

Chair's report - Ron Sandland



Attendees discuss the ANDS EIF Project Plan at the meeting chaired by DIISR in Melbourne

Over the last few months, planning for the new Education Investment Fund (EIF) funding for ANDS has been continuing apace and I must congratulate Ross Wilkinson and his team for an outstanding job in getting the final plans submitted and meeting the very stringent timelines set by Government in doing so.

A crucial part of finalising the plans was a consultation process with all the major stakeholders – universities, government agencies and other groups working in related areas.

The consultation process was a Department of Innovation, Industry, Science and Research (DIISR) initiative and DIISR chaired all the meetings which were held right around Australia – from Townsville to Tasmania and from Sydney to Perth. DIISR presented the Government position and rationale for ANDS and ANDS staff (mainly Ross) presented the initial proposed workplans. We didn't want to go into these meetings with an empty sheet of paper but we were very genuinely seeking input from our many stakeholders.

I was able to get to the first two of these meetings and was impressed by the openness of the discussion, the level of engagement and the strongly constructive support for the objectives of ANDS. One of the most common questions asked was, "How do we engage with ANDS?" I'm also aware of what a herculean effort it was for all involved to cross the country, present, listen, record the feedback and respond to it. The initial plan was largely supported although significant changes were made to the distribution of resources largely to ensure that the high level of demand for engagement at institutional levels could be met.

As an example of the constructive input received, I'd like to cite that of tropical research data. A powerful suggestion came from the tropical institutions for an integrated view of tropical research data, one which could be taken across institutions rather than across research disciplines. And rather than different sets of data held by research groups from individual institutions within the tropics. This was a case of the true spirit of ANDS being reflected back to us.

Inside Issue 02

- » Chair's report – Ron Sandland
- » Executive Director's report – Ross Wilkinson
- » ANDS out and about – Capabilities and Outreach
- » National eResearch Architecture Taskforce (NeAT) project update
- » ANDS Roadshows
- » The University of South Australia's response to the Australian Code of Conduct for the Responsible Conduct of Research
- » Data management solutions to aid Phenomics initiatives
- » Conference reports
- » Focus on Research Excellence – Developing WA marine metadata infrastructure
- » Meet the ANDS Staff - Dr Adrian Burton
- » Supporting the professionals who support eResearch
- » Monash eResearch e-XPO 2009
- » Forthcoming events
- » Taking part in ANDS discussions

And now, having completed our consultation and submitted the final plans that emerged, the next challenge is delivery. We (ie the whole ANDS community) have to spend a significant amount of money in a very short period of time and in doing so ensure that we deliver outcomes of lasting value to Australia's research community.



L to R: Linda O'Brien (Griffith University) and Clare McLaughlin (DIISR) at the Brisbane consultation

Executive Director's report – Ross Wilkinson

ANDS has received \$48M from the Super Science initiative to fund the development of the Australian Research Data Commons. This is an enormous opportunity. The recent ANDS consultation on how to use this funding indicated support for an integrated and comprehensive program of activity. This would result in more research and public data being described, stored, managed, and available to be discovered and exploited. The new ANDS activity will cover high volume research data, hand crafted collections, collections held by Government Departments and by museums - all the publicly funded data relevant to Australian researchers. ANDS aims to enable researchers to share research data more easily.



L to R: Dr Ross Wilkinson (ANDS), Dr Rhys Francis (AeRIC) and Peter Nicholson (DIISR)

But more importantly, researchers will also be able to work on problems in ways not possible without wider data access across the disciplines. This implies significant changes in research practice, and significant potential benefits.

ANDS out and about – ANDS engagements

Capabilities

Part of the work of ANDS is understanding the different capabilities required of the different groups who conduct or support the eResearch process. These groups are diverse, encompassing not just the research team but all those others who provide background policy, procedures, services and facilities.

The ANDS Capabilities Program has been out and about visiting research institutions around the country in order to find out more about how individual institutions are developing their own staff programs to underpin improved data management and sharing. In August, Ian Barnes and Margaret Henty visited the University of Melbourne, Monash University and the CSIRO Marine and Atmospheric Research Division in Hobart, talking with key personnel about their policies, plans and initiatives in this area. These visits followed earlier ones with the University of Sydney, Intersect, iVec, the University of Technology, Sydney, the University of Queensland, Griffith University and the Australian National University.

Outreach – a new approach to connecting with the research community

ANDS has commenced funding Research Data Consultants at the regional level around Australia. These people will form the ANDS Outreach team and, within their local area, will identify the needs of researchers to manage and reuse their research data sets through the data life-cycle. Primary outcomes sought are enhanced data management practices and more research data sets in the Australian data commons. The facilities and tools developed by ARCS and ANDS will enable enhanced capture, storage and access to research collections during a research project from data creation to post publication in a shared environment.

ANDS Outreach Team members will be based either within one of the member of ARCS (MARCS) computing facilities such as eResearchSA, iVEC and Intersect or in some circumstances directly within a publically funded research institution. These appointments

are an ANDS and MARCS collaborative effort. Members of the Outreach Team will have opportunities to meet virtually and in person to share knowledge and experience gained from their engagements with research institutions.

The first appointee as part of the ANDS Outreach service is Dr Andrew Williams who is based at eResearchSA in Adelaide. Dr Williams commenced with ANDS in early September.

National eResearch Architecture Taskforce (NeAT) project update

The first round of National eResearch Architecture Taskforce (NeAT) Projects commenced on or after 1 July 2008. Brief overviews of these projects can be found at <http://ands.org.au/neat-activities.html>, and more detail (including the project plans) at <http://www.pfc.org.au/bin/view/Main/NeATprojects>. The process leading to Round 2 NeAT projects was outlined in the last issue of Share. More details are available at <http://www.pfc.org.au/bin/view/Main/NeATdevelop>.

At the time of writing, the following Round 2 projects have been approved to commence:

- » Aus-e-Stage: will provide performing arts researchers with access to quantifiable research data about performances and collaborations through a mapping service, a network analysis service and an audience contribution service.

- » Australian Node of the Human Variome Project: will create a database and associated data flows to link together information about human genetic variation with its corresponding impacts on development and disease.
- » Phenomics Ontology Driven Data Management (refer separate article in this issue of Share)
- » Remote Computed Tomography (CT) Reconstruction and Visualization Service at the Australian Synchrotron: will develop a service for remote 3D CT image reconstruction, simulation, analysis and visualization at the Australian Synchrotron.

Two other projects are awaiting approval from the ANDS Steering Committee:

- » BioSecurity Collaboration Platform
- » National Criminal Justice Research Data Network.

ANDS Roadshows

The planned series of ANDS Roadshows got off to a promising start with a visit to Perth of four ANDS staff, David Groenewegen, Ian Barnes, Margaret Henty and Frances Watson, in mid-September. Coincidentally, the Perth EIF consultation was planned for the first day of the Roadshow. The two events combined to provide Perth with a strong ANDS presence for a couple of days, enabling much business discussion and strong participation in the Roadshow events.

Attendance was strong at the three Roadshow workshops, with over eighty people in total participating in the two different workshops on Responding to the Australian Code for the Responsible Conduct of Research and a third on ANDS Services. Participants came from universities, other research institutions such as CSIRO, government departments, other educational institutions and iVEC. David Groenewegen also presented a general introduction to ANDS for those unable to attend the daytime workshops.

ANDS would like to thank iVEC, and particularly Valerie Maxville, for their strong support in helping to publicise the Roadshow and their presence at the workshops.



Groups discuss the institutional perspective to the Code at the Perth Roadshow

The second Roadshow was held in Brisbane in the last week of September. Levels of participation were again high, with well over one hundred people registered for one or more workshops. Two speakers from the Queensland University of Technology provided additional interest to workshops. Martin Borchert spoke on the way in which QUT has responded to the Code and Professor Anne Fitzgerald spoke on data and copyright.

More Roadshow visits are planned for October and November, to Newcastle, Adelaide and Hobart.

The University of South Australia's response to the Australian Code of Conduct for the Responsible Conduct of Research

By Dr Mark Hochman, Director, Research and Innovation Services and Vicki Allen, Ethics and Compliance Officer, University of South Australia

In responding to the Australian Code for the Responsible Conduct of Research, the University of South Australia has taken a holistic approach by:

- » Establishing a new research governance framework;
- » Developing a 'Framework for the Responsible Conduct of Research';
- » Systematically working through its suite of research policies, making amendments to existing policies and/or establishing additional policies as required;
- » Moving to embed the Code into the research culture of the institution by reviewing existing training and education processes, identify gaps and developing training programs where appropriate.

The Australian Code requires that institutions provide an appropriate research governance framework through which research is assessed for 'quality, safety, privacy, risk management, financial management and accountabilities of those who play a part in research' (Australian Code Section 1.2.1).

In response, the University's Academic Board approved the establishment of an institutional governance framework in March 2008. In doing so, the University has moved from an organic system with input from separate, disconnected ethics and compliance subcommittees to a more formal reporting relationship for these subcommittees, reporting through to the Academic Board.

This structured and accountable governance framework enables the University to oversee research ethics, integrity and compliance issues and deal with those that require an institutional response.

In creating this framework, the Academic Board established the Research Integrity, Accountability and Compliance Committee (RIACC). RIACC plays a pivotal role within the governance framework as it receives formal reports from its five subcommittees (National Statement Steering Committee, Human Research Ethics Committee, Institutional Biosafety Committee, Chemical Hazards Subcommittee and Animal Ethics Liaison Committee) and addresses issues requiring an institutional response. RIACC is responsible for implementation of the Code.

RIACC has drafted a framework which describes the attributes of responsible research practice through a series of University policies and policy statements and sets out the standards expected of the University and its researchers. It shows how the University complies with the Code from a policy perspective and, in so doing, provides researchers with a 'one-stop-shop' overview of how to conduct research responsibly.

Of particular interest to the ANDS constituency is Section 2 of the Code "Ownership, Storage and Retention of Data." This requires that "institutions must provide facilities for the safe and secure storage of research data and for maintaining records of where research data are stored." Much research data has traditionally been stored by institutions in hard copy with varying degrees of accessibility. UniSA is looking forward to working with ANDS to implement best practice techniques for curation, discovery and storage of research data.

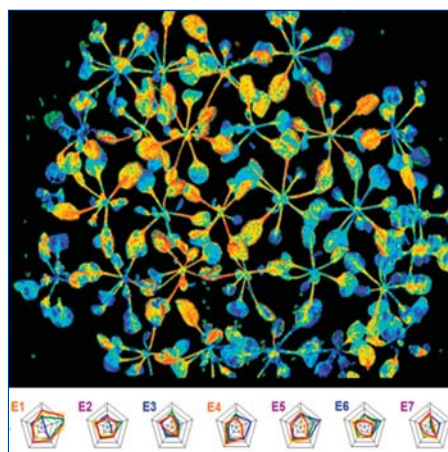
Data management solutions to aid phenomics initiatives

Phenomics is emerging as an exciting science that is providing scientists with the tools to unlock the information coded in genomes. It is the study of how the genetic makeup of an organism determines its appearance, function and performance. The Phenomics Ontology Driven Data Management Project (PODD) (http://www.pfc.org.au/pub/Main/NeATdevelop/PODD_NeAT_May_2009.pdf) is a National e-Research Architecture Taskforce (NeAT) (<http://www.pfc.org.au/bin/view/Main/NeATdevelop>) project co-funded by ANDS and

ARCS. The aim of the project is to develop data management solutions to meet the needs of researchers working at the Australian Plant Phenomics Facility (APPF) (<http://www.plantphenomics.org.au/>) and the Australian Phenomics Network (APN) (<http://www.australianphenomics.org.au/>). Both research communities have the need to gather and annotate data from both high and low throughput phenotyping devices.

Gavin Kennedy is the Project Manager of the PODD data repository project. The development team comprises Xuan-Fang Li who is the Lead Developer and two bioinformaticians, James Eddes and Kai Xu. James Eddes works on the bioinformatics for the plant phenomics clients at the APPF. Kai Xu is a mouse bioinformatician. The development team defines the data models and workflows for the phenotypic data that is to be captured into the repository. Gavin Kennedy says that, "Phenotypic data is an observable/measurable trait of a genotype." In the case of mice, a series of experiments may be run where the mice are put through mutagenesis and their resultant phenotype investigated. This can involve looking for changes in the organism on microscopy slides. In plant phenomics, the biologists look for deviations in breeding, or for variations in the size of the seeds or for improvements in yield.

The image files that are generated by this research range from simple RGB to infrared and 3D and are large. A single investigation can generate tens of gigabytes. In a high throughput environment several terabytes are expected to be generated each year. The bioinformaticians need to identify what metadata needs to be captured. It is important for sufficient metadata to be generated to ensure that the data is organised, can be made accessible to researchers and provides sufficient contextual information for experiments to be understood and reproduced. The metadata capture is done on a case by case basis. Mouse microscopy (histopathology) data is mostly generated automatically, and a parallel EIF project at the University of Melbourne is supporting this metadata capture. Fedora Commons (<http://www.fedora-commons.org/>) repository software is being used to manage, preserve and link data. Data is either left in situ on servers or deposited into a large data store. Gavin Kennedy says, "The intention is to collect and maintain the data



A chlorophyll fluorescence image of Arabidopsis seedlings
Photo courtesy of the Australian Plant Phenomics Facility

that is relevant and store it in perpetuity. This means that data curation and data archiving activities will need to be considered."

In the short term the PODD development team are writing their own methods for encrypted access to the data. Once the Australian Access Federation

(AAF) (<http://www.aaf.edu.au>) becomes available this will be utilized. Generally speaking, biological researchers want to manage their own data and restrict access to those working in their own laboratories. In some instances they may wish to share data with other researchers outside their laboratories. Gavin would like to support open access to data where possible and is hoping to engage with the phenomics community about their data sharing practices. The intention is that PODD is not only a data management tool, but a data publishing tool as well. PODD will also support data discovery through the Atlas of Living Australia and ultimately through the Australian Research Data Commons. A data management workshop is planned for early next year with a focus on data management for bioinformatics. Some of the issues to be explored are access sharing policies within the PODD repository, data usage and attribution of data sources.

Conference reports

Open Repositories

The Open Repositories (OR) series of conferences have become a 'must-attend' event since they were inaugurated in 2005. This year the world wide community of 'open' repository developers and managers gathered together in Atlanta, Georgia, to participate in OR2009. Chris Blackall co-presented a paper about the Global Registries Initiative (www.globalregistries.org), a project supported by ANDS to exchange information about research data collections between registries located worldwide. In addition to presenting, Chris also had the opportunity to attend a number of conference sessions.

Perhaps the highlight was the announcement of DuraSpace (<http://duraspace.org/index.php>) and the sessions taken up with explaining it. DuraSpace is significant because it merges the Fedora Commons (<http://www.fedora-commons.org/>) and the DSpace Foundation (<http://www.dspace.org/>) arguably the largest providers of open source repository software worldwide. The aim of DuraSpace is to sustain and grow its existing flagship repository platforms

(Fedora and DSpace), while at the same time developing new repository services that will live in the 'cloud'. These are designed to respond to the popularity of Web 2.0 environments and to meet new requirements from existing and future users, largely research groups and organizations that need to manage large volumes of data. DuraSpace will feature some of the 'cloud computing' technologies developed by the Grid computing community and represents something of a convergence of interests with the scientific data management community.

Chris also attended several other interesting sessions that focused on overcoming the technical and institutional barriers to recruiting content into repositories. A number of sessions addressed the need for third-party software applications and tools to seamlessly interoperate with repositories. For example, Pablo Fernicola from Microsoft Research successfully demonstrated how one can easily upload files from within Microsoft Word to DSpace using the SWORD protocol.

ESciDoc, CERN workshop on Innovations in Scholarly Communication and SUN PASIG Summer meeting

In June 2009 David Groenewegen visited Europe to attend conferences and meetings. The first of these was the eSciDoc Days in Karlsruhe, Germany June (15-16 June). The eSciDoc Project (<https://www.escidoc.org/>) has been going for 5 years. It was funded by the German Federal Government to develop eResearch tools initially for the Max Planck Gesellschaft (the German equivalent of CSIRO). These tools are based on Fedora, and are all Open Source. David wanted to attend this conference to see what their tools were like, and to explore whether ANDS should be looking at adopting or recommending any of them. The eSciDoc people

were certainly interested in working with ANDS. All the papers (including David's ANDS presentation) can be found at <https://www.escidoc.org/JSPWiki/en/EscidocDays2009Programme>.

From here David headed to the CERN Workshop on Innovations in Scholarly Communication (<http://indico.cern.ch/conferenceDisplay.py?confId=48321>) in Geneva, Switzerland, (17-19 June). This was a pretty big conference and very interesting. After a brief stop in London to meet with Tish Roberts and Dave Cook from JISC to compare notes on organisational issues, it was off to St Juliens, Malta for the SUN Preservation and Archiving Special Interest Group Summer Meeting 2009 (June 24-26). There is a good summary of the papers and the general issues on Carl Grant's blog (<http://commentary.exlibrisgroup.com/2009/06/digital-preservation-observations-from.html>). Most of the papers are at http://events-at-sun.com/pasig_malta09/agenda.php. A very interesting and busy two weeks in all.

Focus on Research Excellence – Developing WA marine metadata infrastructure

By Luke Edwards, Marine Information Officer, iVEC

The organisations of iVEC (the hub of advanced computing in Western Australia - <http://www.ivec.org/>), WAMSI (Western Australian Marine Science Institution - <http://www.wamsi.org.au>) and WASTAC (Western Australian Satellite Technology and Applications Consortium - <http://www.wastac.wa.gov.au/>) are working with IMOS-eMII (Integrated Marine Observing System / eMarine Information Infrastructure - <http://imos.org.au/emii.html>) to help facilitate a lasting legacy of marine data discovery, access and sharing in Western Australia. A major component is the use of the MEST (Metadata Entry and Search Tool - <http://mest.ivec.org/geonetwork/>) which was developed from the GeoNetwork product by the federal BlueNet project.

The initial purpose of installing the MEST at iVEC was to catalogue the large amount of information resulting from the WAMSI research program, which has more than 250 scientists working on 87 research projects. This will help ensure this WAMSI data is discoverable and able to be used into the future. In addition, this MEST forms part of the IMOS distributed data system with Western Australian IMOS metadata to be stored here and harvested to the central IMOS MEST in Tasmania.

The MEST provided an ideal location to store the results of the North West Marine Research Inventory (NWMRI). This was a WAMSI project that created a database of metadata records of the completed, current and planned marine and coastal research in State and Commonwealth waters off Western Australia between



Aerial view of the Kimberley Coast. Photo courtesy of WAMSI

Kalbarri and the Northern Territory border. The plan going forward is to increase the number of metadata records through linking with industry and government databases (such as Interragator+ <https://www.landgate.wa.gov.au/interragatorplus/>) so it becomes a more comprehensive index of all WA marine research projects leading to increased efficiencies and a useful tool for the marine community.

The metadata in the iVEC MEST is currently being harvested into 'Research Data Australia' where it joins other research data collections. ANDS is assisting in this approach through the funding of an eResearch position at iVEC. Part of this person's role will be to extend the metadata in the MEST through outreach into the marine community and will also help achieve uniformity between IMOS's metadata approach and that evolving within NCRIS TERN (Terrestrial Ecosystem Research Network).

Meet the ANDS staff - Dr Adrian Burton



Adrian Burton is a Deputy Director of ANDS and is located at the Canberra office on the Acton Campus of the Australian National University. He has direct responsibilities for the ANDS online utility services that support the data commons as well as for the ANDS "Building Capability" agenda. His Canberra location often involves him in liaison with federal government agencies regarding access to public sector information and the broader policy framework for data re-use.

Adrian worked for some time at the Faculty of Asian Studies at ANU where he completed a doctorate, researched, and taught in South Asian languages, linguistics, politics and religion. He was also language consultant and translator at a number of government departments in Canberra.

Training and working in IT was originally a way of keeping bread on the table as a graduate student but it also opened doors at the ANU into IT systems and services as applied to research and teaching. Adrian soon found himself working as Academic Liaison for central IT services at ANU. A more national focus followed when Adrian became leader of the federally funded Australian Partnership for Sustainable Repositories which brought together leading universities, the National Library, and Australian Partnerships for Advanced Computing.

Adrian is looking forward to the year ahead as it will see ANDS release a number of services to support the publication of data. The ability for research data to be referenced and acknowledged in similar ways to other research publications will be a fundamental "game changer". He is looking forward to a renewed ANDS focus on the broader policy framework for data re-use and the "building capability" agenda.

When on field trips in the sub continent (generally with head in a shawl and sporting a three day growth) the combination of his Irish green eyes, pale complexion, and exotic Hindi accent would often mean he was mistaken for a Kashmiri traveller. Now it's the semantics of research data collections that occupies his mind...

Supporting the professionals who support eResearch



L to R: Margaret Henty (ANDS), Dr Ian Barnes (ANDS), Karen Visser (ANU), Dr Adrian Burton (ANDS) and Janet McDougall (ASSDA) attend the ANDS eResearch Support Forum via a video conference

An important aspect of ANDS aims is recognition of the provision of professional services around the conduct of eResearch. Research in the digital environment requires professional input from a variety of skilled support personnel. These experts may be within the research team, associated with the support infrastructure of the research institution, be part of a state-based eResearch support body or from another institution.

The ANDS eResearch Support Forum has been established to assist the establishment and support of a community of practice for the professionals who support eResearch. The Forum brings together representatives from a range of research institutions to identify and address issues of professional services in support of eResearch. Members come from the University of Sydney, Australian National University, the Australian Social Science Data Archive, Intersect, iVEC, the University of Queensland, the Queensland University of Technology, CSIRO, University of Technology, Sydney, the University of South Australia and National Archives of Australia.

Terms of reference for the Forum were approved at a meeting via video-conference of the Australian Research Data Commons Content Forum which has now changed its name and agreed on its key functions of the new Forum. Of immediate relevance to the Forum is discussion about the provision of materials, such as ANDS Guides, which can assist with staff training.

The next meeting of the Forum will be held at eResearch Australasia in November.

Monash eResearch e-XPO 2009



Photo courtesy of Sharon Walker, Monash eResearch Centre

The Monash University eResearch Centre (<http://www.monash.edu.au/eresearch/>) held an 'e-XPO', Technological innovation transforming research (<http://www.monash.edu.au/eresearch/events/e-xpo/2009/index.html>) on 19-21 August 2009. There were displays and demonstrations of practical solutions in Data Management and storage, data analysis utilising High Performance Computing, state-of-the art Visualisation Services and real-time, global Collaboration Services. ANDS was one of many exhibitors in attendance on Wednesday 19th August. Staff enjoyed the opportunity to talk to researchers about ANDS activities, services and tools. As part of the series of seminars on the final day of the e-XPO, David Groenewegen gave a brief introduction to ANDS. His presentation is available at: <http://ands.org.au/presentations/index.html>

Forthcoming events

Institutional approaches to data management support: exploring different models - Sydney

Workshop to be held in association with eResearch Australasia 2009
9 November, 2009. From 1.30pm - 5.00pm.
Details at <http://www.eresearch.edu.au/>

Integrated data services: ARCS and ANDS - Sydney

Workshop to be held in association with eResearch Australasia 2009
13 November, 2009. From 9.00pm - 5.00pm
Details at <http://www.eresearch.edu.au/>

ANDS Roadshow - Hobart

Two workshops: Research Data and the Code and ANDS Services
1-2 December, 2009
Registration details forthcoming at <http://ands.org.au/events/>

Gumboots for the Data Deluge - Melbourne

To be held in association with VALA2010
Friday, 12th February, 2010. From 9:30am - 1pm
Monash Conference Centre, Collins Street, Melbourne
Registration details forthcoming at <http://ands.org.au/events/>

Taking part in ANDS discussions

ANDS has now launched a bulletin board to facilitate discussion and the exchange of ideas within the ANDS community. Sign up now at <http://community.ands.org.au/>

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.
This newsletter is designed by AK Design (<http://www.akdesign.com.au>)

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/>