

Introduction to ANDS - Ron Sandland, Chair of the Steering Committee



Welcome to the ANDS newsletter. ANDS began its life as a visionary component of the Nation's investment in e-Science. Commencing as part of the Australian Government's NCRIS initiative (the National Collaborative Research Infrastructure Strategy), involving an investment of \$542M over the years 2005-2011, ANDS funding of \$24M lies within the Platforms for Collaboration capability within NCRIS. The decision by the government in its May budget to add \$48M from the Education Investment Fund to pursue the ANDS objectives through its "Super Science" initiative has further demonstrated its commitment to e-Science.

ANDS rests on two critical underpinning ideas:

- » that data is an increasingly important component of the research enterprise as new means of data capture and storage (sensor systems, high performance computing modelling outputs, data repositories etc) become available.

- » that research collaboration is vital to crack some of the major challenges facing humanity in the twenty-first century – energy, climate change and health being just three examples.

Bringing these two ideas together means we have to find new ways to capture and share data across Australian (and ultimately global) research and data producing agencies to ensure relevant, high-quality data is accessible to researchers anywhere in the system. To do this, we have to be able to create accessible repositories and facilitate access to their holdings. This is a non-trivial task that requires us to address such questions as: how can data be captured and stored in such a way as to make its retrieval and re-use as efficient as possible?

I'm very pleased that three research institutions, understanding implicitly the importance of the initiative, agreed to partner to form ANDS: Monash University, ANU and CSIRO. Monash, in particular, has long been a leader in this domain and it has been very important to the initiative that their expertise has been at the forefront of the development of ANDS.

Inside Issue 01

- » Introduction to ANDS from the Chair of the Steering Committee
- » Supporting the Data Lifecycle Symposium
- » ANDS out and about - ANDS engagements
- » National eResearch Architecture Taskforce (NeAT) project update
- » Australian Code for the Responsible Conduct of Research Forums
- » Conference reports
- » Focus on research excellence - TARDIS
- » ANDS launches online services
- » Meet the ANDS staff - Dr Ross Wilkinson
- » ANDS staff visit the Australian Synchrotron

Putting ANDS together involved a major commitment from a long list of players. We were fortunate to have visionary leadership within DIISR in Anne-Marie Lansdown and Clare McLaughlin along with the DIISR team of Peter Nicholson, Jacqueline Cooke and Mardi Jordan. Rhys Francis, who facilitated the NCRIS Platforms for Collaboration capability, and Andrew Treloar played critical roles in shaping ANDS in the early days. The formation of ANDS had important contributions from many others including the members of the Program Management Committee (Paul Bonnington, Alan McMeekin, Cathrine Harboe-Ree, Vic Elliott, Adrian Burton, David Toll, Alex Zelinsky) and the ANDS Establishment Project (Andrew Treloar, Adrian Burton, Margaret Henty, Chris Blackall, Ross Wilkinson, Tracey Hind, John Morrissey).

Robin Stanton, Edwina Cornish and Alex Zelinsky took the lead in their respective founding organizations and provided visionary inputs. The ANDS Technical Forum also played an important role in its foundation (Markus Buchhorn, Adrian Burton, James Dalziel, Rhys Francis, Clare McLaughlin, John Morrissey, Ray Norris, Judith Pearce, Wayne Richards, Nick Tate, Andrew Treloar, Andrew Wilson

and Lesley Wyborn). They were supported by teams from Monash, ANU and CSIRO. We have been very fortunate to attract Ross Wilkinson, formerly of CSIRO, as the Executive Director of ANDS.

Of course ANDS is a far broader initiative than the activities spanned by the three founding partners. As well as representatives from those institutions, our Steering Committee also has representatives from the University of Queensland, the New South Wales Department of Health, and the University of South Australia. My role is that of an independent Chair.

While our job in ANDS is to provide the services and infrastructure that underpin the ability to capture, store, retrieve and re-use data across the Australian research community, we will need to see a parallel change in culture that will result in researchers' seeing depositing and sharing their data as a good thing. We must provide incentives to share as well as infrastructure to facilitate it. We hope the legacy of ANDS will be to facilitate the creation of a true 'data commons' whose existence will mean a more powerful national research capability, and ultimately better research.

Supporting the Data Lifecycle Symposium



L to R: Dr Tim Churches (NSW Department of Health) and Dr Rodney McDuff (University of Queensland)



L to R: David Koch (University of Newcastle) and John Brattan (University of Newcastle)



L to R: Prof Leon Sterling (University of Melbourne) and Prof Paul Turnbull (Griffith University)

Forty people came together at the University of Sydney on February 11 for the *Supporting the Data Lifecycle Symposium*. The aim of the day was to encourage discussion around the stage in the data lifecycle where a dataset first moves into a stable managed environment, to allow for formal collaboration and the preliminary phases of publication. The event was sponsored by ANDS, ARCS and Intersect, and hosted by University of Sydney ICT.

Those present came from varied backgrounds, including research areas, research administration, IT service areas and libraries. The day-long symposium included presentations from ANDS and ARCS, together with presentations from the perspective of the university administration and the researcher.

Major issues of the day included the roles of both ANDS and ARCS in finding solutions to many of the issues posed by improving data management and storage. ANDS and ARCS are not the only groups with a responsibility for supporting researchers, and there is a considerable challenge in the co-ordination of solutions.

Other issues discussed included (from a very long list), disciplinary differences, data citation, metadata description, discover, access, identity management and, curiously, the phenomenon of "petabyte envy".

For further information about this event see:

<http://ands.org.au/events/sdl/index.html>.

ANDS out and about – ANDS engagements

Since the foundation of ANDS the Executive Director, Ross Wilkinson, has travelled to meet with a large number of stakeholders across Australia. As a result of these initial discussions ANDS has entered into closer engagements with a number of organisations. At the time of writing these included:

- » Australian Nuclear Science and Technology Organisation
<http://www.ansto.gov.au/>
- » The Australian National University
<http://www.anu.edu.au>
- » Australian Research Collaboration Service
<http://www.arcs.org.au/>
- » CSIRO
<http://www.csiro.au/>
- » Department of Primary Industries (Victoria)
<http://new.dpi.vic.gov.au/home>
- » eResearch SA
<http://www.eresearchsa.edu.au/>
- » Griffith University
<http://www.griffith.edu.au/>
- » Integrated Marine Observing System
<http://www.imos.org.au/>
- » Intersect
<http://www.intersect.org.au/>
- » iVEC
<http://www.ivec.org/>
- » Monash University
<http://www.monash.edu.au/>
- » Queensland University of Technology
<http://www.qut.edu.au/>
- » TARDIS – Federated Diffraction Image Publication Repository
<http://tardis.edu.au/>
- » Tasmanian Partnership for Advance Computing (TPAC)
<http://www.tpac.org.au/main/>
- » The University of Melbourne
<http://www.unimelb.edu.au/>
- » The University of New South Wales
<http://www.unsw.edu.au/>
- » The University of Newcastle
<http://www.newcastle.edu.au/>

National eResearch Architecture Taskforce (NeAT) projects

- » Atlas of Living Australia – DIAS-B project
- » Australian Animal Health Laboratory – Emergency Animal Disease Bioresponse Collaboration Platform
- » Australian Phenomics Facility – Ontology Driven Data Management for eResearch



Ross Wilkinson gives a presentation at the Australian Research Council. Photo courtesy Brooke Audsley - Australian Research Council



Prof Margaret Sheil and Ross Wilkinson. Photo courtesy Brooke Audsley - Australian Research Council

- » Australian Social Science Data Archive - ASeSS
- » Australian Synchrotron - Remote Computed Tomography Reconstruction and Visualization Service
- » CSIRO – SISS project, Emergency Animal Disease Bioresponse Collaboration Platform and AusCover Workflow, Remote Computed Tomography Reconstruction and Visualization Service
- » Curtin University of Technology– AusCover Workflow project
- » Flinders University – eResearch service around AusStage
- » Griffith University – National Criminal Justice Research Data Network
- » Integrated Marine Observing System – MACDDAP project
- » iVEC – AusCover Workflow project
- » Murdoch University – Bioinformatics Workflows
- » Terrestrial Ecosystem Research Network (TERN) – AusCover Workflow project
- » The University of Adelaide – Australian Plant Phenomics Network
- » The University of Melbourne – Australian Node of the Human Variome project
- » The University of Queensland – Aus-e-Lit project and eResearch service around AusStage and Australian Plant Phenomics Network
- » The University of Sydney – DataMiNX project
- » The University of Western Australia – National Criminal Justice Research Data Network
- » Western Australian Satellite Technology and Applications Consortium – AusCover Workflow project

National eResearch Architecture Taskforce (NeAT) project update

The Platforms for Collaboration website describes the National eResearch Architecture Taskforce (NeAT) as a "committee of experts established under the NCRIS Platforms for Collaboration capability to develop and assist the adoption of new eResearch methods and tools." (<http://www.pfc.org.au/bin/view/Main/NeAT>)

NeAT oversees the process whereby opportunities are identified to develop e-Research tools and turn these into services that are embedded into research communities. These so-called NeAT Projects are co-funded by ARCS and ANDS (because they are expected to result in candidate ANDS or ARCS production services), together with significant in-kind investment from the groups developing them.

There have been two rounds of NeAT projects to date. The first round resulted in six projects being funded and commencing. These are listed at <http://www.pfc.org.au/bin/view/Main/NeATdevelop>. The second round (called late last year) resulted in 69 initial submissions. Of these, eight were approved to move to project proposal stage. These (together with the agreed public versions for some of the proposals) are available at <http://www.pfc.org.au/bin/view/Main/NeATdevelop>. Those with public versions have been approved to proceed to detailed project planning, with the aim of funding flowing from July 1, 2009.

Australian Code for the Responsible Conduct of Research Forums



L to R: Dr Ian Barnes (ANDS), Prof Paul Bonnington (Monash University), Prof Leon Sterling (University of Melbourne), Paul Sherlock (University of South Australia), Anna Shadbolt (University of Melbourne), Cathrine Harboe-Ree (Monash University), Prof John Lynch (University of South Australia), Cynthia Love (CSIRO)

ANDS has held two information gathering forums on the Australian Code for the Responsible Conduct of Research to help us engage with the wider community. The first forum was held in Melbourne on 20th April, 2009 with representatives from Monash University, the University of South Australia, CSIRO, and University of Melbourne. The second was in Sydney, on 21st May, 2009, with representatives from the University of Sydney, Australian National University, University of Newcastle, Intersect and Queensland University of Technology.

These small discussion groups enabled us to compare how institutions are responding to the Code by developing new policies and procedures. ANDS is seeking to understand the issues involved and provide support in data management. Common themes emerged over the two days, along with differences in how each institution responded to those themes.

Strong themes included the importance of support from high levels within the institution, together with a strong governance framework. It became clear that an institutional response across library, IT, research office, archives and records areas is valuable. The leading role may come from any of those areas, but it is of value to have a steering committee or governance structure that facilitates that collaboration.

Issues of access control, storage solutions, training and good practice guidelines on data management and metadata were all discussed, as was the need to engage with the agenda of ethics committees.

Conference reports

International Repositories Workshop

In March, Andrew Treloar attended a number of meetings and workshops in the Netherlands and UK.

The main meeting was a JISC/SURF/DRIVER Workshop on International Repository Infrastructure (held 15 to 17 March). This was an interesting example of a well-facilitated process to try to make the face to face time in Amsterdam as productive as possible. Andrew led the strand dealing with Interoperable Identifier Structure. The resulting output is visible at <http://www.prezi.com/17905/>, and the activity proceeding from the workshop is available for people to see and contribute to at <http://repinf.pbworks.com/>.

While in the Netherlands, Andrew took the opportunity to meet with DANS, the "national organisation responsible for storing and providing permanent access to research data from the Humanities and Social Sciences" (<http://www.dans.knaw.nl/en/>) and SURF, the "the collaborative organisation for higher education institutions and research institutes aimed at breakthrough innovations in ICT" (<http://www.surf.nl/en/Pages/home.aspx>). DANS are active in a whole range of data related projects, including XML for preservation, user direct deposit, persistent identifiers and enhanced publications, and ANDS has identified a number of strong candidates for collaboration.

Andrew also met with Matthew Dovey from JISC and Simon Hodson (ex the University of Hull, now working full time on the JISC Data Programme) to share information about what the respective programmes were seeking to do.

The Dutch co-ordinated approach to data management seems to be synergistic with the approach being taken in Australia, and ANDS anticipates continuing to explore opportunities for collaboration.

DigCCurr and CNI

These two strange acronyms belong respectively to the Digital Curation and Curriculum and the Coalition for Networked Information, which held their Biennial Conference and Spring Meeting respectively in early April, 2009. Margaret Henty represented ANDS at each, and found them very relevant to developments in ANDS.

The focus of the DigCCurr Conference was the practice, promise and prospects for digital curation, with a substantial emphasis on education for digital curation. The ANDS presentation, prepared by Adrian Burton and Margaret Henty, was entitled Building Australia's eResearch Capability: The Challenge of Data Management and based on the challenge facing ANDS in developing eResearch capacity in Australia's research institutions. A copy of the paper can be found at <http://stores.lulu.com/DigCCurr2009>.

As with most conferences, much of the benefit came from the opportunity to meet with others with similar interests and form relationships of mutual benefit.

The CNI Spring Meeting was held in chilly Minneapolis where temperatures were still little above freezing. Fortunately, the air indoors was warm and the presentations absorbing. CNI Meetings always provide the opportunity to learn about the cutting edge, and this was no different. Topics covered included Duraspace, the new alliance of Fedora and DSpace, the role of research libraries in helping researchers manage their data, the publication and citation of research data, researcher identity management, the Open Annotation Collaboration and details of two of the initiatives of the National Science Foundation's DataNet Program. Full details of the program and overheads can be found at <http://www.cni.org/tfms/2009a.spring/schedule.html>.

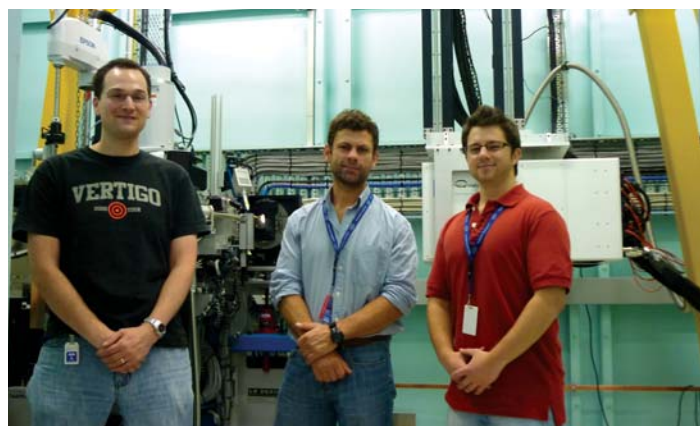
"ANDS began its life as a visionary component of the Nation's investment in e-Science."

Focus on research excellence - TARDIS

TARDIS (<http://tardis.edu.au/>) is a multi-institutional collaborative venture that aims to facilitate the archiving and sharing of raw X-ray diffraction images (collectively known as a 'dataset') from the protein crystallography community. Whereas the model coordinates and (less often) the structure factors (processed experimental data) are stored in the Worldwide Protein Data Bank the raw diffraction data are often not available. Addressing this shortcoming is the central aim of TARDIS. The establishment of TARDIS has been led by Associate Professor Ashley Buckle of the Department of Biochemistry and Molecular Biology at Monash University and NHMRC Senior Research Fellow (<http://www.med.monash.edu.au/biochem/staff/abuckle.html>). The technical wizardry behind TARDIS is thanks to Steve Androulakis.

The archiving and curation of raw X-ray diffraction data that TARDIS facilitates is a critical part of the scholarly record. The determination of protein structures can only be truly repeated or verified when the raw data is stored, shareable and discoverable. In addition, the creators of software to undertake image analysis and processing need sample data with which to work.

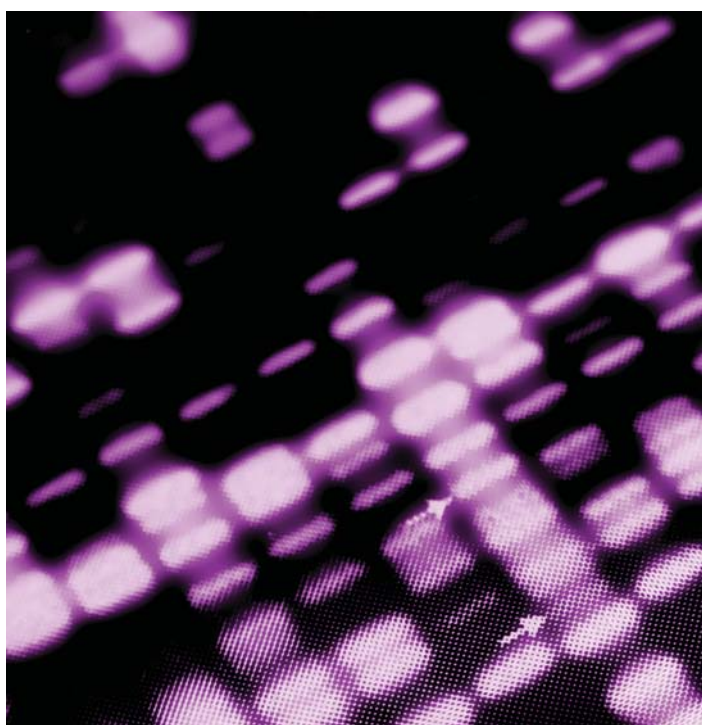
The Australian National Data Service has been working with TARDIS to make this possible. Building on the federated TARDIS



L to R: Dr Tom Caradoc-Davies, Assoc Prof Ashley Buckle and Steve Androulakis in front of the px2 protein crystallography beamline at the Australian Synchrotron

approach, ANDS harvests dataset collection descriptions which the TARDIS software makes available in RIF-CS format (<http://globalregistries.org/rifcs.html>). These descriptions are then loaded into the ANDS Collections Registry and converted into pages that are harvested by web search engines. The end results can be seen at <http://services.ands.org.au/home/orca/rda/list.php?group=Monash+University&class=Collection>.

ANDS launches online services

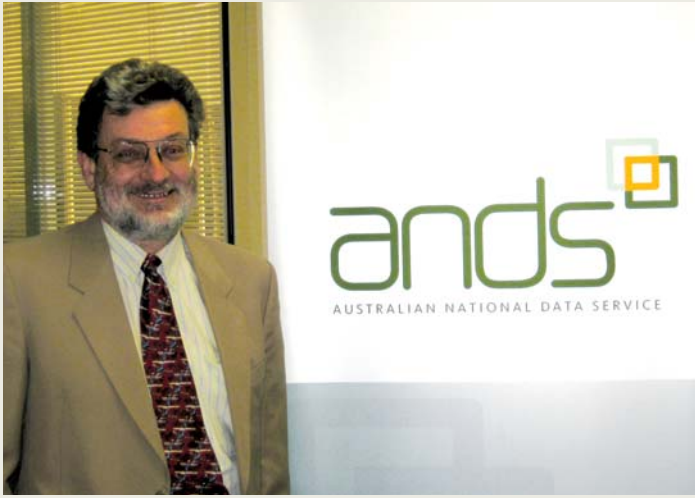


ANDS is launching a number of online services to support data intensive research. The first of these were made public in the month of May:

- » Register My Data (<http://ands.org.au/services/register-my-data.html>)
- » Identify My Data (<http://ands.org.au/services/identify-my-data.html>)

The Register My Data services allow you to register descriptions of your research data. These descriptions are then published in a number of discovery environments. The first of these is the Research Data Australia gateway (to be launched by ANDS this month) which aspires to include any Australian publicly funded data relevant to research and enable innovative cross-disciplinary re-use. Data descriptions registered with ANDS are also fed into other data discovery portals in Australia and internationally, including the big search engines such as Google. The Identify My Data services allocate persistent identifiers to data. These identifiers enable continuity of access even when the location of the data on the internet changes. These new services have been established by ANDS to help realise the vision of a global research data commons.

Meet the ANDS staff - Dr Ross Wilkinson



Dr Ross Wilkinson is the Executive Director of ANDS. He joined ANDS in November 2008 and works from the Monash University office.

Ross obtained his PhD in mathematics at Monash University in 1982. After teaching in the United States and La Trobe University, he joined the Department of Computer Science at RMIT and worked there until 1997, when he joined the Division of Mathematical and Information Sciences at CSIRO as a Principal Research Scientist. In 1998, CSIRO worked with the Victorian government's Public Records Office to develop the Victorian Electronic Records Strategy (VERS), a framework to reliably and authentically archive electronic documents created in government. Ross was the technical director of this project.

From 2000-2004 he was Science and Industry Manager and from 2005-2008 he led the Information Engineering laboratory at CSIRO. During this time Funnelback, the Enterprise Search Company was a spin off from the Information Engineering Laboratory.

His principal research interests are in the area of information environments. Some of his areas of research are document retrieval effectiveness, personal information delivery, structured documents, and the use of XML in document retrieval. He has published over 80 research papers, has served on many program committees and was a program co-chair for both SIGIR'96 and SIGIR'98. In 1998 he co-authored the book entitled Document Computing: Technologies for Managing Electronic Document Collections, published by Kluwer. In the Australian National Data Service his passion is to enable more researchers to reuse research data more often.

Ross believes that the key challenge of ANDS is to build an environment in which researchers expect to store and share their data, and expect to be able to find data, projects and researchers, just as they do for publications. To do that ANDS needs to create a populated data commons, tools to support its use, and partner with institutions to ensure there is an appropriate environment to use it. "For me, it is very simple - I believe in the idea of ANDS - researchers re-using research data - and I want to do what I can to help achieve this idea."

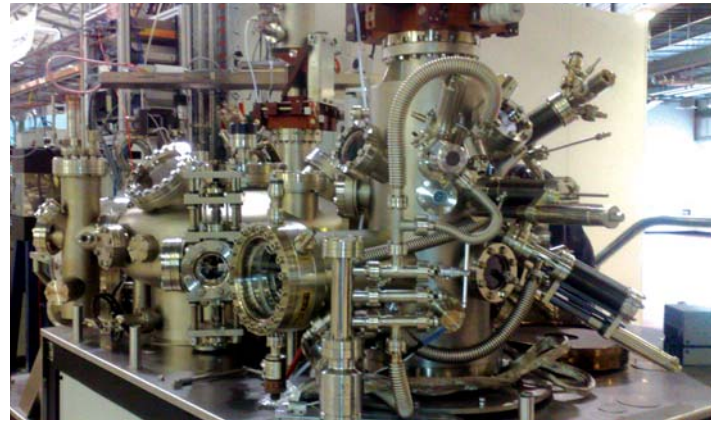
In his spare time Ross recently built a straw bale shack by hand. "It was great fun and very good therapy!"

"For me, it is very simple - I believe in the idea of ANDS - researchers re-using research data - and I want to do what I can to help achieve this idea."

ANDS staff visit the Australian Synchrotron



Richard Farnsworth (centre) shows the ANDS team around the Australian Synchrotron



Soft x-ray Spectroscopy

ANDS staff visited the Australian Synchrotron on Monday 30 June. Richard Farnsworth, the Head of Computing provided an informative tour of the Synchrotron's world-class facilities. Officially opened in July 2007, the Australian Synchrotron (<http://www.synchrotron.org.au/>) is one of fewer than 40 similar facilities around the world. It is the largest stand-alone piece of scientific infrastructure in the southern hemisphere. Over 1000 individual users from research groups have used the synchrotron's state-of-the-art beamlines to further their research objectives.

The group visited the control room where the operation of the beamlines is monitored and had the opportunity to look at a number of the beamline workstations including those for infrared spectroscopy, image and medical and soft x-ray spectroscopy. It was interesting to see how the protein samples are loaded into cartridges for x-ray crystallography research.

Chris Myers, Chief Infrastructure Architect with VerSI showed how researchers are able to login from a remote location and check how their experiment is proceeding using the high throughput protein crystallography beamline. Researchers are able to operate the robotic arm remotely and have the option to mount the protein crystals manually. Chris also showed the group the educational virtual beam line software. The educational virtual beamline demonstrates, in real time, Thomas Young's classic double-slit diffraction experiment. Secondary school students are able to login and view diffraction images and control the slits remotely.

"It is the largest stand-alone piece of scientific infrastructure in the southern hemisphere."

Get notified about our forthcoming newsletters via RSS feed: <http://ands.org.au/newsletter>. For more news, alerts, announcements and discussion subscribe to the ANDS General group: <http://groups.google.com.au/group/ands-general/subscribe>



ANDS Office
 1st Floor, 700 Blackburn Rd
 Building 203, Monash University
 Clayton, VIC 3800
 Telephone: 03 9902 0585
 Email: contact@ands.org.au

ANDS is supported by the Australian Government through the National Collaborative Research Infrastructure Strategy program and the Super Science Initiative.

ANDS Project Partners:



This work is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 2.5 Australia License <http://creativecommons.org/licenses/by-nc/2.5/au/>