AONS pilot service launched

On 17 August 2007, a beta version of the Automatic Obsolescence Notification Service (AONS) was launched as a pilot service. The service provides support to repositories anywhere around Australia (or in fact globally) for managing format obsolescence of digital materials.

Users can register with the service, and some already have, by providing a URL to your repository's format scan summary. The AONS service will display the summary and allow the repository manager to compare the formats of items in their repository with information from format registries such as PRONOM and the Library of Congress. These registries flag any formats that are likely to become obsolete. Repository managers can then make curation decisions about any items at risk, such as migrating the formats to something current.

By downloading and installing the AONS software locally, an institution can also take advantage of the pilot risk metrics implementation. This feature requires manual data entry in order to assign weightings to particular formats to create a risk profile for each format.

Over time, by sharing this information between AONS instances, it may be possible to build up accurate risk profiles for many common formats.

The AONS software is the result of the AONS II project funded through APSR and developed by David Pearson, David Levy and Matthew Walker from the National Library of Australia (NLA) with an administrative user interface developed by David Berriman at ANU.

The software can be downloaded from Sourceforge at http://sourceforge.net/projects/aons and a mailing list is also available for support and feedback. As this is a beta release, we welcome feedback through the Sourceforge mailing list.

Please try out the pilot service by sending an email to cosi@apsr.edu.au to register with the service, and tell us which institution you are from.

Time is running out to register for Clever Collections!

Clever Collections: a national showcase of technical innovations for digital collections will be held on November 28 and 29 in Melbourne. If you are involved with the technology underpinning eResearch, a researcher engaged or interested in innovative ways of enhancing research through technology, a repository developer or a curator of databases of research materials, you should register now! Registration and program details can be found at: http://www.apsr.edu.au/clevercollections/. There will also be a workshop on ICE (Integrated Content Environment) on the day preceding Clever Collections.
Loving open access

On Valentine’s Day next year, APSR, in conjunction with QULOC and the University of Queensland (UQ) library, will be hosting Open Access Collections in Brisbane. This will be a one day workshop on the challenges and opportunities of open access publishing for Australian universities. Repositories are being built in every major institution in Australia in preparation for the RQF. This offers a unique opportunity for opening up access to Australian research data and outputs.

Open Access Collections will bring together speakers who work in the area of open access advocacy. This is an opportunity to find out the positive benefits of opening up access to research. Discussion areas will include increased citation rates - with a flow into potentially increased university ranking status, and the economic benefits to both the institution and the country of having openly accessible research outputs.

Two keynote speakers have accepted invitations to take part. Dr Peter Murray Rust works in molecular informatics at Cambridge University and has been active in promoting open source and open data. His group is developing a peer-to-peer system for publishing molecular information at source so it becomes globally available. Kevin Stranak is the Librarian with the Public Knowledge Project at the Simon Fraser University. He works with editors, publishers and librarians from around the world on the use of open source software for open access publishing, and is the author of "OJS in an Hour", "OCS in an Hour", "Getting Found, Staying Found", and other documents published by the Project.

The day will feature several academics who have found benefit in providing open access to their works. Professor Hubert Chanson from UQ, and Professor Ray Frost from the School of Physical & Chemical Sciences at the Queensland University of Technology (QUT) are both currently using repositories to house their work, and will discuss how this change in work practice has helped them in their work and in their citation rankings. Professor Frost is the most downloaded author from QUT’s repository. This event will give repository managers and others working in the area of repositories or libraries some real-life examples to assist them in increasing the uptake and use of their repository.

The conference is conveniently scheduled for the week after VALA2008. We are pleased to announce we have once again secured the historic Customs House, overlooking the Brisbane River for the event. More information about the event can be found at: http://www.apsr.edu.au/open_access_collections/

Applying repositories in an information environment

The tradition started in January 2006 by APSR continues to move from strength to strength. The third Open Repositories conference will be held in 2008. The four day conference, between April 1 and 4, will be hosted by the School of Electronics and Computer Science at the University of Southampton. This group, which developed the EPrints repository platform, have a 20 year history of research into hypertext, multimedia, digital libraries and open access.

The aim of this conference is to address the technical, managerial, practical and theoretical issues that arise from diverse applications of repositories in the increasingly pervasive information environment. Open Repositories 2008 will provide focused workshops and tutorials, followed by general conference sessions that cover cross-cutting and overarching issues and EPrints/DSpace and Fedora user group meetings.

The conference organisers are inviting developers, researchers and practitioners to submit papers describing novel experiences or developments in the construction and use of repositories. Submissions must be in by Friday 7th December 2007, and authors will be notified of acceptance on January 21st 2008.

The themes of the conference include:

- Transformational change in the knowledge workplace
- Professionalism and practice
- Sustainability
- Legal issues
- Successful interoperability
- Models, architectures and frameworks
- Value chains and scholarly communications
- Services built on repositories
- Use cases for repositories

More information about the conference and submission requirements can be found at: http://or08.ecs.soton.ac.uk/
The research data challenge

APSR’s recent forum, Long-lived Collections: The Future of Australia’s Research Data was a big success, attended by government, university and research leaders. The forum discussed the recommendations of Data for Science (published in December 2006) and trends in research data management and preservation.

The event was designed to turn full attention on the recommendations of Data for Science, which was commissioned by the Prime Minister’s Science, Engineering and Innovation Council (PMSEIC) to guide national policy in the management and preservation of research data collections. Three groups of speakers addressed the forum, from the point of view of policy makers, of the researcher and of professional organisations or institutional stakeholders.

Professor Ian Chubb, Vice Chancellor of ANU, pointed out the dilemmas facing research-intensive universities supporting the public-good aspects of research while at the same time endeavouring to commercialise research and innovate. Dr Warwick Cathro from the NLA spoke as a member of the Data for Science Working Group (the Group) and set out the issues for the longer term, especially the task of developing coordinated and sustainable national approaches to data management and preservation. Dr Evan Arthur, a member of the Group and in charge of the Innovation and Research Systems Group of DEST, discussed the NCRIS Platforms for Collaboration initiatives and how they fit into the overall DEST-funded initiatives. Finally, Professor Elim Papadakis, representing the Australian Research Council, discussed the role of the Council as a funding agency and the options it might exercise to acknowledge researchers for preserving their research data.

The second group of speakers addressed data management from a researcher perspective. Dr Ray Norris, CSIRO Australia Telescope National Facility, discussed the international contexts for setting international standards and priorities for managing research data in Astronomy and argued that Australian peak research bodies need to play a greater role here. Dr Kaarin Anstey, Centre for Mental Health Research, ANU, illustrated the complexities of conducting community-based health research and the challenges this raises for data management; specifically, accessing government health data and the resulting privacy and intellectual property issues. Finally, Dr Anne Fitzgerald, from the QUT OAKLaw project, explored the legal and regulatory challenges in providing public access to research data and advocated the adoption of ‘open’ intellectual property licences, such as Creative Commons, that would lower the access barriers to data for researchers and the general public.

The third group of speakers included Professor Graeme Turner, President of the Australian Academy of the Humanities, who discussed the rapidly changing nature of humanities research and strongly argued that these special data requirements should not be overlooked. Dr Markus Buchhorn, ANU, spoke about the findings of Sustainability Issues for Australian Research Data (co-authored with Paul McNamara and extensively quoted in Data for Science). Dr Buchhorn outlined many of the data management challenges ahead, for example, the need to acknowledge and reward researchers for creating and managing data and special training to facilitate the process. Finally, Steve Matheson from the Australian Bureau of Statistics discussed the National Data Network initiative and addressed the opportunities for closer collaboration between the higher education and research sector and Australian Government authorities to better share key data sets. He also suggested that a more cohesive approach to developing national information policy was required that incorporated the interests of government, higher education/research and industry.

Another highlight of the day was an open forum hosted by Professor Warwick Anderson, Director of the NHMRC. Themes such as identifying who has the responsibility for managing and preserving research data collections for the long-term, the need for cultural change at an institutional level to better train and reward researchers for managing and preserving research data, and the need for overall governance mechanisms of the research data domain were all discussed.

Long-lived Collections successfully addressed issues of vital national concern, and although the speakers clearly agreed on many points, the discrepancies between the visionary recommendations of Data for Science and the reality on the ground was clear for all to see. This suggests that solving the research data challenge will remain on the national policy agenda for some time.

Presentations and podcasts from the event are now online: http://www.apsr.edu.au/long_lived_collections/
As many repository managers have discovered from first-hand experience, the process of getting digital resources into repositories lacks automation and is prone to mishap. APSR has contributed to solving this problem through an ambitious project to develop standards and repository software that enable the academic community to automatically transfer new resources and data into repositories.

The project, METS Profile Development, provided the technical foundation for the RIFF (Repository Interoperability Framework) projects to improve the interoperability of repositories. Because most repository software applications have their own, non-standard, methods of ingesting new items, the academic community has faced barriers when uploading or exporting new content into repositories. In addition, few repositories support user-friendly batch-uploading functions, and those that do are incompatible with each other. The RIFF projects aimed to address this. Repositories clearly needed common standards - but which standard? To overcome this problem, APSR turned for help to the National Library of Australia (NLA) and its internationally recognised expertise in standards.

As with many projects, the biggest hurdle was establishing what was needed before work began, explained David Pearson from the NLA. An initial meeting was held in March with representatives from ARROW and the APSR partners. “We were defining how the profile would work, did we want a generic profile or one for each content type? It wasn’t clear cut,” he said.

METS, which stands for Metadata Encoding and Transmission Standard, was originally developed as an outcome of the Making of America II project funded through the Library of Congress, which digitised historical collections for web access. During this process it became apparent that there were no standardised methods for ‘packaging’ the collections so that they could be automatically transported between institutions for submission to repositories. So the Library of Congress developed the METS protocol for packaging digitised content—much like an addressed envelope would package a letter for posting. This package (commonly referred to as a Submission Information Package by repository managers) comprises digital objects (or files) and metadata; that is, information about the digital objects. Metadata comes in three categories—descriptive, administrative and structural—and is expressed as XML (eXtensible Markup Language) files, a universally used format.

While METS is by no means the only packaging protocol available, it has rapidly become the de facto standard adopted by the commonly used repository software platforms. But METS only provides users with a ‘scaffold’ to package content. The other half of the package must be developed and described by communities in response to their particular usage requirements. These requirements are expressed as a ‘METS application profile’. The purpose of an application profile is to detail the precise ways that digital content is described and structured using the METS scaffold so that it can be easily uploaded and processed by the repository software.

“At first we thought it would be a relatively simple task”

Creating a new METS application profile is no easy matter; it requires community consensus and considerable technical expertise and attention to detail. To get the ball rolling, the NLA, through involvement in the RIFF projects, took the lead in developing a number of METS application profiles for specific publication types: e-journals issues, conference proceedings, image collections and so on. For example, working with the University of Sydney, a METS application profile was developed for automatically ingesting electronic journals (published using the Open Journal System) into repositories. In practical terms, this means that an e-journal editor can simply click a button when a new journal issue is completed and it will be automatically uploaded to a target repository for safekeeping and guaranteed public access.

“At first we thought it would be a relatively simple task,” said Judith Pearce of the NLA. “But we found that there were still a lot of standards unwritten, there was still much to be done.” People working in Australia are an important part of the international community of those developing standards, explains Judith. “In Australia we often contribute internationally to standards by taking a holistic perspective. We have a good national infrastructure, and what we end up putting out is what we consider to be best practice. I anticipate that other communities might not have done it in the same way but our work will help them go further than if they start from scratch.”

Clearly, developing METS application profiles is a complex and time consuming activity, beyond the resources of most Australian universities. Having the expertise and support of the NLA for this task has meant that Australia has developed reliable, user-friendly, methods for increase the flow of new materials entering its repositories. APSR is keen to share the benefits of this work and the METS application profile will registered with and maintained by the Library of Congress in late 2007, and can be freely used by the international community.

David Pearson and Megan Williams of the NLA are to be congratulated for their leadership and hard work on making this project a success. Thanks also to go Dr Judith Pearce, Colin Webb, Bronwyn Lee and Gerard Clifton of the NLA for their contributions to the project. More about METS: [http://www.apsr.edu.au/nla-mets/](http://www.apsr.edu.au/nla-mets/)
Creating the BEST statistics

As research institutions increasingly mandate the provision of open access copies of members’ research outcomes, usage data is becoming important for informing policy and funding decisions, and assessing the impact of research at an individual, subject and institutional level. The reliability of this information is limited, as there is currently no standard form of presenting and filtering the data and no service which aggregates the information across institutions to allow higher level analysis and comparisons.

The Benchmark Statistics Project (BEST) being conducted by APSR in the second half of 2007 will create a service to fill these gaps and allow research to be assessed at various levels to inform a wide range of decisions in the academic world. The BEST project will work with a select number of research repositories to define a standard format for the presentation of individual repositories’ usage statistics, and agree on a ‘code of practice’ for the handling of aggregated data.

The project aims to design a pilot aggregation service and document the capacity and needs of the participating repositories. The service will be designed to allow the inclusion of other types of repositories in the future, and to accommodate anticipated needs.

A reference group has been established and met on October 17 to agree on a specification that could be used to develop software for the Australian higher education sector.

For more information, see the BEST wiki http://pilot.apsr.edu.au/wiki/index.php/BEST

Looking after your data

More than 60 researchers, postgraduate students and repository managers attended a workshop designed as a practical introduction to the topic of creating a research project data management plan. The seminar was held at the University of Queensland (UQ) library on Friday, 12 October.

Speakers at Creating a research project data management plan stressed the importance of good data management practice for a wide variety of reasons - the importance of research data as a valuable national asset, the potential it opens up for data sharing or re-use, which has benefits to research as a whole, as well as the fact that well-managed data is more likely to be preserved, which is crucial for unique and historically significant data such as observational datasets.

Well-managed data is also more likely to be discovered by others. This can facilitate the development of genuinely multidisciplinary approaches to issues of national importance. Greater visibility of data can also reduce the waste of funds on duplicate data collection. Belinda Weaver of UQ eSpace presented a checklist for researchers constructing a plan from scratch. Jonathan Manton from the Australian Research Council stressed the importance of good data management from the perspective of the funding bodies, which now make a data management plan mandatory as part of funding conditions. Professor Deborah Terry of UQ made the point that accessible data helps the research and knowledge base develop more rapidly.

ORCA Support Network members are eligible for funding from APSR to support their new training activities on condition that training materials are shared and made available via the web. As a result of this, the UQ training day has been spotted from elsewhere, prompting the following query from the London School of Economics:

The LSE Data Library, in partnership with Edinburgh, Oxford and Southampton, is currently involved in a project exploring digital repositories of research generated data and the policies and procedures in which to embed those. The project is called DataShare: http://www.disc-uk.org/datashare.html and one of the issues we are facing is training our researchers on data management issues. This is why I see synergies with our activities and would appreciate any information that we could share about those.

Presentations from the workshop are online at http://www.library.uq.edu.au/escholarship/orca.html. More information about ORCA training on offer from the University of Sydney and the University of Melbourne can be found at http://www.apsr.edu.au/orca/orca_training.htm.
Long term data management

Sydney’s murky drizzle failed to dampen the high-spirits of the people participating in the Data Preservation Management workshop held on October 25 and 26 at the Darlington Centre at the University of Sydney. It was an engaging and challenging two days of presentations, discussion and class work.

The workshop gathered people from a range of occupations; librarians, archivists, computer scientists, directors, data managers and others to explore and reinforce all the facets of establishing and managing data preservation frameworks.

Both presenters radiated an innate understanding and unwavering enthusiasm about the issue. Anne Kenney, from Cornell University, examined the Open Archival Information System (or OAIS) and delved into the issues facing people heavily dependent on accurate information. Nancy McGovern, from the University of Michigan, further reinforced that data preservation management required clear and practical policies in order to keep pace with the changing technology-driven world. The dynamic interplay between the two speakers created a lively and involving atmosphere and fuelled many fascinating discussions within the workshop.

APSR has received exceptionally positive feedback:

“Thank you for two excellent days!”

“Workshop was very comprehensive and highly relevant; presenters were highly knowledgeable on the subject.”

“The two presenters were great and it was fantastic to get a breakdown of the OAIS model. This was really useful and the document now makes more sense after the explanation Nancy and Anne gave.”

The Data Preservation Management two-day workshop was provided by the University of Michigan and Cornell University and organised by APSR in conjunction with the University of Sydney.

Report from the field – Kevin Bradley at JTS

APSR assisted Kevin Bradley’s trip to Toronto to attend the Joint Technical Symposium (JTS) (http://www.jts2007.org/) on June 28-20 this year. Kevin has reported here on his experiences:

JTS is an international meeting of all the technical participants of the international enterprise to preserve the world’s audiovisual heritage. It is held under the auspices of UNESCO through the Coordinating Council of Audiovisual Archives Associations. It is held every three or four years. AMIA was the host organisation in Toronto.

JTS is a forum for many of the big players in audio-visual archiving, such as major film and audio content providers, as well as equipment manufacturers and service providers. It is also where these large private industry concerns meet with libraries and archives, such as the Library of Congress, British Library, the national libraries of many countries, national sound archives and many public broadcast companies. UNESCO also funds a number of attendances on a needs basis.

When it comes to audio and audio-visual archiving, it is critical that manufacturers, private industry content providers and creators, libraries (most often publicly funded) and archival institutions come together if attempts to maintain access to this historical content are to be successful. While in the world of print publishing and bibliographic standards, the major libraries of the world are an integral part of standards discussions, audio-visual archiving is primarily driven by the broadcast, film and recording industries, and libraries need to argue for a place in the discussion. Successful control and preservation of audiovisual collections needs standardised metadata in the same way as print, but the archival institutions are only included in these discussions if they find the opportunity to participate. This is why JTS is so important. It is one place where these opportunities occur.
I attended JTS to speak about the recently published UNESCO report *Towards an Open Source Archival Repository and Preservation System: Recommendations on the Implementation of an Open Source Digital Archival and Preservation System and on Related Software Development* [http://www.unesco.org/webworld/en/mow-open-source](http://www.unesco.org/webworld/en/mow-open-source), which I prepared on behalf of the Sub Committee on Technology (SCoT) with Junran Lei and Chris Blackall of APSR. This report was jointly funded by UNESCO Memory of the World (MoW) and APSR.

The report was well received. There is a clear recognition that a reliable, low-cost solution is necessary for many communities, regardless of whether they are national libraries in developing countries, or the local library in a wealthy country. There was some significant discussion out-of-session about suitable technologies and some useful recommendations on integration approaches. During the meeting, and in emails since, a number of people have continued to discuss the issues raised. UNESCO MoW has recently agreed to support a pilot project and is looking for partners in the project.

There has been an approach to UNESCO to translate this document and its partner publication *Risks Associated with the Use of Recordable CDs and DVDs as Reliable Storage Media in Archival Collections - Strategies and Alternatives* ([http://unesdoc.unesco.org/images/0014/001477/147782E.pdf](http://unesdoc.unesco.org/images/0014/001477/147782E.pdf)) which was launched at SCoT September 2006 meeting in Mexico City. They will appear in Spanish in print form, bound into a single book.

Since the JTS meeting and the International Association of Sound and Audio Visual Archives, where Kevin also spoke, there has been a development. UNESCO has provided significant funding to develop such a system, and the Memory of the World Sub Committee on Technology is actively looking for partners to help resource the development of the system.

### APSR’s public profile

It is good to see the general media is taking an interest in the important issues that APSR works hard to address. APSR has been in the news recently, with top billing on the Computing page of the *Canberra Times* on 1 October. The article, by Simon Grose, called “Scientific slant on data storage”, described the Automatic Obsolescence Notification Service (AONS) and discussed the Australian eResearch Sustainability Survey Project report in some detail.


### Hidden heroes – James Blanden

Our Hidden Hero for October is James Blanden, who joined APSR at the beginning of 2007 as a developer on secondment from the Enterprise Systems Program—part of the ANU Division of Information. His work with APSR has been on the ORCA (Online Research Collections Australia) Registry project and the supporting Collection Infrastructure Services (COSI) web applications framework.

The ORCA-Registry is the first implementation of the ISO2146 Draft Standard—Registry Services for Libraries and Related Organisations. The process of developing a functioning prototype has generated many ideas on where improvements and refinements could be made to the standard, and has also delivered a compelling demonstration of the potential benefits of this type of service.

The software is based on the open source (and cross-platform) technologies of PHP and PostgreSQL, and will ultimately be released for use by anyone who’s interested. The pilot is running at [https://pilot.apsr.edu.au/cosi/](https://pilot.apsr.edu.au/cosi/), and further information can be found at [http://www.apsr.edu.au/orca/index.htm#registry](http://www.apsr.edu.au/orca/index.htm#registry).

### Staff news

As APSR moves towards the end of its funding period, a flurry of activity is occurring to finish off the various projects. To assist with this process, APSR has taken on extra staff members. Laura Vallee has joined the team as the manager of the BEST Project. Tran Le has taken on an administrative role to assist with finances and web publishing. Chris Blackall has rejoined the fold as a consultant and APSR is grateful to have Leo Monus on secondment from ANU who is undertaking additional programming work.
Conferences and events

2007 15-16 November, Appraisal in a Digital World, Rome, Italy
http://www.dpc.delos.info/rethinking_appraisal/

2007 12-13 December, Curating our Digital Scientific Heritage: a Global Collaborative Challenge, Washington DC, USA
http://www.dcc.ac.uk/events/dcc-2007/

2008 5-7 February, VALA2008: Libraries/Changing Spaces, Virtual Places, Melbourne, Australia

2008 1-4 April, Open Repositories 2008, Southampton, United Kingdom
http://or08.ecs.soton.ac.uk/

2008 21-23 April, Fourth Nordic Conference on Scholarly Communication, NCSC 2008, Lund, Sweden
http://www.lub.lu.se/en/ncsc.html

Publications (reading)

This chapter is aimed at those who are involved in planning, setting up and running a digital repository for an institution or community. It covers making a business case, costs, staffing requirements, managing growth and change and other sustainability issues. A number of repository case studies across Europe were used to derive data to inform the study.
Report downloadable as a Word document from: http://eprints.ecs.soton.ac.uk/14455/

Based on actual work-practice studies of students and faculty -- original research conducted at the University of Rochester -- this visionary guide maps potentially vibrant futures for academic libraries.
More information: http://www.alastore.ala.org/SiteSolution.taf?_sn=catalog2&_pn=product_detail&_op=2341

Nancy Fried Foster and Susan Gibbons (eds), Studying Students: The Undergraduate Research Project at the University of Rochester, ACRL 2007.
This book provides a view into the groundbreaking application of ethnographic tools and techniques to the understanding of undergraduate students and their use of information. The publication describes findings of the work at the University of Rochester River Campus Libraries and provides insight into how academic librarians might use these techniques on their own campuses
More information: http://www.alastore.ala.org/SiteSolution.taf?_sn=catalog2&_pn=product_detail&_op=2434
Report downloadable as pdf: http://docushare.lib.rochester.edu/docushare/dsweb/View/Collection-4436

Next issue?
Be the first to know … join the APSR Announcements email service at:

Contributors to this newsletter were Margaret Henty, Danny Kingsley, Chris Blackall, Tran Le, Belinda Weaver and Kevin Bradley.

The APSR partners

Supported by

This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 2.5 License:
http://creativecommons.org/licenses/by-nc-nd/2.5/au/