



Sustaining Repositories

newsletter of the APSR Project

August 2006

www.apsr.edu.au

In this issue:

- APSR Looks to the Future
- Identifier Management
- Field Researchers - help has arrived
- A Successful Workshop
- Brace Yourself - The Net Generation is coming!
- Who is Accessing Your Repository?
- RQF and Repositories
- News from the Partners
- Meet Danny Kingsley
- Around the Traps

■ APSR Looks to the Future

*"The time has come," the Walrus said,
"To talk of many things:
Of shoes--and ships--and sealing-wax--
Of cabbages--and kings—..."*

Well, perhaps not cabbages and kings, but the time has come for APSR to think about future directions.

APSR was originally established with a three-year time table and three-year funding which come to a finish at the end of 2006. A submission was presented to the Department of Education, Science and Training (DEST) in May outlining future directions for APSR and seeking the funds to support them. Approval of funding was signed off by the Minister, the Hon. Julie Bishop, on July 31.

The activities of APSR have been divided into a number of programs, all of which are to continue in 2007.

Digital Sustainability This program has primarily involved the National Library of Australia to create sustainable solutions for digital preservation. Amongst its successes are important reports, *The APSR Sustainability Issues Discussion Paper* and the *Survey of Data Collections* issued in 2005, and the *PREMIS Requirement Statement Project Report* issued in July 2006.

Further projects are underway in the areas of risk analysis and obsolescence notification. Future funding for this program will make it possible to take the work of these projects through to operational integration with DSpace and Fedora and to make the activities and expertise of the Digital Sustainability program more widely available through outreach events.

National Outreach This program will be familiar to many repository managers, librarians and university administrators though its schedule of workshops, seminars, colloquia, forums and public lectures. The high level of attendance at these events has demonstrated the widespread need for information, training, and advice on digital collections and institutional repositories throughout the higher-education sector, as well as from cultural and other government institutions.

Other aspects of the Outreach Program include the APSR website which publishes project reports and repository tools and provides a forum for the sharing of information.

In 2007, APSR plans to build on its strengths in this area with an augmented national services program. This will be aimed at building technical and management skills which underlie eResearch in the higher education sector.

A program of major events and seminar series is in planning, to include major events on research management and interoperability and eResearch, and a travelling workshop series to go to all capital cities on repository issues. And undoubtedly more.

Repository Development The APSR partnership framework allows a coordinated program for the development of repository software and brings the partners together to address common issues.

Core development programs among the partners are block-funded by APSR. There are currently three of these, each based on different service models for their institutions at the Australian National University, the University of Queensland and the University of Sydney.

Strategic development projects are narrower in focus and more independent development activities that bring together small teams for defined periods. Future areas of focus will be collaborative repository infrastructures, e-Research facilitation, user-friendly repositories, spatially-oriented repositories, copyright management, repository interoperability and integration, and professional development and training. APSR associate members can participate in these projects and a number of universities have expressed interest in joining APSR for this purpose.

International Linkages APSR has established extensive international linkages through its use of open source software and through partner interactions with other institutions working on similar projects.

The development and maintenance of linkages is a long term commitment and builds on previous associations as well as new relationships emerging from the current projects. During the 2006-7 period, linkages will include working with DSpace, Fedora and other communities to improve the interoperability of repository systems; collaborating with groups such as JISC and SURF to identify opportunities for collaborative research and development; continuing international representation at workshops, seminars, and conferences and increasing APSR relationships with Australian and international standards groups.

The Research Data Consultancy Service In 2006, APSR conducted a survey of digital collections and more generally of e-research sustainability. This identified widespread demand for 'hands-on' advice and assistance with establishing and maintaining significant collections of digital data.

The APSR consultancy service is currently staffed and funded mainly through partner in-kind contributions. This will be consolidated and extended and will operate in tandem with the APSR workshop and seminar series to include customised information exchange events with research communities.

■ Identifier Management



The issue of identifier management is critical to the success of repositories and one which still needs a definitive long-term solution. Based on the experience of APSR partners, there is widespread acknowledgement that the higher education sector requires timely and cost-effective services to automate the assignment and resolution of persistent identifiers. We need a coherent national strategy to ensure that critical parts of the national information infrastructure are not endangered.

One component of funding provided to the ARROW Project concerns identifiers and work has commenced on defining the technical, governance and policy needs surrounding them. The three cheerful faces to the left belong to Rob Wilson (consultant), Nigel Ward (IMS Australia) and Kerry Blinco (RUBRIC) who have been given the task of doing this on behalf of ARROW. The trio met with a group from APSR,

APAC and ANU on August 2nd to discuss the issue and, as you can see, the meeting was full of laughter (and practical suggestions).

Some of the questions addressed included: who should have responsibility for governance? How should it be funded? At what level should persistent identifiers be resolved: institutional, regional, national or international? How might institutions migrate their existing identifier schema to national and international ones?

■ Field Researchers - help has arrived

The FieldHelper tool, developed as part of the FIDAS project at the University of Sydney is about to be released. The "Sustainable Data from Digital Fieldwork" APSR-EthnoER Conference and Workshop will run from Mon 4 to Wed 6 December 2006 at the University of Sydney, including a hands-on computer workshop. Watch this space for more details!

■ Brace Yourself - The Net Generation is coming!

Our future customers have a very different approach to their work and life, and libraries will have to change the way they work with their customers to accommodate these changes. This was the message given to a packed lecture room at the University of Queensland on June 30 by Susan Gibbons, Associate Dean of Public Services and Collection Development at the University of Rochester. "Redefining Academic Librarianship for the Net Generation" was a fascinating description of her latest research into the way young American university students are using technology.

As the first generation to have grown up in a digital environment, the 'Net Generation' is defined as those born after 1982. Unlike these 'born digital' students, those of us from earlier generations are 'digital migrants' and will always speak with a 'technological accent'. Susan described this difference as a seismic shift, and those at the lecture walked out realising that new and innovating ways of engaging these students will be needed to avoid the library becoming an insignificant part of future campus life.

The (edited) powerpoints from her talk are available at: <http://www.apsr.edu.au/presentations/successful.html>

■ Who is Accessing your Repository?

The answer to this question was difficult to ascertain until recently, but now ANU has developed a simple and robust statistical analysis tool for DSpace users. The ANU team recognised that while DSpace logs user access to held items, there is no tool to filter, analyse and represent this data. So they set about creating one.

Details of access rates are of interest to anyone with a Collection, so Scott Yeadon and Peter Raftos decided to develop a simple and robust statistical filtering and analysis tool on behalf of the DSpace community.

The purpose of the exercise was to produce a DSpace extension, which allows the on-demand generation of meaningful statistical information for DSpace administrators and Collection owners.

The statistical analysis tool has been well-received in the DSpace community, with the Universidade do Minho, Portugal, DSpace group building on the work done at the ANU (information about their plug-in is available at http://dsug2006.uib.no/index.php?p=pres#eloy_rodrigues)

■ RQF and Repositories

The forthcoming Australian Research Quality Framework (RQF) is a hot topic in universities throughout the country. While some busy themselves with defining impact or worrying about metrics and research funding, others are concerned about the practicalities of how to do it. APSR is providing advice and assistance to those who want to know more through its new web page at www.apsr.edu.au/rqf. Here you will find several helpful documents which provide an indication of how APSR partner universities have gone about it, web links to other sources of information and something about the proposed colloquium to be sponsored jointly by APSR and ARROW and held when more is known about the DEST specifications for the RQF, probably early next year.

APSR has been represented on the RQF IT Working Group, with John Shipp attending the first of the three meetings and Adrian Burton the other two. Belinda Weaver was appointed to a subgroup of the IT Working Group and has attended one meeting where the meeting discussed possible RQF models.

■ News from the Partners

The Australian National University

The ANU testbed has been working with Library staff on the current ANU Academic Promotions round. For promotion, applicants must supply examples of their work to assessors. As some of the material supplied to assessors is in confidence, this process requires strong access controls.

To address the issue, the materials supplied by the academics are being loaded into a DSpace installation where access controls and accounts for the items are being created automatically.

As the academic community braces itself for an upcoming RQF, all testbeds are keeping watching briefs on the RQF. There are strong similarities between the needs of the assessment round and the expected requirements of the RQF and this project has provided a good opportunity to develop techniques and further build on the experience gained from the 2004 research assessment exercise. *Contact: Peter Raftos, peter.raftos@anu.edu.au*

The University of Queensland

Work continues here on Fez development. Changes to the back end have made the system run faster for both indexing and searching. The MAMS/Shibboleth authentication has been integrated into Fez, which means a person from another institution can log into the system via a log in at their home institution. The Fez system picks up their eduPerson attributes via Shibboleth and assigns them rights accordingly.

Suggest has been integrated into Fez. When a term is entered into a Fez search box, the Suggest feature supplies the number of matches expected before a user commits to a search. The suggest database is built from the repository contents and so is updated regularly. The aim is to reduce frustration by giving users the opportunity to narrow or broaden their search terms, depending on the number of expected matches. Further functionality such as multi-word or phrase searching via Suggest will be added.

Statistical reporting has been incorporated into Fez. Current statistics supply the top 50 authors and papers, as well as a breakdown of statistics by latest week, year, all years and so on. While individual download statistics per paper by author are not currently supplied, this information is included with the number of abstract views and file downloads when an author's works are listed.

The launch of UQ eSpace has been delayed to allow Keith Webster, the newly appointed UQ Librarian, to oversee the launch. Also we are currently moving from Dublin Core with extensions to MODS as our basic metadata schema and want to complete this work before the migration of ePrints and ADT material occurs, rather than reverse engineer this. The documentation of migration has begun, but will not be fully complete until the migration has finished. The use of MODS will make the repository software more interoperable. Data has already been exchanged between ANU's DSpace system and Fez.

eScholarshipUQ staff have created a FezWiki to foster communication between developers and users. The Fez developers' email list also supports the developers at a wide range of organisations testing the software. Test sites currently exist in Latvia, the UK, the US, Norway, Switzerland, Ireland and New Zealand. The wiki is at: http://dev-repo.library.uq.edu.au/wiki/index.php/Main_Page. Contact: *Belinda Weaver, b.weaver@library.uq.edu.au*

The University of Sydney

One of the University of Sydney APSR activities for 2006 is to examine processes around complex publication and sustainability. This activity – Rivermouth to Wangga – draws on the experience of ANU in their Rivermouth project, using DSpace and Cocoon to store and render complex multimedia through to publication. This process is being tested in the publication, through Sydney University Press of *Wangga, Songs of North Western Australia*. The authoring, editorial, technical and production processes were recently workshopped at the University of Sydney with the authors (Allan Marret and Linda Barwick) and technical production staff from ANU and Sydney. This collaborative project will generate general workflows and processes that can be applied in the archiving and re-use of complex content. Contact: *Ross Coleman, r.coleman@library.usyd.edu.au*

The University of Melbourne

The Melbourne team has been focusing on reporting the findings from our project: "Sustainable paths for Data-intensive Research Communities at the University of Melbourne" and identifying the implications of these findings at the local and national level. The final project report has been written and is currently undergoing scrutiny from our participating research communities and key stakeholders prior to its release. This process will bring the current Melbourne project to a close. We are keen to build stronger links with other APSR partners as we move into a new phase of engagement with APSR.

Planning for information exchange events is also underway. On September 4th Robin Dale from RLG will be spending a day helping us at Melbourne to grapple with the RLG-NARA Audit Checklist for Certifying Digital Repositories. Later in the year, we will host an eResearch Expo for our researchers. This will provide an opportunity to present some of the findings from the "Sustainable paths for Data-intensive Research Communities at the University of Melbourne" project as well as showcasing the expertise and capability across the university. Contact: *Anna Shadbolt, annams@unimelb.edu.au*

Meet Danny Kingsley

APSR is expanding! Danny Kingsley will be assisting Margaret in the National Services program as a Project Officer two days a week. As a current ANU PhD student looking into Open Access issues, Danny brings a mixture of administrative skills and knowledge to the team. Her position is for an initial six-month term, and APSR is looking forward to a highly productive second half of 2006.

■ Around the traps ... Send your suggestions to contact@apsr.edu.au

Conferences and events

2006 November 21-22, Digital Data Curation in Practice. 2nd International Digital Curation Conference 2006, Glasgow <http://www.dcc.ac.uk/events/dcc-2006/>

2006 November 27-30, 9th International Conference on Asian Digital Libraries, Kyoto, Japan <http://www.icadl2006.org/>

2006 December 4-5, CNI Fall Task Force Meeting, Washington DC <http://www.cni.org/tfms/tf.meetings.html>

2006 December 4-6, Sustainable Data from Digital Fieldwork: APSR-EthnoER Conference and Workshop, University of Sydney

2006 December 5-8, International Conference on Digital Libraries (ICDL 2006) India Habitat Center, New Delhi, India. <http://www.teriin.org/events/icdl>

2007 January 30-February 1, Information Online Conference 2007: 13th conference and exhibition. Australian Library and Information Association, Sydney, NSW. <http://www.information-online.com.au/>

2007 February, Open Repositories 2007: The 2nd International Conference on Open Repositories San Antonio, Texas. <http://openrepositories.org>

Publications

Long-term storage of documents is a pressing issue for repository managers. Word-processing documents and even pdf files are unsuitable in the longer term, and need to be converted into an archival format that retains much of the structure of the original (as opposed to saving as a text only file). There are potential legal problems of ownership and access for documents that are stored in a commercial format (for example MS Word). Ian Barnes has looked at the problem of long term storage and developed a tool – the Digital Scholar's Workbench - which helps academics create their documents in a format that is easily converted into an archival format. The issue of documents in LaTeX is very different and Ian investigated this separately. His findings are published in two easy to read publications:

■ Ian Barnes, *The Preservation of Word Processing Documents*. APSR, Canberra, June 2006. Available as a pdf and as html, downloadable from: www.apsr.edu.au

■ Ian Barnes, *Preservation of TeX/LaTeX Documents*. APSR, Canberra, July 2006. Available as a pdf (html version forthcoming), downloadable from: www.apsr.edu.au

The PRESTA project aimed to develop a requirements specification for preservation metadata based on the PREMIS (PREservation Metadata: Implementation Strategies) final report, the Data Dictionary for Preservation Metadata (May 2005). Bronwyn Lee, Gerard Clifton and Somaya Langley of the National Library of Australia prepared the *PREMIS Requirement Statement Project Report* which will be showcased at Long-term Repositories: Taking the Shock out of the Future, held at the National Library of Australia August 31-September 1.

■ Bronwyn Lee, Gerard Clifton and Somaya Langley, *PREMIS Requirement Statement Project Report*. APSR, Canberra, July 2006. Available in pdf (485kb) and in html, downloadable from: www.aspr.edu.au

Next issue?

Be the first to know ... join the APSR Announcements email service at:
http://mailman.anu.edu.au/mailman/listinfo/apsr_announcements

Contributors to this newsletter were Margaret Henty, Danny Kingsley, Belinda Weaver, Anna Shadbolt and Peter Raftos.

The APSR partners



Supported by



An Innovative Action Plan for the Future.