The Australian National Data Service – outcomes and achievements after 4 years

David Groenewegen¹, Adrian Burton², Cynthia Love³, Andrew Treloar⁴, Ross Wilkinson⁵

¹Australian National Data Service, Melbourne, Australia, david.groenewegen@ands.org.au
²Australian National Data Service, Canberra, Australia, adrian.burton@ands.org.au
³Australian National Data Service, Melbourne, Australia, cynthia.love@ands.org.au
⁴Australian National Data Service, Melbourne, Australia, andrew.treloar@ands.org.au
⁵Australian National Data Service, Melbourne, Australia, ross.wilkinson@ands.org.au

INTRODUCTION

This paper will describe the work and the achievements of the Australian National Data Service (ANDS) as it approaches the end of its fourth year of operation. The paper will look at ANDS' role in the sector, the underpinnings of ANDS' strategy, how this strategy has been applied and what the Service has accomplished over this period.

ABOUT THE AUSTRALIAN NATIONAL DATA SERVICE

The Australian National Data Service (ANDS) was established in January 2009 following the successful completion of the ANDS Establishment Project. ANDS was originally funded under the National Collaborative Research Infrastructure Strategy (NCRIS) initiative to ensure that research data is used as effectively as possible by Australian researchers. The Super Science Initiative announced in May 2009 provided additional funding from the Education Investment Fund (EIF) to establish the Australian Research Data Commons (ARDC). This provided the opportunity to leverage both NCRIS and EIF funds to build a co-ordinated set of programs through to June 2013.

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 is that Australia’s research data as a whole is steadily becoming a nationally strategic resource.

THE ERESEARCH SECTOR AND ANDS’ PLACE IN IT

The 2011 Strategic Roadmap for Australian Research Infrastructure notes that “NCRIS and the Super Science Initiative have resulted in investments of over $1.4 billion in a broad range of research infrastructure at the national, collaborative scale” [1]. This funding covers a wide range of priority capability areas, new facilities and expansion of existing capacity. ANDS’ role within this, which is to create the infrastructure to enable Australian researchers to easily publish, discover, access, and use research data, and its relationship to other funded areas will be discussed in the paper.

ANDS’ STRATEGY AND ACHIEVEMENTS TO DATE

ANDS is 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 eight inter-related programs of activity (Frameworks and Capabilities, Seeding the Commons, National Collections, Data Capture, Metadata Stores, Public Sector Data, ARDC Core and ARDC Applications). These programs support ANDS partners through a combination of project funding, advice and support engagements.

As a result of these goals, initial activity, and consultations, ANDS’ role can be best described as enabling Australia’s research data to be transformed:

<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 form a nationally significant resource so that Australian researchers can easily publish, discover, access and use Australian research data.
ANDS is contributing to this resource by creating the ARDC, a combination of the set of shareable Australian research collections, the descriptions of those collections including the information required to support their re-use, the relationships between the various elements involved (the data, the researchers who produced it, the instruments that collected it and the institutions where they work), and the infrastructure needed to enable, populate and support the commons. ANDS does not hold the actual data, but points to the location where the data can be accessed.

The outcome of ANDS activity is that Australian researchers have access to infrastructure that enables them to: systematically, reliably and authoritatively connect their research data to project, institutional and disciplinary descriptions, and simultaneously publish citable research data collections through institutional, disciplinary and national services all supported by a wide-ranging set of coherent software, policy and process outputs.

This will ensure that Australia has a mature, globally leading capability in research data, making it a key locus for data intensive research. It will be possible to do this as a result of ANDS and its partners developing a wide-ranging set of coherent software, policy and process outputs that support the Australian Research Data Commons.

Substantial infrastructure has already been created and this will be discussed in detail in the paper. ANDS has or will soon have in place a range of national services, including; a data collections registration service; a data collection description publication service; a national gazetteer service; a data citation service; a researcher identification service; a “see-also” service that enables other discovery tools to use the ANDS discovery service; a research activity identification service; a research project identification service; and an enhanced Research Data Australia – a data collections discovery service.

This is combined with a range of coherent institutional research data infrastructure, including tools deployed to automatically capture rich metadata along with the data for a wide range of instruments; operational metadata stores for this metadata; collections descriptions feeds to ANDS, both Research Institutions and Public Sector data holders; institutions responsible for nationally significant collections will have been supported in making those collections connected, visible and available, often through an RDSI node; over 40,000 collections are available for discovery through Research Data Australia; and discipline oriented portals are cross connected to Research Data Australia.

The following tools, frameworks and capability are in place to further exploit the ARDC: improved institutional guidance for internal institutional data management; improved institutional guidance for responding to national instruments such as the Australian Code for the Responsible Conduct of Research; institution wide research data management planning frameworks or improved research data management at many research institutions; increased institutional capability for research data management with staff trained in this area; and tools that enable more effective re-use of research data. This capability will be demonstrated by high profile researchers showing the value of this infrastructure.

CONCLUSION

The ARDC represents a new scale and quality of data infrastructure to support new forms of research and help answer the important questions that society wants researchers and policy-makers to address. This infrastructure also lets Australian researchers and research organizations join and form new international data-intensive research collaborations.

ACKNOWLEDGMENT

ANDS is supported by the Australian Government through the National Collaborative Research Infrastructure Strategy Program and the Education Investment Fund (EIF) Super Science Initiative.

REFERENCES


ABOUT THE AUTHOR(S)

Mr David Groenewegen

Mr David Groenewegen is currently the Director of Research Data at the Australian National Data Service. Previously he has been ARROW Project Manager and ARCHER Project Director, and spent a number of years working in the areas of electronic information provision and information literacy at Monash University, and in information resources at the University of Ballarat.
Dr Adrian Burton

Dr Adrian Burton is the Director of Services in the Australian National Data Service with particular interest in online services, human capability, and national policy in support of the Australian Research Data Commons. Adrian has taught and researched in South Asian Linguistics at the Australian National University and contributed to ICT infrastructure and support for teaching and research in a number of local, national, and international programs.

Ms Cynthia Love

Ms Cynthia Love is the Director of Public Sector Data with the Australian National Data Service (ANDS). Her professional background is in research libraries and information management, developing services associated with both. She had 20 years experience working with CSIRO, Australia’s premier public research institution, prior to joining ANDS.

Dr Andrew Treloar

Dr Andrew Treloar is the Director of Technology for the Australian National Data Service, with particular responsibility this year for demonstrating the value of bringing together data from different disciplines to answer new questions. In 2008 he led the project to establish ANDS. Prior to that he was associated with a number of e-research projects as Director or Technical Architect: ARCHER (http://archer.edu.au/- an e-Research support environment), DART (http://dart.edu.au - data acquisition and analysis), and ARROW (http://arrow.edu.au/- institutional repository software), as well as the development of an Information Management Strategy for Monash University. He has a B.A. hons. (first class), majoring in Germanic Languages and Linguistics, a Grad. Dip. in computer science, and a research Masters in English Literature, all from Melbourne University. In 1999 he received his Ph. D. from Monash University with the topic Hypermedia Online Publishing - Transformation of the Scholarly Journal. His research interests include data management, institutional repositories and scholarly communication.

Dr Ross Wilkinson

Dr Ross Wilkinson is the Executive Director of the Australian National Data Service, dedicated to enabling more researchers re-use data more often.

His research career commenced with his Ph. D. in mathematics at Monash University before researching in computer science at La Trobe University, RMIT and at CSIRO. Some of his areas of research have been document retrieval effectiveness, structured documents retrieval, and most recently on technologies that support people to interact with their information environments. He has published over 90 research papers, has served on many program committees and was a program co-chair for both SIGIR’96 and SIGIR’98.

He is now leading the Australian National Data Service creating tools, information, frameworks and the skills to enable Australia’s researchers to more effectively use and re-use research data, wherever it comes from.