

Transforming scholarly communication: a Monash University perspective

Introduction

New technologies have dramatically altered scholarly publishing and communication¹ over the past decade, and have considerable potential to deliver further transformation and innovation. The traditional system of publishing peer reviewed print journals is said to have been in existence, largely unchanged, since the 17th Century.² In a little over a decade the basic publishing medium, at least for scholarly journals, has swung from print to electronic,³ and there are now many institutions and researchers experimenting with new ways of disseminating their research findings outside the formal publishing model, especially through the establishment of electronic publishing channels of their own, and through the development of institutional repositories of research publications.

While the driver for many of these initiatives has been the need to tackle the pricing crisis that has accompanied the conversion from print to electronic publishing, it has rapidly become apparent that these new directions also provide exceptional opportunities for improving access to research and, as a consequence, increasing its impact.⁴ These new methods of dissemination supplement, complement and sometimes bypass traditional scholarly monographs and peer-reviewed print journals.

¹ Scholarly communication is defined for the American Research Libraries organisation thus:
Scholarly communication can be defined as the process by which scholars and scientists conduct their research and make the results of their work known. Encompassing both formal and informal means, scholarly communication is critical to the advancement of knowledge and a scholar's career. In the formal process of publishing, researchers, building on the works of others, write up their findings and give them essentially without charge to publishers. In turn, publishers manage the peer-review process, provide editorial improvements, and distribute the work widely through the inclusion in scientific journals. The journals are then purchased by libraries which organize, provide access to, and preserve them for future generations of [scholars].

Mary M. Case, 'Igniting Change in Scholarly Communication: SPARC, Its Past, Present and Future', *Advances in Librarianship*, 26, 2002. [online] http://www.arl.org/sparc/SPARC_Advances.pdf. (Accessed: 8 June 2004).

² The first peer reviewed journal is generally agreed to have been the *Philosophical Transactions* of the Royal Society of London, first published in 1665. Quoted in David J. Solomon 'Talking Past Each Other: Making Sense of the Debate over Electronic Publication', *Firstmonday*, 2002. [online] http://firstmonday.org/issues/issue7_8/solomon/index.html (Accessed: 8 June 2004).

³ Van Orsdel and Born argue that electronic journals have now become the predominant media. See Lee Van Orsdel and Kathleen Born 'Periodicals Price Survey 2002: Doing the Digital Flip', *Library Journal*, 15 April 2002. [online] <http://libraryjournal.reviewsnews.com/index.asp?layout=article&articleid=CA206383&publicat> (Accessed 8 June 2004).

⁴ A discussion paper – *Managing Australian research output for increased return on investment: the role of open access institutional repositories* – which focuses on the issue of impact is appended.

The drivers behind the transformation of scholarly communication include:

- unacceptable price increases, especially of commercially published journals⁵
- in relation to commercially published material, barriers to access imposed by licensing terms and technology, restrictions that disproportionately affect less affluent countries or institutions
- a desire to increase the impact of the research of individual scholars, and of faculty or institutional research
- an interest in exploiting new technologies to improve institutional management of information
- the need for institutions to preserve their research output through new iterations of technology
- an interest in utilising new technologies to support innovation.

With regard to this last point, Clifford Lynch, one of the key spokespeople on the digital revolution in scholarly publishing, argues that the exploitation of the World Wide Web to disseminate scholarship is now following a relatively conservative path. In his view, the urgent task at the present time is to develop institutions, and cultures within them, that will nurture and support the exploration of authorship in the new digital medium.⁶ While this may be true, there are still very few universities that have managed to incorporate the foundation activities into their operations systematically in ways that achieve substantial benefits and support ongoing innovation. By creating and implementing an information management strategy, and embracing the transformation of scholarly communication within this strategy, Monash University is positioned to achieve this.

The open access movement

A great deal of the debate about the transformation of scholarly communication has been preoccupied with the issue of providing open access⁷ to research output. This is sometimes motivated by philosophical arguments, and sometimes by economic. There are now many models of electronic publishing to examine, from open access supported fully by institutions (usually the university library or individual faculty members), to author pays, such as BioMed Central⁸ and *PLoS Biology*,⁹ to low cost subscription, of which there are many examples.

⁵ Data collected by the American Research Libraries in 2001 revealed that the unit cost paid by research libraries for serials increased by 226% between 1986 and 2000. In comparison, over the same period, the consumer price index increased by 57%. (Mary Case, op.cit.). This picture is paralleled in Australia, with this university still estimating annual journal increase of 14%.

⁶ Clifford A. Lynch, 'Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age', *ARL Bimonthly Report* 226, February 2003. [online] <http://www.arl.org/newsltr/226/ir.html> (Accessed: 8 June 2004).

⁷ The term open access is used here to mean "free and unrestricted online availability", which is the definition used by the Budapest Open Access Initiative. See [online] <http://www.soros.org/openaccess/> (Accessed 8 June 2004). For further information about open access initiatives refer also to one of a number of statements, particularly the *Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities*. [online] <http://www.zim.mpg.de/openaccess-berlin/berlindeclaration.html> (Accessed: 8 June 2004); *IFLA Statement on Open Access to Scholarly Literature and Research Documentation*. [online] <http://www.ifla.org/V/cdoc/open-access04.html> (Accessed: 8 June 2004).

⁸ The Library currently subscribes to BioMed Central on behalf of the university. This allows all Monash researchers to publish with BioMed Central free of charge. See [online] <http://www.biomedcentral.com> (Accessed: 8 June 2004).

There is, of course, no such thing as no-cost publishing, although the costs may – for a time – be hidden. This paper takes the position that, ultimately, open access publishing is a desirable goal from the point of view of encouraging and supporting research, but particularly from the point of view of maximising research opportunities for those nations, organisations and individuals that cannot afford access to high-quality information. The approach being taken by Monash University is one of philosophical support for open access publishing within a framework of attempting to find sustainable publishing models (there being quite a gap between Monash University ePress’s partial cost-recovery model, for example, and the recent annual billion dollar profit achieved by a major commercial publisher).

It is noted here that:

- The Australian Vice Chancellors’ Committee is a signatory to the Budapest Open Access Initiative,¹⁰ which invites “governments, universities, libraries, journal editors, publishers, foundations, learned societies, professional associations, and individual scholars who share [the] vision to join ... in the task of removing the barriers to open access and building a future in which research and education in every part of the world are that much more free to flourish”.¹¹
- The Group of Eight Vice Chancellors have recently agreed to a statement on open access to scholarly information.¹² The stance taken by the Group of Eight

⁹ *PLoS Biology* is a new (2004) open-access journal published by the Public Library of Science, an international non-profit organisation of scientists and physicians. See [online] <http://www.plosbiology.org> (Accessed 8 June 2004).

¹⁰ See [online] <http://www.soros.org/openaccess/> (Accessed: 8 June 2004).

¹¹ See [online] <http://www.soros.org/openaccess/read.shtml> (Accessed 8 June 2004).

¹² The following draft Group of Eight statement on open access to scholarly information is reproduced with permission from the Group of Eight Secretariat:

The Group of Eight vice-chancellors, representing Australia’s pre-eminent research universities, record their commitment to open access initiatives that will enhance global access to scholarly information for the public good.

The vice-chancellors note that:

- information, if it is to achieve maximum benefit for society, must be readily available to a global audience
- the rapid development of digital communication technologies provides expanded opportunities for the widespread dissemination of scholarly information
- new business models are required to ensure that scholarly publishing is cost effective
- any development in digital publishing must incorporate the current framework of scholarly publishing standards relating to the quality of inquiry and reporting
- digital publishing initiatives must appropriately recognise and protect the intellectual property of the authors and require accepted standards of attribution
- the Group of Eight universities are providing leadership in the development of digital publishing initiatives in Australia.

The vice-chancellors support:

- ongoing development of open access initiatives in Group of Eight universities
- digital publishing practices that underpin the timely and cost-effective dissemination of the highest quality scholarly information with a commitment to good practice
- further examination of criteria for promotion in new publishing models.

April 2004

is to encourage cost effective solutions to the issue of disseminating research publications as widely as possible.

A strategic response for Monash University

Monash University has a tradition of innovation, not least in the use of emerging technologies to advance learning, teaching and research, and it has a commitment to “improve the human condition by advancing and transmitting knowledge through research and education and by a commitment to social justice, human rights and a sustainable environment”.¹³ In addition, Monash University is highly motivated to improve its research standing. The initiatives described in the following sections of this paper are intended to help the university achieve its goals.

Over the past decade Monash University Library has experimented with and deployed new technologies in support of the university’s objectives. This has included early leadership in the adoption of a web presence for the university, extensive development of electronic resources, and initiatives such as the electronic readings and reserve service, Monash University Lectures Online and the creation of a past examination papers database. The Library, in collaboration with the Information Technology Services division and the Monash University community, is well positioned to explore innovative approaches to scholarly communication.

Monash University ePress

Several years ago, in response to interest from the university’s academic community, the Library and the Information Technology Services division started experimenting with electronic publishing,¹⁴ and in 1993 the university agreed to fund a prototype electronic press. Advised by a board and based in the Library, the Monash University ePress will be launched in March 2005 with three electronic journals and a monograph, with other publications in production.¹⁵ The business model for the ePress is based on partial cost recovery, derived in the main from subscription and pay per view charging. The intellectual property contracts will protect the rights of the authors and allow them to make reasonable use of their material. After a period of time, Monash material published by the ePress will be made available on an open access basis.

The aims of the ePress are to:

- provide a more direct link between researchers and readers
- use innovative technology to publish scholarly material
- promote Monash University’s research activities and intellectual capital
- provide a sustainable publishing model
- provide a leadership role within the university, and

¹³ *Monash University: Excellence and Diversity: Strategic Framework - 2004–2008*. Clayton, Monash University, 2004.

¹⁴ See, for example, *People and Place*. [online] <http://elecpress.monash.edu.au/pnp/> (Accessed: 8 June 2004).

¹⁵ See www.epress.monash.edu.au

- build partnerships with others

The ePress's policies and procedures are intended to ensure that only high quality material is published, and all new material published will attract Department of Education Science and Training (DEST) research credits. EPress features will include: incorporation of colour images; reference linking,¹⁶ content automatically included in indexing and abstracting services; tools for managing the submission and refereeing processes, leading to reductions in publication lead times; usage statistics; flexible subscription and pay-per-view offerings; sophisticated sales and marketing functionality; and, repackaging of content.

ARROW

The Library and the Information Technology Service division's involvement in the ePress and other technology initiatives, and discussions about the need to improve and better integrate the university's management of its information resources, led to a grant of \$3.66 million from DEST under the Commonwealth Government's Backing Australia's Ability initiative.

This initiative, known as ARROW (Australian Research Repositories Online to the World), is a three-year consortium project for the establishment of institutional repositories and associated resource discovery mechanisms. Monash University is the lead institution, and the other partner members are the University of New South Wales, Swinburne University of Technology and the National Library of Australia. It is intended that, once the technology infrastructure and workflows are in place, other universities will be invited to participate.

ARROW repositories will include a wide range of digital formats, including pre- and post publication journal articles and conference proceedings, electronic versions of higher degree theses, and digital collections (such as image databases). The ARROW solution is also intended to support self-published open-access journals. If successful, it is anticipated that these repositories will evolve to include research data and teaching and learning objects, supported by associated authentication and rights management software.

ARROW's objectives are to:

- create a service framework that achieves national relevance by using interoperable solutions that will have wide applicability, and, by working with national and international standards, projects and initiatives
- gain high levels of stakeholder engagement with the project's aims, activities and developments
- establish affordable, sustainable and cost-effective solutions supported by institutional commitment and appropriate business models

¹⁶ References (citations) linked to the full-text versions wherever possible, allowing readers to jump directly from the citation to its full-text version. ePress's participation in [CrossRef](http://www.CrossRef.org) (See [online] <http://www.CrossRef.org>) (Accessed: 8 June 2004) and its use of Digital Object Identifiers (DOIs) will drive this functionality. For more information on reference linking see Amy Brand and Kristen Fisher, 'Linking Evolved: the Future of Online Research', *Research Information*, no. 5, Spring 2003: 12–14.

- implement hardware and software that is, in its elements and the way it is put together: scalable, modular, interoperable, open-standards based, referencing a common services layer, supporting value-added services, implementing digital rights management and taking into account archiving and preservation issues
- achieve critical mass of content
- enhance access to and increase the use of the scholarly research output of participating institutions
- improve scholarly communication by encouraging open or low-cost access to research output in a copyright environment that protects the individual's intellectual property and enhances the impact of their research.

The biggest challenge facing any institutional repository appears to be that of achieving a critical mass of content; that is, determining whether or not the concept itself will be of sufficient interest to the host institutions for its services to be taken up widely. The experience of almost all early initiators of repositories is that the idea is endorsed, but that in itself does not create the behavioural change necessary to get researchers to lodge their material in the repository. Through the ARROW project Monash University has a unique opportunity to create a repository that manages access to, and dissemination and preservation of, its research output in ways that enhance the university's research standing and impact. This paper outlines a number of strategies designed to achieve this.

a) Research directory

The ARROW project is designed to link to the research directory that is compiled annually to submit to DEST for research credits. The ARROW team aims to use this directory to capture articles as they are reported annually, gradually building a comprehensive source of information about Monash University research publications. The Library will manage the permissions process. Links from the directory will be created to the ARROW repository, to the Library catalogue and to other sources of information about the author. The directory will also be a useful tool for communicating with Monash authors about best practice intellectual property management.

A key to the success of this strategy is the overall University commitment to embed ARROW into its workflows and information management framework.

b) Memoranda of understanding

The Library has signed memoranda of understanding with the Faculties of Business and Economics and Information Technology to form the basis of support from the faculty to assist with the location and retrieval of research papers (both published and grey literature) for the repository. This strategy is most advanced with the Faculty of Business and Economics, which is supporting the project by allocating a staff member to gather and lodge working papers in the repository, and which has changed its workflows to accommodate the repository.

c) Academic promotion procedures

Much of the apparent inertia on the part of researchers about depositing research in repositories appears to be the result of a lack of immediate incentive. Monash University intends to encourage academics to lodge their

papers in the repository when they are seeking promotion. This would be counted as evidence of professional service, one of the four areas of activity assessed during the promotions process. In this way the prompt to take action would be incorporated in an existing procedure. At least in the first few years of establishing the repository the Library will seek the necessary copyright permissions and will manage the submission process for academics.

d) Theses

Monash University, through the Monash Research Graduate School has agreed to mandate electronic deposit of accepted theses with the Library from 1 January 2005. Electronic copies of doctoral theses will be stored in the Monash ARROW repository, and will be exposed for searching to abstract level. This is to gain the benefit of promoting the theses without the problems of third party copyright clearance (which is more stringent for electronic publishing) and issues about potential commercialisation of theses. The full text of theses will be available by inter library loan, as applies at present. The Library will include multiple formats in the repository.

The ARROW project has close links with the Australian Digital Theses project, which harvests metadata for preservation and resource discovery purposes. The ARROW project will incorporate the software necessary to support deposit of theses into the repository. Over 600 recent theses have been retrospectively digitised for inclusion in the repository, and it is likely that this will be expanded.

e) Individuals

In addition to the more significant strategies outlined in a) to d) above, the repository is promoted to individuals, and efforts to obtain research papers continue. Key individuals in the university, such as the Deputy Vice Chancellor Research, are asked to champion the repository, and to demonstrate this by lodging their papers in it.

Mandating deposit of research papers

There has been some debate about the issue of mandating deposit of research publications into institutional repositories.¹⁷ On one end of the continuum, Queensland University of Technology has received international attention for being the first higher education institution globally to mandate that all research publications be lodged in its repository.¹⁸ Clifford Lynch, on the other hand, argues that administrators (including librarians) may restrict the creative potential of institutional repositories by burdening them with too many policy and administrative constraints.¹⁹ The strategies outlined in this paper, which can be seen as a compromise between these two positions, are informed by a key recently released research paper into

¹⁷ Note that some categories of material would be exempt, such as material likely to have commercial application, or culturally sensitive material.

¹⁸ See [online] http://www.qut.edu.au/admin/mopp/F/F_01_03.html (Accessed: 8 June 2004).

¹⁹ Clifford A. Lynch, op. cit.

institutional repositories.²⁰ This authoritative and well-researched report says that the take-up rates of institutional repositories have to date been very patchy, and it concludes that this is the biggest problem facing those setting up these repositories.

To quote:

Outside a few disciplines (e.g. physics, computer science and economics) there is little tradition of preprints or working papers and apparently still little interest in self archiving. Academics may be radical in their thought but they are conservative in their behaviour, and there is a great deal of inertia in the current publishing systems.

The ARROW grant gives Monash University the opportunity to commit resources to advocacy and publicity, as well as resolving technical, workflow and sustainability issues. This paper argues that this should be supported by strategies that attempt to create cultural change. A richly populated repository will demonstrate its value within two or three years, after which self deposit will be more likely.

The approach being adopted at Monash University is based on a belief that institutional repositories will only be sustainable if they add value in a systemic way. To this end, the ARROW repository will be an integral part of the annual process of gathering and reporting on research output. The landscape will change dramatically if and when government or other incentives for depositing research output into institutional repositories are brought into play. This could occur, for example, through mandating that publications arising from Commonwealth Government grants be deposited, or by increasing points allocated to publications if they have been deposited. Before this can be done significant policy, copyright, resourcing and administrative issues would need to have been resolved.

A further development is that Monash University, on behalf of the ARROW project, is a member of the committee advising Thomson ISI on the development of its new product *Web Citation Index*, which will examine citations to and from material held in institutional repositories. ARROW's involvement in this project means that information about the impact of the repository will become available.

Beyond this phase

The initiatives described in this paper (Monash University ePress and the ARROW project) are designed as building blocks for future development. In particular, there is a need to consider management of converging e-learning and e-research requirements and output, with a view to ensuring integration and interoperability of these key university activities, and as a way of underpinning the teaching research nexus. The intention is to develop standard information management techniques and underlying infrastructure to support multiple data types (including e-learning resources, digital library materials, research output and research resources, such as datasets).

²⁰ Mark Ware Consulting Ltd, *Publisher and Library Learning Solutions (PALS) Pathfinder Research on Web-based Repositories: Final Report*, 2004. [online] <http://www.palsgroup.org.uk/palsweb/palsweb.nsf/> (Accessed: 8 June 2004).

Conclusion

With the decision to fund the establishment of an electronic press, and the subsequent Commonwealth Government grant to create an institutional repository, Monash University currently has a unique opportunity to transform its approach to scholarly communication, in ways that enhance its research standing and improve management of its research information. This paper concludes by quoting Clifford Lynch; he is specifically referring to the importance of institutional repositories, but the comments are equally relevant to electronic publishing.

At the most basic and fundamental level, an institutional repository is a recognition that the intellectual life and scholarship of our universities will increasingly be represented, documented, and shared in digital form, and that a primary responsibility of our universities is to exercise stewardship over these riches: both to make them available and to preserve them. An institutional repository is the means by which our universities will address this responsibility both to the members of their communities and to the public. It is a new channel for structuring the university's contribution to the broader world, and as such invites policy and cultural reassessment of this relationship.²¹

Cathrine Harboe-Ree
Monash University Librarian
4 March 2005

²¹ Clifford A. Lynch, *op.cit.*

MANAGING AUSTRALIAN RESEARCH OUTPUT FOR INCREASED RETURN ON INVESTMENT: THE ROLE OF OPEN ACCESS INSTITUTIONAL REPOSITORIES

DISCUSSION PAPER

Purpose

The Australian Government has identified research as one of its strategic policy priorities in order to improve Australia's competitiveness. The government invests heavily in research and research infrastructure through recurrent funding and grants. Currently support for research is being supplemented through *Backing Australia's Ability – Building Our Future through Science and Technology*, a five-year \$3 billion program.

Robin Batterham, the Chief Scientist, has highlighted the importance of “[creating] environments in which the best ideas are identified quickly and easily, to promote our contribution to the global knowledge pool and to respond to business and community needs.” He has also said that “the cycle of innovation must be fed by new ideas and basic knowledge which are capable of being transferred and accepted by end users.”²²

This paper seeks to demonstrate the importance of open access²³ institutional repositories²⁴ in supporting the government's objectives, thereby optimising its return on investment and creating an innovative environment. Open access institutional repositories can quickly and efficiently disseminate Australian research results in an environment increasingly dominated by the world wide web.

This paper also scopes the issues associated with the establishment of open access institutional repositories in Australian higher education institutions at this time.

Environmental scan

Batterham describes the advantages of implementing an investment strategy based on people and culture, ideas and commercialisation²⁵. The primary ways that research ideas are communicated are through collaboration, publication and presentation. Institutional repositories facilitate these activities by using technology to capture, store, expose and promote research output.

There is an increasing body of bibliometric analysis providing evidence that open access articles²⁶ have a significant impact on research accessibility. Four examples follow.

²² Robin Batterham, *The chance to change: final report by the Chief Scientist*. Canberra, Commonwealth of Australia, Canberra, November 2002.

²³ In this paper the term open access means “free and unrestricted online availability”, which is the definition used by the Budapest Open Access Initiative. See [online] <http://www.soros.org/openaccess/> (accessed 8 June 2004).

²⁴ The term institutional repository is defined on page 5 of this paper.

²⁵ *Ibid.*

²⁶ One of the few studies examining journals, as opposed to articles, *The impact of open access journals: a citation study from Thomson ISI*, 2004, argues that there is insufficient evidence to date that citation rates will increase for open access journals, although the article acknowledges that the potential is there to garner greater readership.

- a. *Astrophysical Journal* articles that are also on the pre-print server have a citation rate twice that of papers not on the pre-print server.²⁷
- b. Steve Lawrence from the NEC Research Institute at Princeton has undertaken studies that demonstrate that “there is a clear correlation between the number of times an article is cited and the probability that the article is [free] online. The mean number of citations to offline articles is 2.74, and the mean number of citations to online articles is 7.03, an increase of 157%.”²⁸
- c. A more recent study by Kristen Antelman from North Carolina State University Libraries confirms Lawrence’s result and claims that “the data show a significant difference in the mean citation rates of open-access articles and those that are not freely available online. The relative increase in citations for open-access articles ranged from a low of 45% in philosophy to 51% in electrical and electrical engineering, 86% in political science, and 91% in mathematics.”²⁹
- d. Steve Harnad and Tim Brody³⁰, while observing that “access is not a sufficient condition for citation, but a necessary one”, cite a number of studies that reveal dramatic citation advantages for open access (for example, a study of physics articles published each year between 1992 and 2001 reveals variation on an annual basis of between 2.5 to 5.8 times more citations for open access articles compared to closed access articles³¹).

In parallel with this growth in evidence that open access makes a significant impact on citation rates, the volume of information available online and its relative ease of access is dramatically changing researcher behaviour. The fact that search engine providers are now beginning to recognise the importance of the space means that these trends are likely to accelerate.

Lawrence points out that “the volume of scientific literature far exceeds the ability of scientists to identify and use all relevant information. The ability to locate relevant research quickly will dramatically improve communication and scientific progress.”³² He goes on to say that “to maximize impact, minimize redundancy and speed scientific progress, authors and publishers should aim to make research easy to access.”

Antelman observes that “as more research is available online, readers lower the threshold of effort they are willing to expend to retrieve documents that present any barriers to access.

²⁷ Sarah Stevens-Rayburn, *ASTRO: report from the AAS Publications Board meeting*, November 2003. [online] <http://listserv.nd.edu/cgi-bin/wa?A2=ind0311&L=pamnet&F=&S=&P=1632> (Accessed 18 January 2005).

²⁸ Steve Lawrence, ‘Free online availability substantially increases a paper’s impact’, *Nature*, May 2001: 411, 521.

²⁹ Kristen Antelman, ‘Do open-access articles have a greater research impact?’, *College & research libraries*, 65, no. 5: 372-382. [Online] <http://eprints.rclis.org/archive/00002309/> (Accessed 18 January 2005).

³⁰ Steve Harnad, and T Brody, ‘Comparing the impact of open access (OA) vs. non-OA articles in the same journals’, *D-Lib Magazine*, 10, no. 6, June 2004. [online] <http://www.dlib.org/dlib/june04/harnad/06harnad.html> (Accessed: 10 December 2004).

³¹ T. Brody, H. Stamerjohanns, S. Harnad, Y. Gingras, F. Vallieres and C. Oppenheim, ‘The effect of open access on citation impact’, presented at *National policies on open access (OA) provision for university research output: an international meeting*. Southampton University, Southampton UK, 19 February 2004. [online] <http://opcit.eprints.org/feb19prog.html>. (Accessed: 10 December 2004).

³² Lawrence, *op cit*.

This indicates both a 'push' away from print and a pull towards open access, which may strengthen the association between open access and research impact."³³

In a recent survey commissioned by Elsevier, scientists were asked (unprompted) to name the top scientific and medical resources they use or are aware of. By a large margin the highest score went to Google (Google 39%, Yahoo 10%, PubMed 9%, Science Direct 2%, ISI Web of Science 2%, Medline 2%).³⁴

Google has recently launched Google Scholar,³⁵ through which links to academic and scholarly material are being established, and they have an agreement with major libraries to convert their out-of-copyright holdings into digital files.³⁶ The libraries currently involved include Oxford, Stanford, Harvard and the New York Public Library.

What we see, then, is an huge swing (indeed, an avalanche) towards online activity, both for publishing (in the broadest sense of the word) and consuming of information, with print being a necessary but increasingly historic source of information, and with licensed material likely to be utilised only after easier to obtain online information does not yield the desired (or good enough) information. In fact, Google searching currently favours the open-access version; if an open access copy of an article is available, and has been indexed by Google, it will appear ahead of the restricted copy or copies in search results display.

What then of impact and quality in an increasingly open access environment? Antelman³⁷ states that "citations cannot in themselves be said to measure research impact, nevertheless, citedness as measured by ISI is a measure that is commonly relied on as a surrogate for such impact." She also notes that "citedness ... is itself viewed by scholars as an objective."

Quality is often tied to peer reviewing and the ranking of a journal, but these measures have known flaws, which are becoming more of an issue in an era where the definition of publishing is changing, where pre-prints are appearing more regularly as citations, and where open access exposure is attracting a wider range of readers. Antelman draws attention to a study performed using the Citebase data, which found that the more often a paper is downloaded, the more likely it is to be cited.³⁸ Antelman also says that "open-access articles published in lower-impact journals may ... have a greater relative research impact because [the journals] are not so widely available to authors through personal and institutional subscriptions."³⁹

In the absence of more transparent or easily identifiable quality measures, it is reasonable to argue that citation rate should be a measure of accessibility and of quality. The British Research Assessment Exercise (RAE) rates publications in high-impact journals as a measure of quality, although it should be noted, as described by the studies above, that there is a divergence between individual scholars' citation rates, and those of journals. Thomson ISI provides information about both journal impact factors and individual citations, however only the individual citation information provides an accurate measure of individual performance and a basis for measuring changing accessibility in an open access environment.

³³ Antelman, *op cit.*

³⁴ *Library Connect* (Elsevier newsletter), October 2004.

³⁵ *Google Scholar Beta launched November 2004* [online] <http://scholar.google.com> (Accessed 18 January 2005).

³⁶ *Google checks out library books* [online] http://www.google.com/intl/en/press/pressrel/print_library.html (Accessed 18 January 2005)..

³⁷ Antelman, *op cit.*

³⁸ In Antelman, *op cit.* Correlation Generator available online at

<http://citebase.eprints.org/analysis/correlation/php>

³⁹ Antelman, *op cit.*

In response to this, Thomson ISI is currently establishing a new indexing tool, *Web Citation Index*, which will report citations to and from institutional repositories and other web resources. Seven institutions active in the establishment of institutional repositories form an advisory group to this project. These are the Australian National University and Monash University from Australia (because of their leadership in the current government funded demonstrator projects), NASA Langley, Cornell University, Rochester University, CalTech and the Max Planck Society. The *Web Citation Index* will assess individual citation performance, including citations of pre-print material.

The debate about the importance of the link between quality and peer reviewing is likely to continue for some time, and to be monitored closely as pre-prints and other previously unpublished materials appear on citation indexes. Some disciplines, such as physics and business, have a culture of using material that has not been peer reviewed, whereas other disciplines continue to stress the need to retain peer reviewing. In the context of this paper, the relevant point is that this is a time of great change, during which much will be tested and debated. It is important that Australian research is included in the mix of material being exposed and evaluated in these ways.

Publication and citation rating are key factors in determining position on three recent university ranking systems:

- a. The Times Higher Education Supplement World Ranking gives 20% weighting to citations per faculty member.
- b. The Shanghai University Jiaotong Academic Ranking of World Universities gives 20% weighting to the number of highly cited researchers in 21 broad subject categories in life sciences, medicine, physical sciences, engineering and social sciences, 20% weighting to articles in Thomson ISI's *Science Citation Index* and *Social Science Index* and 20% weighting to articles published in *Nature* and *Science*.
- c. The Melbourne Institute Index of International Standing of Australian Universities, released on 24 November 2004, gives 40% weighting to quality and international standing of academic staff, based on research output and citations, and other factors.

Quality and Accessibility Framework

The Australian Government is in the process of establishing Quality and Accessibility Frameworks for Publicly Funded Research⁴⁰. The Government states that the Frameworks should:

- be transparent to government and taxpayers so that they are better informed about the results of the public investment in research;
- ensure that all publicly funded research agencies and research providers are encouraged to focus on the quality and relevance of their research; and
- avoid a high cost of implementation and imposing a high administration burden on research providers.

This paper argues that the best way to inform the nation and the international community about the results of public investment in research is to make research outcomes as accessible as possible, and that the best way to do this is to expose research output through open access repositories. This will lead to a greater assertion of Australian research contribution through

⁴⁰ *Research Quality Framework Consultation Discussion Starter* [online] http://www.dest.gov.au/resqual/documents/rqf_disc_paper.pdf (Accessed 18 January 2005)

increased citation, which in turn will affect the global ranking of Australian universities and research centres.

The argument in favour of making research outcomes more accessible has also been articulated on behalf of the government by Evan Arthur, who has said that “in the vast majority of cases, value from research accrues through effective dissemination of the results of research, leading to advancement of understanding and societal/industry uptake of results. Government ... has an interest in facilitating the effective dissemination of knowledge.”⁴¹

Establishing institutional repositories: background

For the purposes of this paper an institutional repository is defined as the software, hardware and services required to accept, store, make available online and manage a wide range of digital content, including the research output categories of pre- and post-publication journal articles, working papers, conference papers and higher degree theses, on an institutional basis – a university in most cases. Institutional repositories are also important for management of other content, including interactive models and data sets, research resources (such as digital image collections) and learning objects. Repositories will provide much greater functionality and support for collaboration, as well as exposing research activity in ways that have not been possible before, further increasing the return on investment in publicly funded research. These wider uses, while important, are not the focus of this paper.

Campbell, Blinco and Mason⁴² address a range of issues to do with the management and sustainability of repositories in various domains, including e-sciences, grid computing and e-learning. They point out the increasing occurrence of cross-domain communication, and the accompanying need for interoperability and integration between distributed systems and services within a domain also applies across domains. Government funded projects that are developing tools and services to support inter- and intra-institutional and cross-domain communication are detailed later in this paper. These tools and services will facilitate greater functionality for researchers at their desktop, including the ability get access to information from a variety of sources seamlessly, exchange information collaboratively with colleagues and manage and disseminate the results of their research through institution supported repositories.

A 2003 Council of Australian University Librarians (CAUL) survey on institutional repositories⁴³ identified six universities that had established e-print repositories; these are basic repositories that can manage pre- and post-publication journal articles, conference proceedings and working papers. A further fourteen universities were considering establishing a repository (it can be assumed that some of these will have been established in 2004), and ten responded that they had no plans for a repository. It is probable that none of the eight universities that did not respond to the survey is currently planning a repository.

Twenty-five universities participating in the Australian Digital Theses (ADT) project have started to store higher degree theses digitally and make them available online. The software currently being used for this task is different to the software being used for the e-prints

⁴¹ Arthur, Evan, ‘Government policy frameworks’, presented at *Changing research practices in the digital information and communication environment*, a National Scholarly Communication Forum roundtable. Canberra, June 2004. [online] [http://www.humanities.org.au/NSCF/PowerPoints/NSCF%20\(Arthur\).ppt](http://www.humanities.org.au/NSCF/PowerPoints/NSCF%20(Arthur).ppt) (Accessed 18 January 2005).

⁴² Campbell, Lorna, Blinco, Kerry and Mason, Jon, *Repository management and implementation: a white paper for alt-i-lab 2004*. Prepared on behalf of DEST (Australia) and JISC-CETIS (UK) [online] http://www.educationau.edu.au/papers/repositories_Altilab.pdf (Accessed 20 January 2004).

⁴³ [Online] <http://www.anu.edu.au/caul/surveys/eprint-repositories2003.xls> (Accessed 18 January 2005).

activity described in the preceding paragraph, although some of the institutions that have already established a repository are able to manage both of these categories of material in the one repository. (The government is currently funding the development of a new ADT metadata repository to improve access to Australian theses).

Benefits of institutional repositories

Clifford Lynch, Executive Director of the Coalition for Networked Information and one of the key spokespeople on the digital revolution in scholarly communication, expresses the view that the urgent task at the present time is to develop institutions, and cultures within them, that will nurture and support the exploitation of authorship in the new digital medium. It is worth quoting Lynch on the importance of institutional repositories. He says that “at the most basic and fundamental level, an institutional repository is a recognition that the intellectual life and scholarship of our universities will increasingly be represented, documented, and shared in digital form, and that a primary responsibility of our universities is to exercise stewardship over these riches: both to make them available and to preserve them. An institutional repository is the means by which our universities will address this responsibility both to members of their communities and to the public. It is a new channel for structuring the university’s contribution to the broader world, and as such invites policy and cultural reassessment of this relationship.”⁴⁴

In summary, institutional repositories add value to Australian research and increase its impact by:

- making it more accessible;
- promoting it through greater exposure;
- providing a durable location;
- reducing the time between authorship and consumption;
- supporting collaboration;
- providing the basis for improved university ranking; and
- encouraging innovation in, and refinement of, accessibility and quality measures.

They also have the potential to improve organisational efficiency by utilising the one mechanism to improve research accessibility, rationalise information management within each institution and provide research assessment data to the government and other funding agencies.

Establishing institutional repositories: issues for consideration

The rest of this paper concentrates primarily on the factors that need to be taken into account for the establishment of institutional repositories to manage research publications, especially those that are reported to the Australian government annually (although many of the issues are common to institutional repositories regardless of the content included in them). In order to manage research publications, many elements need to be in place. These include appropriate hardware and software (the technology), supporting workflows, policy and regulatory frameworks and administrative arrangements, and resources, especially staff resources. In addition, there are copyright and other legal considerations, and technical standards issues, and the issue of sustainable repository management services needs to be considered.

⁴⁴ Clifford A. Lynch, ‘Institutional repositories: essential infrastructure for scholarship in the digital age’, *ARL Bimonthly Report* 226, February 2003. [online] <http://www.arl.org/newsltr/226/ir.html> (Accessed: 8 June 2004).

Technology: hardware

This is not a major issue. While most institutions would need to acquire a server to manage a repository, these are not particularly expensive (circa \$10-15,000), and it is assumed that this cost could be absorbed by institutions.

Technology: software

This is a major issue. Most of the institutions in Australia currently managing a repository are using eprints.org⁴⁵ open source software. One or two are using MIT's open source DSpace software.⁴⁶ It has taken each institution already running a repository a significant amount of time to establish the software and develop the expertise necessary to manage it. There are a number of other possible software solutions available at present, including the above open source solutions and some proprietary content management systems. Eprints.org software has been a useful starting point, but it has limited functionality and no support or training base. DSpace has greater functionality, but it has not had any support or training base in Australia until recently. The work being done by the Australian Partnership for Sustainable Repositories (APSR) and Australian Research Repositories Online to the World (ARROW) projects is discussed below.

Most proprietary content management systems are not designed for this purpose, may be quite expensive and are not generally OAI compliant.⁴⁷ OAI compliance is important in allowing harvesting of metadata from multiple repositories for indexing in research specific search engines.

Government funded demonstrator projects are currently exploring a range of issues to do with the establishment and maintenance of institutional repositories, as follows:

The ARROW project is specifically developing software, based on the open source FEDORA software, to manage the full range of universities' research outputs, to support the reports submitted annually to the Department of Education, Science and Training (DEST), and to provide easier access to the material. The software will support capture of content and its indexing in a research directory or directories, present a web version of the full directory, link from the directory level to full text (where possible – see also the section following on copyright) and support once-only creation of metadata and other information needed to report correctly and manage the directory, the repository and a resource discovery service. The ARROW project is building a sustainable solution by involving a library systems supplier to provide software development, installation, maintenance and training for the ARROW repositories.

The ARROW software will be developed by early 2005 and tested and improved during 2005. A licensing and support strategy for any, or all, Australian universities has been negotiated with the supplier. The estimated total cost for this would be dependent on the number of institutions that wish to participate.

The APSR project is focusing on the critical issues of the continuity of access to, and sustainability of, digital information collections. It is also building on a base of demonstrator systems creating interoperability amongst discrete projects, embedded in development repository facilities within partner institutions, and it is contributing to national strength in these areas by encouraging the development of skills and expertise and providing coordination throughout the sector. Progress has been made in developing DSpace

⁴⁵ Refer <http://www.eprints.org/>

⁴⁶ Refer <http://www.dspace.org/>

⁴⁷ Software is Open Archives Initiative (OAI) compliant when the metadata can be harvested from repositories running the software using the OAI metadata harvesting protocol. This is an essential standard for facilitating discovery of open access research.

repositories, scoping common protocols for sustainability, accessing software packages to support an integrated gateway for repositories.⁴⁸ The Australian National University, through the APSR project, has recently released an out-of-the-box version of DSpace, which the University of Sydney is successfully using.

The Meta Access Management System (MAMS) project is developing the essential software elements to underpin the exchange of information within and between universities. This includes considerations of advocacy, and consideration of international standards and intellectual property which would enable this to work.⁴⁹ and the development of middleware systems for access and identity management (intra- and inter-institutional authentication and authorisation). Both the ARROW and APSR projects are intending to incorporate MAMS middleware solutions as their repository access protocols.

Taking all of the above into account, most Australian universities should have the capacity to establish a basic repository using eprints.org software now, generally using the resources of the Library and the IT department, however this is an optimistic assessment that probably does not fully take into account the frustration of trying to use under-developed software with inadequate functionality. The APSR and ARROW solutions either are, or will soon be, ready for trialling in other institutions. The support provided by these projects at least during the establishment phase will be essential. Support provided beyond the term of the current projects is discussed later in this paper.

Metadata, standards and persistent identifiers

There are numerous issues relating to software that may act as barriers to the effective management of repositories, but which will not be resolved in the short term. These include developing and agreeing on standards, the current early state of work towards a coherent approach to metadata in higher education, and the need for a cost-effective and sustainable system of persistent identifiers. None of these issues will prevent institutions from establishing a repository, but there may be a role for the Australian Government in ensuring these are addressed or resolved from a national perspective. The current projects (MAMS, ARROW and APSR) are exploring these issues, and will be able to inform future directions.

Regulatory, administrative and legal environment

There are several issues that need to be considered here. A communication strategy, including consultation, addressing the needs of the Australian Vice Chancellors Committee, vice chancellors, deputy vice chancellors research and research management units will be necessary to address the concerns likely to be expressed about these kinds of issues, and achieve necessary policy changes and cultural acceptance. Some of the specific issues that need to be addressed follow.

- **Incentives**

There will be no significant development of open access repositories without the provision of incentives to encourage deposit of research outputs. The processes of intra-institutional communication, advocacy and policy development will not of themselves fundamentally change behaviours. This is a key issue for the government's consideration. The current system, with its strong links to promotion and recognition for individuals and financial rewards to institutions, is firmly entrenched.⁵⁰

⁴⁸ *Progress Report: Australian Research Information Infrastructure Committee (ARIIC) Activities 2004-2004*, an ARIIC Committee meeting paper, 17 December 2004.

⁴⁹ *Ibid.*

⁵⁰ This issue is discussed in more detail in my paper 'Transforming scholarly communication: a Monash University perspective', 10 June 2004, which is available separately.

- **Copyright**

Experience to date as institutions attempt to establish repositories suggests that copyright is one of the biggest problems, and guidance will definitely be needed to support more general establishment of repositories. As the material being captured for the repositories has already been published (particularly the items submitted to DEST for the research publications data collection) permissions may have to be obtained from publishers before full text can be made available online. In the short time this material may be able to be suppressed, or exposed only to abstract level. A longer term strategy would see researchers being advised in advance to vary standard publisher contracts in order to allow deposit in open access repositories, but this will take some years to achieve.

The ARROW and APSR projects are producing documentation to assist in the management of this issue, including sample clauses for researchers to use in discussions with publishers, responses to frequently asked questions about copyright and pro formas for identifying copyright status at the time of capture of the material. Also to be addressed are embedded third party copyright materials and approvals for depositing multiple author works.

Institutions must consider all of these issues and be well informed about the level of risk attaching to any particular decision or approach. A national perspective would be helpful. The work that Queensland University of Technology is doing as part of Creative Commons is likely to be useful in this regard, noting that "Creative Commons aims to make copyright content more "active" by ensuring that content can be reutilised with a minimum of transactional effort."⁵¹

It needs to be accepted that, while DEST research publications directories may soon be web based, full open access to 100% of the material listed in them will not be feasible for the foreseeable future, because of copyright issues.

- **Work flow support**

It will be easier for institutions to establish repositories linked to research directories if work flows are supported through pro formas and guidelines. These would include the copyright clearance forms identified above, as well as forms designed to capture metadata and publishing details required by DEST and for management of the repository. It is important that the information required does not have to be sought or created twice.

A best practice approach to business planning for open access of institutional research output would support end-to-end work flow. This would start at the point of public funding of research, and move through generation of the article, capture, reporting, dissemination via an open access repository, and management of the information institutionally and at DEST for a range of purposes. The more steps in this chain that are supported by repository software, the regulatory framework and work flow processes, the greater the efficiency that will be achieved, and the better the exposure of Australian research output.

- **Resourcing**

All institutions currently establishing repositories have found that the process requires dedicated staff. At present the tasks are time consuming because they are experimental, and because many of the tools referred to in this paper are not yet fully in place. Much of what is being done through MAMS, ARROW and APSR will allow other institutions to save time. Nonetheless, all institutions will face an establishment phase requiring project management, and there are recurrent staffing implications. Both the APSR and ARROW

⁵¹ See [online] <http://www.creativecommons.org.au>. (Accessed 19 January 2005).

projects will support the establishment and operation of institutional repositories during the life of the projects. Beyond that, the Australian Government may consider it appropriate to support the establishment phase, either through a support team that moves from institution to institution helping with setup and training, or through institutional grants, or a combination of these, or some other means. Some national coordination to ensure consistent communication and effective dissemination of experience, processes, tools and documentation would be desirable.

Neil McLean has suggested a practical way forward, and a means of providing repository management services in a cost effective, sustainable manner, in a paper entitled *Australian Repository Management Service (ARMS)*⁵². He details a number of necessary services, including hosting tools and services, assistance to institutions in implementing repositories, developing and maintaining discrete modules of web services, collecting and managing appropriate open source software, negotiating software agreements, managing identifier and resolution services, maintaining registries of schemes and vocabularies, liaising with interested sectors and bodies, producing good practice guides and identifying areas for further research and development.

Planning: 2006 and beyond

From the above discussion it can be concluded that it should be feasible to achieve the following in time for implementation by the 2006 academic year.

- 1 Repository software at, or available to, all Australian universities, facilitating the online availability of a significant proportion of (but not all) published and unpublished Australian research outputs through the APSR and ARROW projects.
- 2 A mechanism for submission and compilation of information required for the annual DEST research publications data collection, through the ARROW project.
- 3 Prototyping of authentication and authorisation capability to support cross-domain and intra and inter institutional communication involving repositories, through the MAMS project.
- 4 A national research discovery service based on metadata harvested from OAI compliant repositories in Australian research institutions receiving public funding developed by the National Library as part of the ARROW project.

For sustainable solutions beyond 2006 Neil McLean's ARMS concept will need to be developed, or some other approach adopted.

Cathrine Harboe-Ree
Monash University Librarian
8 February 2005

⁵² *Progress Report: Australian Research Information Infrastructure Committee (ARIIC). op cit.*