host institution’s representative on the Steering Committee.
Directors normally have a high degree of autonomy within their areas of responsibility but work under the leadership of the Executive Director.

If there is disagreement or conflict between the Executive Director and a Director the matter is discussed with the Supervisor in the first instance, after which it can be escalated to the Chair of the Steering Committee and, if necessary, the Steering Committee.

Any alterations to this arrangement will be as a result of, and documented in, a revised Collaboration Agreement that takes account of this Project.

ANDS staff work collaboratively with each other and support activities across ANDS. Some will be located at ANDS Member institutions and others out ‘in the field’. These field locations may include state based eResearch organisations, a Division of CSIRO or major data federating institutions.

ANDS staff within or appointed by an ANDS Member institution report to the relevant Director, or as otherwise negotiated for staff located in other institutions. These staff are appointed in consultation with the Executive Director.

If necessary, the Executive Director can direct, through the Directors and Managers, or other supervisory arrangements applicable at other institutions, the work of ANDS staff located in any institution.

The ANDS central office at Monash provides administrative support to ANDS and its staff, including communications, branding, and website maintenance.
9 Risk Management

ANDS maintains a Risk Register. The risk assessment methodology, adapted from the Australian Risk Management Standard AS/NZS 4360:2004, 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.

9.1 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 datasets 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 Departmental 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

### 9.2 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.
- 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.

**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.
- 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.

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 2013 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 Research Infrastructure 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 eResearch 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.

9.3 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 ‘reusers’ 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 reuse 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).
9.4 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.

10 Key Performance Indicators

Since an initial condition of the NCRIS Funding Agreement for ANDS is that “Key Performance Indicators (KPIs) acceptable to DIISRTE must be developed” ANDS has been using a consistent set of KPIs.

10.1 Key Performance Indicator Series

1. The number and coverage of data repositories providing metadata feeds to the national registry compared to the number of data repositories.
2. The number and coverage of institutions and number of research groups with which ANDS has engaged.
3. The number of institutions with research data management policies and practices consistent with ANDS recommendations.
4. The number of times a search is initiated with an ANDS discovery service.
5. The number of times an ANDS data page (defined below) is accessed.
6. The satisfaction of researchers and partners (see below) with ANDS services as measured by an annual survey.
7. The number of data access and sharing agreements with stakeholders – principally research institutions, government data agencies, government research agencies.

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

8. The number of research data sets in harvestable repositories.
9. The number of research data sets with persistent identifiers.

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

These KPIs address ANDS objectives (refer section 2) as follows:
- The commons: KPIs 1, 2, 4, 5, 7, 8, 9 and the long-term measure 10 address objective A.
- **Data management:** KPIs 3, 6 and ANDS’ long-term measure address objectives B and D.

- **Repositories:** KPIs 3, 8 and 9 address objective C.

- **Access:** KPIs 4, 5, 6, and 7 address objective E.

- **Use:** KPIs 4, 5, 6, 7 and the long term aspirational measure 10 address objectives F and G. (Note – when KPIs 4 and 5 are being measured, not only use will be noted, but where it is initiated so that analysis can be done both within and across disciplinary use. The satisfaction survey will be qualitative, enabling an understanding of how well disciplinary, cross-disciplinary and multinational interaction is being facilitated.)

The form in which ANDS services are offered will be shaped by adherence to the guidance provided above. This guidance will be reflected in the business plans, and adherence to this guidance will be determined in discussion with stakeholders.

**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, etc., 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 will again set benchmarks, but also help inform future surveys.
## 10.2 Estimates against Key Performance Indicators for 2014-15

<table>
<thead>
<tr>
<th>Performance Indicator</th>
<th>Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The number and coverage of data repositories providing metadata feeds to the national registry compared to the number of data repositories.</td>
<td>1. ANDS intends to maintain at least 100 metadata feeds. This will cover at least 60 research data-holding institutions.</td>
</tr>
<tr>
<td>2. The number and coverage of institutions and number of research groups with which ANDS has engaged.</td>
<td>2. ANDS will continue to engage with all Australian universities, PFROs, and 20 major Government data providers this year, and through them at least 50 research groups.</td>
</tr>
<tr>
<td>3. The number of institutions with research data management policies and practices consistent with ANDS recommendations.</td>
<td>3. The target is 31</td>
</tr>
<tr>
<td>4. The number of times a search is initiated with an ANDS discovery service.</td>
<td>4. The target is 50,000</td>
</tr>
<tr>
<td>5. The number of times an ANDS data page is accessed.</td>
<td>5. The target is 120,000 (now counting filtered page views using Google Analytics standard excluding robots spiders etc).</td>
</tr>
<tr>
<td>6. The satisfaction of researchers and partners (see below) with ANDS services as measured by an annual survey.</td>
<td>6. No number can be given here, but a report will be provided.</td>
</tr>
<tr>
<td>7. The number of data access and sharing agreements with stakeholders – principally research institutions, government data agencies, government research agencies.</td>
<td>7. ANDS will aim to maintain at least 50 agreements to make information available.</td>
</tr>
<tr>
<td>8. The number of research data sets in the ARDC.</td>
<td>8. Access to more than 80,000 collections will be maintained.</td>
</tr>
<tr>
<td>9. The number of research data sets with persistent identifiers.</td>
<td>9. The target is 14,000</td>
</tr>
<tr>
<td>10. The number of times a data set is reused and referenced.</td>
<td>10. This will be tracked but cannot yet be reported.</td>
</tr>
</tbody>
</table>