



Data Management Practice in Australia – the ANDS perspective

David Groenewegen and Andrew
Treloar

What is ANDS?

- ANDS is supported by the Australian Government
- Began in 2009, currently funded to December 2014
 - Collaborative Research Infrastructure Scheme (CRIS)
- Collaboration between Monash University, CSIRO and the Australian National University
- Staff in 6 cities across the country
- Funding around 250 projects

ANDS role:

- Establish several national services in support of research data
- Help populate the Australian Research Data Commons with data collections that are managed, connected, and discoverable and reusable
- Partner with institutions to establish coherent institutional research data infrastructure
- Improve the ability of the Australian research system to exploit its research data through the lens of the ARDC using tools, policy and human capability

National Services

- Research Data Australia
- Register My Data enables the harvesting of description information about collections and the activities, parties and services
- Publish My Data allows individuals to manually register a collection description and to obtain a persistent identifier for their collection
- Cite My Data provides cost free creation, resolution and management support for persistent identifiers (DataCite DOIs)
- Describe My Data enables communities to publish and consume existing vocabularies
- Party Infrastructure for independent party (person/organisation) identifies
- Online Gazetteer provides a publicly available interface (both web and machine readable) to an authoritative national gazetteer of place names and other useful spatial information
- Activity Infrastructure persistent identifiers for ARC and NH&MRC research grants

Research Data Australia

- researchdata.ands.org.au
 - ~44,000 collections of data
 - associated with 8,000+ linked researchers or research staff
 - harvested through 62 separate data sources
 - and covering all 22 top-level ANZSRC FOR codes.
- 2011 usage: **120,457** page views and **68,031** unique page views over the last year.
- Page views doubled from last year.

ANDS as a funder

- Funding primarily provided at institutional level
- Infrastructure programs
 - Metadata Stores
 - Data Capture
 - Public Sector Data
 - National Collections
 - Applications
- Human capacity programs
 - Seeding the Commons
 - Frameworks & Capabilities

Tools, frameworks and capability

- 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
- tools that enable more effective re-use of research data
- tools 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

Changes in the sector

- 2009
 - 8 institutions with some DM capability (Mostly policy and/or storage)
 - Virtually no discovery or systematic understanding of what they have
 - Limited deployment of DM tools
- 2012
 - 38 with some capacity, across a range of areas
 - 34 with discoverable data
 - 28 have done systematic looks at what they have
 - Wide range of tools and policy

Pulsar Data Archive

- Chinese astronomers have used CSIRO's Parkes Observatory Pulsar Data Archive to look for glitch events in pulsar data sets.
- Results of this work, jointly with CSIRO, are soon to be published.
- Archive is also serving as a major resource in an international search for gravitational waves.



Repository of Antibiotic Resistance Cassettes



- database at the University of New South Wales
- instrumental in the discovery of more than 40 new genes in bacteria, including about 10 that may confer antibiotic resistance and are thus clinically important

Life Patterns

- Project of the University of Melbourne's Youth Research Centre (edfac.unimelb.edu.au/ycrc)
- Online Heritage Resource Manager (OHRM)
- New opportunities:
 - Researcher funded by a postdoctoral fellowship from the Swedish government will be collaborating with the Life Patterns research team, to undertake a new analysis of the data.
 - The establishment of the OHRM and associated data management processes has also had a demonstration effect, resulting in a successful ARC Linkage project with new partners

RedMap



- Observations of marine organisms by non-scientists can be added to traditional scientific sources
- Researcher had just picked up an unusual observation of a gloomy octopus (*Octopus tetricus*), traditionally found off Sydney and the east coast of Australia but not in Tasmania.
- Mexican graduate student now looking at why.

Australian National Corpus



- New research projects include:
 - Irish influences on Australian English
 - Further developing an interlingua ontology for use across spoken data collections
- <http://www.ausnc.org.au/>
- The establishment of the AusNC has also seeded collaborative projects including:
 - the development of a virtual Human Science Communication lab
 - the deployment of language data from the AusNC in educational contexts

International

- Research Data Alliance
- ANDS, NSF, iCORDI (EU)
- Purpose:
 - accelerate international data-driven innovation and discovery by facilitating research data sharing and exchange, use and re-use, standards harmonization, and discoverability
 - through the development and adoption of infrastructure, policy, practice, standards, and other deliverables.
- To be launched in March 2013 in Gothenburg.
- Research Data Alliance <http://rd-alliance.org/>

Thank you

- Questions?

David.Groenewegen@ands.org.au

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