



Australian National Data Service (ANDS)

National Collaborative Research Infrastructure Strategy 2013 (NCRIS 2013)

Progress Report 1

For the period 1 July 2013 to 31 December 2013

31st January 2014

Contents

1	Introduction	3
2	National Collections Program	4
	2.1 Highlights, Difficulties and Breakthroughs	5
	2.2 Activities Undertaken	5
	2.3 Performance against agreed Key Performance Indicators	5
3	National Services Program	6
	3.1 Highlights, Difficulties and Breakthroughs	6
	3.2 Activities Undertaken	6
	3.3 Performance against agreed Key Performance Indicators	6
4	Institutional Engagement Program	7
	4.1 Highlights, Difficulties and Breakthroughs	8
	4.2 Activities Undertaken	9
	4.3 Performance against agreed Key Performance Indicators	9
5	International Collaboration Program	10
	5.1 Highlights, Difficulties and Breakthroughs	10
	5.2 Activities Undertaken	10
	5.3 Performance against agreed Key Performance Indicators	10
6	A summary table of actual expenditure for the period covered by this report	11

1 Introduction

This progress report outlines the current progress against the intended actions as per the submitted NCRIS 2013 Project Plan.

This project plan describes the intended actions to be carried out in the period September 2013 until June 2015 using NCRIS 2013 funding. The project plan is derived from an overall two year plan that describes the plan for ANDS activity over that whole time using NCRIS, EIF, DWF, CRIS and NCRIS 2013 funds. The programs, or components, are the same as have been agreed for the CRIS funding.

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; and
- **International Collaboration**, ensuring that the research data collaboration internationally is maintained, principally through the established international partnerships, most 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, however there will be 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 that are of enduring significance to the institutions and that support their research strategy and research partnerships strategy. The result of these engagements would be that at least 26 internationally significant open data collections are supported by the institutions and 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 already have a strong ANDS partnership, to 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.

- The Maintaining Connections engagements would ensure that connections between institutional metadata stores and RDSI node data registries enable institutions to have a comprehensive view of their research data assets, as well as ensuring connections are made 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.

2 National Collections Program

The scope of this program includes collections of government, institutional, discipline and national facility data.

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

The National Collections program will be responsible for the Access to Data for Research Infrastructure Capability Areas engagement, which will enable ANDS to partner and fund the six capability area that are data intensive, and already have a strong ANDS partnership, to 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.

The Access to Data for Research Infrastructure Capability Areas engagements will provide \$200,000 in funds and strong ANDS engagement with each of the following areas that have been identified in the *2011 Strategic Roadmap for Australian Research Infrastructure*:

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

This will enable ANDS to continue its engagement that has informed the thinking of the Research Data Infrastructure Committee, and enable further engagement with IMOS, TERN, ALA, AuScope, AURIN, AA, PHRN, BPA, APPN, NIF, and AMMRF.

2.1 Highlights, Difficulties and Breakthroughs

There is nothing to report as the money has not been expended.

2.2 Activities Undertaken

There is nothing to report as the money has not been expended.

2.3 Performance against agreed Key Performance Indicators

There has been no progress against Key Performance Indicators as the money has not been expended.

3 National Services Program

The scope of this program includes collection, registration, publication, discovery, and advisory services.

In its first four years of operation ANDS has established a national registry of research data, a national discovery portal, and a data publication identifier service. These are nationally focused, value-adding services which are more efficient to run as a single instance for the sector. For example ANDS has joined the global DataCite infrastructure consortium and provides data Digital Object Identifiers to the whole Australian research sector. Through the EIF and NCRIS infrastructure, many research organisations have established data archives and metadata provider infrastructure to feed the ANDS national registry, which aims to provide a national cross-discipline view of the data outputs from (and inputs to) Australian research.

The NCRIS project has enabled ANDS to establish a capacity to support, advise and exchange information about data infrastructure establishment at Australian research institutions. As well it provides targeted assistance for those using ANDS national services. ANDS has also established an expertise and advice service around questions of national policy promoting reuse of research data.

The National Service program will be responsible for the Maintaining Connections engagements. It will ensure that connections between institutional metadata stores and RDSI node data registries enable institutions to have a comprehensive view of their research data assets. In addition it will ensure 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 engagements will enable ANDS, RDSI, and NeCTAR to fund Nodes to ensure that there is an easy way to use NeCTAR funded services over data described in Research Data Australia and held on an RDSI node. There will be connectivity between the data and the services, and the data will be maintained and visible. \$1,000,000 is allocated to this engagement with the 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.

3.1 Highlights, Difficulties and Breakthroughs

There is nothing to report as the money has not been expended.

3.2 Activities Undertaken

There is nothing to report as the money has not been expended.

3.3 Performance against agreed Key Performance Indicators

There has been no progress against Key Performance Indicators as the money has not been expended.

4 Institutional Engagement Program

The program will ensure that institutional research data infrastructure is maintained and operated at all of Australia's major research organisations to ensure an continuing coherent Australian Research Data Commons, and that the research data assets of the institutions are made as valuable as possible.

This 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; improved data capture and management across the research sector; funded the establishment of infrastructure in institutions for depositing data and metadata into well-managed stores; and funded the creation of a number of demonstrations of the value of bringing data together to answer new questions.

The final year of NCRIS and EIF funding has enabled new data focused partnerships that will assist the organisations involved to achieve their research data ambitions through the continuation of a series of targeted engagements. These engagements are supported by a service offering that includes advice, tools and training and can be offered within a set of categories such as management, publication citation, re-use.

Under NCRIS 2013, the Institutional Engagement Program will be responsible for the Major Open Data Collections engagement. This will enable ANDS to partner and provide \$200,000 to the 26 institutions that have installed data capture and/or institutional metadata stores to establish long-term open access to one or more collections of strategic significance to the institution. The institutions are:

- 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

An additional Major Open Data Collection is intended to be established in partnership with the Regional University Network to enable participation and deliver partnership benefits.

The collections are expected to have the following characteristics:

- Substantial collections of significant data already exist.
- There is a demonstrated research need for the data beyond the scope of the institution
- The data should be well described and well linked having richly connected data collections and sub-collections.
- Open access to the data will be provided, although allowing for the possible requirement for registration.
- The data collection must be 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 this data collection.

ANDS will engage with senior representatives of each institution to establish their desire to establish a Major Open Research Data Collection. The institution may choose to establish more than one such collection, but the ANDS financial support is limited to \$200,000. A project charter will be formed with an agreed division of responsibilities and activities between the institution and ANDS, and this will be the basis of a contract with that institution.

This \$5.2 million dollar engagement strategy is intended to ensure that Australian research institutional data assets are operated and exploited to attract international attention to twenty six 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.

4.1 Highlights, Difficulties and Breakthroughs

The focus of the first reporting period has been on finalising the project plan and commencing conversations with the institutions ANDS is planning to work with. There have been some delays in commencing project plans and contracting. This has occurred due to stakeholders being on leave due to the Christmas shut down and holiday period.

All of the institutions ANDS has met with so far have been supportive of the proposed engagements. However the majority have needed to undertake further internal discussions to determine which of their collections would have the greatest impact. Furthermore they have also had to ascertain as to

whether or not they will be able to deliver collections with the characteristics required for the *Major Open Data Collections* engagement. All the representatives from the institutional meetings held have recognised and supported the need for open data and the value of their institutional data assets.

In December 2013 the drafting of a project plan, which forms the basis of the contract with ANDS, commenced with two institutions; Monash University and James Cook University. The draft project plans will commence in January 2014 with four institutions; Griffith University, Queensland University of Technology, University of Western Sydney, and CSIRO.

4.2 Activities Undertaken

During the reporting period the Executive Director and the Partnerships Programs Manager (PPM) for Institutional Engagement undertook a series of high level meetings and subsequent follow up discussions with 22 of the 26 selected organisations. Due to time constraints and changes within some of the institutions meetings with four of the selected 26 institutions were not held. These meetings will take place in the early part of 2014.

4.3 Performance against agreed Key Performance Indicators

There has been no progress against Key Performance Indicators as yet.

5 International Collaboration Program

The scope of this program includes collaboration with data infrastructure providers to ensure that Australian research data infrastructure is compatible with international approaches.

Until 2012, international collaboration activities within ANDS occurred across a number of different programs, including for example, selection of DataCite as our preferred Digital Object Identifier solution, and engaging with international partners to adopt our technology as a way of future proofing the investment and expanding the pool of potential developers.

The Research Data Alliance has now become the main focus for ANDS' involvement in international activity, 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 re-use, 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 founded through commitments by Australian Commonwealth Government, the European Commission, and United States of America. The work of the Research Data Alliance will primarily be undertaken through its working groups and interest groups. Participation in working groups and interest groups, starting new working groups, and attendance at the twice-yearly plenary meetings is open to all.

In this period, Australia will fund the position of the Research Data Alliance Secretary General. This is both important to the Research Data Alliance in its first period of operation, but also particularly important in Australia's relationship with the US and Europe, cementing the leading role of Australia in research data internationally.

Note that this program will be conducted using a combination of NCRIS, EIF, CRIS, and NCRIS 2013 funds, as well as separate funding under a contract entitled "Data Web Forum".

5.1 Highlights, Difficulties and Breakthroughs

There is nothing to report as the money has not been expended.

5.2 Activities Undertaken

There is nothing to report as the money has not been expended.

5.3 Performance against agreed Key Performance Indicators

There has been no progress against Key Performance Indicators as the money has not been expended.

6 A summary table of actual expenditure for the period covered by this report

Program Title	Actual Expenditure for the period
National Collections	
operation, management and governance costs	\$ -
salaries and on costs for technical staff	\$ -
infrastructure maintenance	\$ -
utilities	\$ -
rent	\$ -
consumables	\$ -
other	\$ -
National Services	
operation, management and governance costs	\$ -
salaries and on costs for technical staff	\$ -
infrastructure maintenance	\$ -
utilities	\$ -
rent	\$ -
consumables	\$ -
other	\$ -
Institutional Engagement	
operation, management and governance costs	\$ 10,327
salaries and on costs for technical staff	\$ 10,162
infrastructure maintenance	\$ -
utilities	\$ -
rent	\$ -
consumables	\$ -
other	\$ 1,379
International Collaboration	
operation, management and governance costs	\$ -
salaries and on costs for technical staff	\$ -
infrastructure maintenance	\$ -
utilities	\$ -
rent	\$ -
consumables	\$ -
other	\$ -
TOTAL EXPENDITURE	\$ 21,868
INCOME RECEIVED	\$ 2,879,006
BALANCE REMAINING	\$ 2,857,138