

ANDS Commentary on the ARDC EIF Draft Final Plan

1. Overview

The Australian Research Data Commons supports the discovery of, and access to, research data held in Australian universities, publicly funded research agencies and government organisations. This investment will enable the construction of a range of ICT utilities to capitalise on and ensure greater use and re-use of existing data resources, as well as better management of new data generated in Australian research.

Specifically the Project will:

- establish the core infrastructure of the Australian Research Data Commons (ARDC)
- construct and install systems that identify and describe significant data holdings for the ARDC
- construct and install systems that permit researchers to explore Australia's research data holdings, and
- construct and install systems that permit researchers to re-use research data through the ARDC.

The rationale for the project is that research is more data intensive, and that there is need for an approach on a national scale.

The work has commenced through involvement in our NCRIS programs, with a collections registry, persistent identifier service, and data discovery service already established. However, the EIF ARDC project will accelerate the benefits by increasing the scale of the ARDC. Work has already commenced in the first 2 months of funding to enable descriptions of the data of some high profile researchers to be made visible through the Commons.

The Australian Research Data Commons Project represents a very significant change in the approach to research data in Australia, enabling Australia's research data as a whole to become a nationally strategic resource. Through EIF Funding, the ARDC will be a populated information infrastructure that will achieve the critical mass necessary to become the primary means by which researchers routinely discover and access Australia's research data.

To succeed in this project, ANDS needs to partner with institutions that create, own and curate research data. Consequently, the Draft Final Plan describes a variety of methods that will be used to engage with others in creating the ARDC together.

2. Programs

The programs of work that constitute the ARDC project are tightly connected to ANDS NCRIS funded programs:

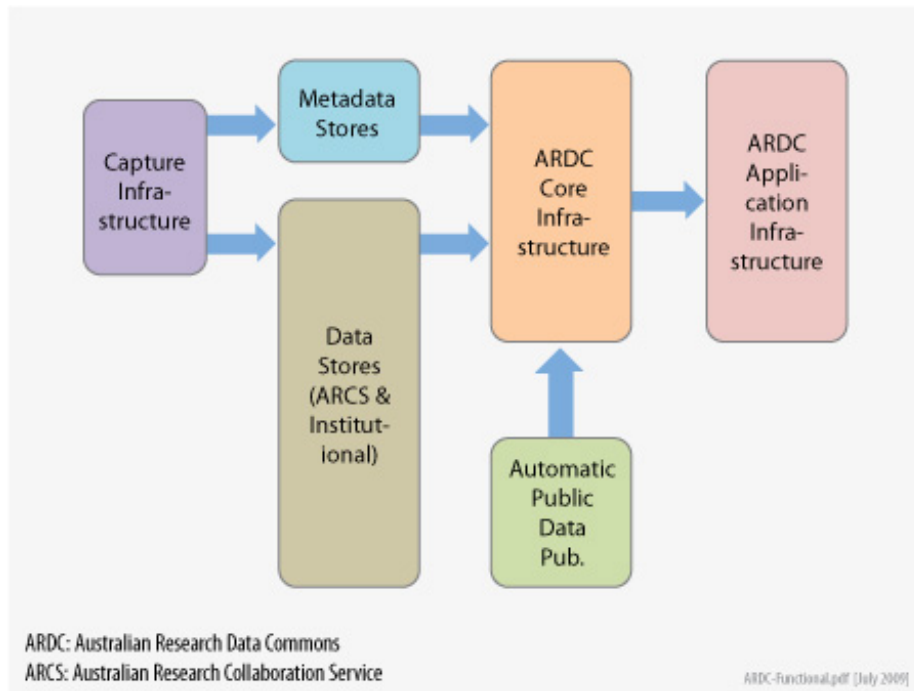


- 1 **Frameworks and Capabilities** – an ANDS driven program ensuring that institutions have the capability and the research system have the structures in place to enable researchers to manage, publish, share and re-use research data (NCRIS-funded)
- 2 **Data Capture** – an institutionally based program to automate the capture of data and metadata from instruments in data intensive research
- 3 **Seeding the Commons** – an institutionally based program to ensure that well managed data is made available through the ARDC with a focus on data that cannot be automatically captured (NCRIS-funded)
- 4 **Public Data Access** – an outsourced program to make more public data collections visible and available through the ARDC
- 5 **Metadata storage** – an institutionally based program that enables metadata to be stored coherently across an institution that supports data management, publishing, sharing and re-use
- 6 **ARDC Core** – an ANDS driven program that puts in place the national services that enable research data to be published and discovered
- 7 **ARDC Applications** – an outsourced program that establishes national services to enable the exploitation of the data in the ARDC.

The next table shows the intended size and focus of the programs.

Programs	EIF (\$)	NCRIS (\$)	%AGE	Focus
Frameworks and Capabilities		4	6%	ANDS
Metadata Storage	8		12%	Institutions
Data Capture	13		19.5%	Institutions
Seeding the Commons		12	18%	Institutions
Public Data Access	10		15%	Contractors
ARDC Core	8		12%	ANDS
ARDC Applications	8		12%	Contractors
EIF Project Management	1		1.5%	ANDS
Project Office		2	3%	ANDS
Total (\$66M)	48	18	100%	

The programs interact as shown in the diagram illustrating how data is made available through the ARDC and is then exploited through the Core and Applications programs.



3. Engagement Methodology

To effectively deliver the ARDC project there are different methods of engagement for each program.

Programs	Focus
Frameworks and Capabilities	ANDS will partner with institutions through small scale engagements both with research institutions and with policy setting agencies.
Metadata Storage	ANDS will fund some representative institutions to install systems that allow them to manage their research metadata in the first instance so that such systems can be rolled out to other institutions.
Data Capture	Institutions will be funded to install systems that capture data and associated metadata from research intensive instruments. The level of funding will be based on the research quantum, and decisions about focus areas will be made jointly by ANDS and the research institute.
Seeding the Commons	Institutions to make research collections available through the commons – this work will be more focused on smaller collections and often consist of a partnership between ANDS and the research institute. The level of funding will be based on the research quantum, and selection will be made jointly by ANDS and the research institute.
Public Data Access	Contractors will be employed to make collections of data available through the ARDC. Selection will occur through discussion with both the data holders and representatives of research disciplines and capabilities.



ARDC Core	ANDS will build some of the core services and through an EOI it will select contractors to build other services that are specified in the ANDS technical roadmap.
ARDC Applications	Through an EOI ANDS will select contractors to build other services that are specified in the ANDS technical roadmap.

4. Consultation Questions

Through this consultation, the Department of Innovation and ANDS seeks to gain your views on the proposed approach to the ARDC project, and your interest in participating. We have 4 specific issues that we seek comment on:

- 1 Are the programs appropriate to achieve the aims of the ARDC project? Is there a program that should not be present? Is there a program that is missing? Should one of the programs be re-focused?
- 2 Is the balance right? Should ANDS spend more effort on one program at the expense of another program?
- 3 Are the engagement approaches appropriate to each program? If not, what engagement approach would you suggest?
- 4 Is your institution willing to engage? How?