2007

Outsourcing open access: Digital Commons at the University of Wollongong, Australia

Helen E. Mandl
University of Wollongong, helenma@uow.edu.au

Michael K. Organ
University of Wollongong, morgan@uow.edu.au

Publication Details
This article was originally published as Organ, M and Mandl, H, Outsourcing open access: Digital Commons at the University of Wollongong, Australia, OCLC Systems & Services - International Digital Library Perspectives, 23(4), 2007, 353-362. Original article available here.
Outsourcing open access: Digital Commons at the University of Wollongong, Australia

Abstract
In October 2005 the University of Wollongong gave approval for the implementation of an institutional repository utilising Proquest's Digital Commons software, the first such instance in Australia. The project sought, over a 2 year period, to make available online a significant portion of the university’s research output for the years 2000-5. It was envisaged that improved accessibility to journal articles and conference papers would assist in enhancing the research reputation of the university. This paper outlines why Digital Commons, re-branded as Research Online (ro.uow.edu.au), was chosen rather than an open source software solution. Issues arising from the utilisation of an outsourced, off-campus institutional repository system are discussed. This case study focuses on the University of Wollongong’s experience in regards to planning and implementation. It also addresses areas of collaboration, both external (nationally and internationally) and internal, and provides cogent examples of ongoing issues, lessons learnt and mistakes to avoid.

Disciplines
Arts and Humanities | Social and Behavioral Sciences

Publication Details
This article was originally published as Organ, M and Mandl, H, Outsourcing open access: Digital Commons at the University of Wollongong, Australia, OCLC Systems & Services - International Digital Library Perspectives, 23(4), 2007, 353-362. Original article available here.
Abstract

Purpose – The purpose of this article is to outline the experiences of an Australian university in selecting a proprietary solution for its open access digital repository requirements.

Design / methodology / approach – An overview is presented of the environment leading up to the decision to select Digital Commons over an open source software solution. The paper also outlines subsequent experiences during a one-year period in operating the outsourced solution.

Findings – Outsourcing is an appropriate digital repository option for higher education institutions when costs are considered and compared with open source solutions, and especially when on-site IT support is limited. Outsourcing allows local staff to concentrate on liaison with faculty in promoting and populating the repository.

Practical implications – A useful resource for those considering the use of proprietary or open source software for their institutional repository.

Originality / value – This paper deals with a little discussed area of the relatively new subject of open access institutional repositories.

Keywords Digital storage, Outsourcing

Paper type Case study
In October 2005 the University of Wollongong gave approval for the implementation of an institutional repository utilising Proquest’s Digital Commons software, the first such instance in Australia. The project sought, over a 2 year period, to make available online a significant portion of the university’s research output for the years 2000-5. It was envisaged that improved accessibility to journal articles and conference papers would assist in enhancing the research reputation of the university. This paper outlines why Digital Commons, re-branded as Research Online (ro.uow.edu.au), was chosen rather than an open source software solution. Issues arising from the utilisation of an outsourced, off-campus institutional repository system are discussed. This case study focuses on the University of Wollongong’s experience in regards to planning and implementation. It also addresses areas of collaboration, both external (nationally and internationally) and internal, and provides cogent examples of ongoing issues, lessons learnt and mistakes to avoid.

Introduction

Institutional repositories (IRs) are relatively new phenomena. Arriving with the change of millennia, they were “well on their way to becoming mainstream technology” by the middle of the decade (Swanepoel 2005). IRs allow organisations to promote their research outputs in new ways and provide, often for the first time, an element of control of that output without impinging upon copyright or intellectual property rights. They are essentially databases of electronic objects such as text and image files, the former comprising journal articles, conference papers, theses, book chapters, reports and the like. Combined, these represent the majority of the research output of a typical university or research centre. Specially designed software allows these digital objects to be archived and associated metadata and open access protocols enable them to be easily found using web search engines such as Google and Yahoo.

Institutional repositories can be a win-win for organisations and staff seeking wider research community exposure to material which has traditionally been locked away in print subscriptions or in password-protected online databases. This material is now made freely available to any researcher via the internet. The success of institutional repositories in broaching national borders is evidenced by the fact that in its first nine months of operation Digital Commons at the University of Wollongong was accessed by
researchers from 100 different specific country domains, or 100 different identifiable countries of origin (Research Online 2006).

Though discipline-based digital repositories of research material have been in use since the early 1990s (viz. the physics eprint server at arXiv.org), the move by individual research institutions and universities to implement local equivalents began to take shape following the sporadic development of eprint servers towards the end of the decade. In Australia, for example, a network of such servers was set up in 1998-9 to make research theses available. The Australian Digital Theses (ADT) project utilised Virginia Tech software and was quickly adopted by a majority of Australian universities. As of 2006 it had 37 members out of the total of 39 eligible institutions and the program was being expanded to New Zealand (Lynch 2005, ADT 2006).

Spurred on by the contemporaneous development and release of the open source ePrints software in 2000 and MIT and Hewlett Packard's DSpace at the end of 2002, and philosophically backed, if not driven, by the open access movement, the library sector naturally took to the role of promoters and implementers of institutional repositories (Mircea 2005, Allard et al. 2005). As a result, there has been a rush of activity on campuses around the world in recent years to put such facilities in place. As network speeds improve and storage capacities become almost limitless, the practicality and sustainability of IRs is enhanced, as is their rate of adoption.

A 2005 Coalition for Networked Information (CNI) survey into the deployment of institutional repositories in the United States and other countries found a great diversity in the type of object content, disciplinary coverage, and software used (van Westrienen and Lynch 2005). Of the US doctoral-granting institutions that replied to their survey, 40% had some type of repository in place, whilst the majority of the others were in various stages of planning. DSpace (58%) and Digital Commons (21%) were the main software solutions adopted in the US at the time of the survey. Since 2002 the value of open access institutional repositories has gained wide acceptance, and their uses within various environments is expanding (Lynch 2003, Cochrane 2006). These repositories are dealing with an ever increasing variety of digital content, ranging from traditional research publications through to teaching and learning materials and new media such as podcasts.

This chapter focuses predominantly on the experience of the University of Wollongong, Australia, in selecting and implementing an institutional repository to house academic publications.
The University of Wollongong

The University of Wollongong is a reasonably young institution. Established in 1975, its 2006 equivalent full-time student load was 13,347 and the academic staff population stood at 723. Over recent years the university has won awards for its teaching and been named Australian University of the Year three years in a row. Wollongong has gained an impressive reputation for its research intensity relative to size and in 2006 was ranked 196 on the Times Higher Education Supplement ranking of the top 200 universities world-wide (Times Higher Education Supplement 2006). Developing its research profile has been a key goal for the university with a range of strategies in place to promote the research it is undertaking, its outcomes and the academic staff involved.

It was observed at the end of 2004 that, despite increasing numbers of publications being authored by University of Wollongong academics, there were no centralised mechanisms or infrastructure in place to make this material easily available online. Some staff had personal home pages, while others with fewer technological skills had none, or at best a list of publications in a curriculum vitae format. Concerns over the complexities of copyright law hindered initiatives in this area. Meanwhile, the omnipresence of Google and other search engines as a starting point for information searching was a compelling argument for improving access to local research output. Students and researchers were voraciously seeking information, with ever increasing expectations of easy access to full-text material. External studies were also showing that the more accessible research is, the more it is cited (Hitchcock et al. 2003, Hajjem et al. 2006, Open Citation Project 2006). There quickly developed an understanding that the institutions leading the way with repositories of research output had an advantage in affecting citation rates over those without such a facility. Governments and funding bodies were also entering the fray, seeking greater accountability and improved access to research outputs in return for their investment. Renewed impetus for IRs in Australia came with the 2005 announcement that the government was investigating a new means of funding research, similar to the Research Assessment Exercise (RAE) in the United Kingdom. This is yet to be finalised, though a 2008 implementation date has been floated. Likewise, the issue of open access and institutional repositories has been much discussed in the US Congress over recent years.

The University of Wollongong had some initial success with the ADT program and implemented mandatory submission of digital theses in 2001. Managed by the University Library, ADT at Wollongong assisted staff in developing a basic understanding of issues
surrounding the operation of an online, open access digital repository before the idea of an institutional repository as we now know it had fully evolved. The move toward considering implementation of such a facility locally in 2005 was therefore a natural one. However, while senior executives of research institutions are attracted to the idea of digital repositories promising improved exposure of outputs and enhanced institutional and academic reputation, the decision to take the next step and allocate resources and recurrent funding is not easily taken. The reality of untested software, unproven budgetary frameworks and unknown cost benefit ratios has resulted in the tentative implementation of institutional repositories across the higher education sector, despite the encouragement of faculty, the cogent arguments of the open access movement and the practical support of librarians, archivists and IT specialists. Institutions are also faced with the challenges of implementing a broad range of limited-access repositories catering to content such as personnel and administrative records, organisational web sites, online teaching and learning materials and large primary data sets. The associated hardware and network development issues are also considerable.

Open Source vs. Proprietary Solutions

Once the decision to go down the open access institutional repository path is taken, the next big hurdle is the selection of software, and here the waters become muddy. A 2005 content analysis of the IR literature suggested a ‘one size fits all’ solution would be difficult to find and a range of approaches could be compared and contrasted (Allard et al. 2005). A number of commentators have pointed out that software is the least of one’s problems when it comes to setting up an institutional repository (Foster 2005, Gibbons 2006). The main hurdles are connected with implementation and affecting the necessary social and cultural change on campus to achieve acceptance and use by academic staff. Nevertheless, selection of the most appropriate software solution is important, if not of prime importance (Thomas et al. 2005).

A determining factor in selection can be cost, and once again appropriate funding models have been slow to develop. Early estimates put the cost of operating an institutional repository at around $200K per annum (ARL 2002, Barton and Walker 2002), though more recent studies have shown that the actual figures can lie in the range $10K to $2M (Houghton 2006). While initial hardware and software costs may be relatively small, the ongoing staffing costs to manage, develop and encourage users are more substantial.
Since 2001 the Australian Federal Government has funded a range of strategic infrastructure initiatives, including a number of repository projects (DEST 2006). A small group of universities have led investigations into different repository packages including ePrints, DSpace and Fedora. The outcomes of the projects have been well documented and distributed, however many could be classed as experimental and requiring a certain level of programming expertise to assist in implementation and ongoing development. In dealing with immature technology, projects locally and abroad have required time and staff support for technical work and developments with less emphasis on building content to show off the capabilities of the repository.

Library staff at the University of Wollongong watched with interest and waited for the arrival of a robust and relatively simple solution to this complex problem. DSpace was an early leader in this area. Proprietary software companies largely stood back to wait for a market to develop, though Proquest linked up with the University of Berkeley’s BePress in 2004 to develop a specific solution following on earlier collaboration in the area of digital theses (BePress 2006). By 2005 a number of products were on the market, including Digital Commons, Harvest Road Hive and The Learning Edge. DSpace supported the greatest number of installations and Digital Commons was the most popular off-the-shelf product in this rapidly evolving area of information technology (Nolan & Costanza 2006, Proquest 2006).

Planning the way forward

By the end of 2004 librarians at the University of Wollongong were aware of the open access movement and the desirable features of institutional repositories. Learning developers were also interested in exploring a repository solution, or learning content management system, for the storage of teaching and learning objects in a limited access environment. As a result, a campus-wide Content Management System (CMS) Evaluation Team was set up at the beginning of 2005, comprising members of faculty and staff from the IT, educational development and library units. It was tasked with engaging the university community in the merits or otherwise of a CMS. The Team subsequently carried out a campus-wide needs assessment and undertook a preliminary investigation into available software solutions and associated costs. A report was prepared in April 2005 following discussions with six Australian higher education institutions which then had in place some form of open access repository, or were well down the path of implementation. These comprised the Australian National University,
the University of Queensland, Curtin University of Technology, the University of Melbourne, Monash University and Queensland University of Technology. By October 2006 this number had swelled to 19 with active institutional repository programs, separate from those involved in ADT (Sale 2006).

While the work of the Team proceeded, external events gave added impetus to the implementation of an institutional repository locally. Through external networks, and in addition to the library’s own lobbying, University of Wollongong senior executive were made aware of the merits of IRs in promoting the research output of the institution. An “upload imperative” arose, with an expressed desire to implement a solution as soon as possible. DSpace was given serious consideration and became the initial frontrunner in the software selection process. A pilot version was installed locally in June 2005 and it proved of immense use in familiarising the Team with the many practical issues involved in IR setup, implementation and on-going maintenance and support. However the decision was made in October 2005 to go with the externally-housed Proquest Digital Commons. Outsourcing was adopted rather then proceeding down the path of open source software and internal development. A separate decision was also made around the same time by the university’s new Graduate School of Medicine to adopt The Learning Edge (TLE) learning content management system for teaching and learning digital objects.

In December 2005 the CMS Evaluation Team final report was completed. It concluded that the University of Wollongong needed to plan for its long-term institutional repository needs as teaching and research staff were increasingly working in an electronic environment (CMS Team 2005). This required the provision of infrastructure to enhance, capture, re-use and promote the intellectual capital of its community, alongside programs aimed at increasing the skills of general, academic and IT staff across campus. Finally, the report noted that the adoption of Digital Commons and The Learning Edge as short-term software solutions for research and teaching objects respectively would provide an opportunity for university staff and students to acquire knowledge in the operation and use of digital repositories. It was recognised that the rapidly evolving information technology and infrastructure environment would require the university to take on board new solutions and adapt to changes imposed by government and funding providers.

Why outsource?
The decision to go with Digital Commons, an outsourced, proprietary software solution as opposed to a locally housed, open source package such as DSpace or ePrints was taken for a variety of reasons. These included uncertainty with regard to local IT support for open source software, the desire to move as quickly as possible into the new area of IRs, and comparative cost. Digital Commons had the required features, and Proquest would provide training and support beyond setup.

The University of Wollongong decision reflected a similar one taken by the Florida State University when it adopted BePress software in 2002-3 to deal with theses (Thomas et al. 2005). In that instance, outsourcing was seen as a “more affordable alternative” to the open source solutions requiring substantial local IT support. Florida believed it would also be quicker to get up and running – 1 month for the outsourced solution as opposed to 1 year for internal development and full implementation using open source software.

In the process of selecting Digital Commons, University of Wollongong staff noted the developing state of the institutional repository environment globally and the fact that available software solutions were, in some respects, immature. It was accepted that any solution may be an interim one. The key point was to deploy a repository consistent with a range of standards so that material loaded could be transferred if necessary at a later date to a different system. The university was also keen not to be caught in a development loop but to focus efforts on increasing content and enhancing the impact of the institution’s research output. Digital Commons was an attractive package in that it was housed externally and supported centrally, therefore bringing down implementation costs related to the purchase of infrastructure and staffing. Digital Commons was ultimately assessed to be cheaper, faster to implement and simpler to manage.

Implementation @ Wollongong

The Proquest contract was signed in November 2005. A budget of approximately $130K per annum over two years was used to employ two library staff full-time to manage the project, populate the repository and promote it to faculty. The local version of Digital Commons – re-branded as Research Online (ro.uow.edu.au) – went live in December 2005, with templates specifically designed for Wollongong. The first papers were loaded just over a month later in January 2006.

An important part of the implementation process was deciding on the structure of the local Digital Commons instance, reflecting the organisational structure of the University
of Wollongong and its research framework. Most institutional repository solutions had adopted a system of communities or groups under which material could be stored. Using experience gained from the DSpace pilot project, the Wollongong implementation team opted for a simple structure, based on major faculty and administrative units. This would complement the visually sparse screens and improve useability. It was expected that the majority of people who accessed material in Research Online would do so externally via an initial search through a search engine such as Google. Only a small percentage of visitors to the site, including University of Wollongong staff, were likely to access it via the Research Online homepage and make use of the structure to locate material. For this reason the community structure was kept at the broadest faculty level, rather than mirroring the complexities of an ever-changing school, department and research centre structure.

Implementation proved relatively straightforward. The Digital Commons system was simple to understand and the backup support provided by Proquest and Digital Commons was fast and client focused. On-site training was provided along with instruction manuals. Questions sent by email to the support centre were responded to within 48 hours and a number of problems and small adjustments made quickly in answer to queries.

Living with an outsourced solution

As of the end of October 2006 some 1,060 papers had been uploaded to Research Online, generating over 54,000 hits to the site, including 32,000 full-text downloads of repository content and on average 2,000+ full-text downloads per week. Local stakeholders have, to date, deemed the facility a success in promoting University of Wollongong research output. The impact of the access provided by Research Online in regards to improving the citation rates of individual academics is a long term element yet to be quantified.

Procedures for the uploading of material have been developed to deal with the sticky issue of copyright and obtaining publisher permissions where necessary. The Sherpa web site (www.sherpa.ac.uk) has been useful at an international level, but Australian and small publishing houses have had to be contacted individually to obtain relevant permissions. Visits to faculty meetings and to individual academics by project staff has been an important tool in gaining acceptance of the role of the repository. The message given during such engagement is that Research Online would improve citation rates and
enhance the visibility of individual research. As part of the ongoing promotion of the product, good news stories and testimonials have been noted in university newsletters and frequent use is made of a range of statistics provided by the Digital Commons software. Individual academics have expressed appreciation of the monthly full-text download reports automatically generated by the system. They like the fact that researchers are actually downloading and reading their publications.

Since the University of Wollongong’s implementation, seven other Australian and New Zealand universities have adopted Digital Commons as their repository solution. This has allowed the development of a local Australian and New Zealand network for the discussion of issues and developments. User groups for ePrints, DSpace and Fedora already operate and meet in Australia. Research Online staff have also been active participants in the international Digital Commons Online Forum, where news of new sites and software enhancements are discussed and disseminated. In a similar vein the Digital Commons RSS feed facility allows individuals to monitor on a daily basis new input to the more than 50 sites which are currently live world-wide. Both the Forum and the RSS feed promote collaboration between local Digital Commons administrators who, upon seeing a feature which may have application to their own site, are able to follow this up with the Digital Commons support team, or with the individual site manager. This in some ways reflects the network which supports the ongoing development of open source software such as DSpace.

An active eye is kept on the repository projects underway in Australia with a view to taking advantage of the situation once the technology matures and is more stable. In the meantime, an active and growing repository exists in the form of Research Online, with benefits for academics and the university community.

The experience of implementing Digital Commons has, in general, been a positive one for the University of Wollongong, with few issues of concern. The platform is stable, and in the first 10 months of operation only one major outage was experienced, lasting approximately 3 days as a result of both off-site and local IT issues. The latter has also arisen in regards to the way in which the University of Wollongong proxy servers deal with an off-site database linked by a local URL, however this has not in any way affected off-campus access, which is the primary function of the database. Staff managing the Digital Commons software have found it eminently usable, with the upload process relatively refined and simple considering the complexity of similar databases. Response times from the central servers located in the United States are fast, both for administrative functions and user search processes.
One of the downsides of going with an external, centralised system is the inability to quickly implement desired changes to functionality and appearance. This has not been a significant issue at Wollongong as the Research Online team was aware of the available functionality at the time of purchase and considered this appropriate. The relative simplicity of the Digital Commons administrator and user interface was one of the features which influenced the decision to adopt the software, and this has proven an important factor in the ongoing process of selling the system to faculty. Individual academic submission of papers to Research Online has been successfully implemented, with few difficulties. A number of ‘trojan horse’ academic staff members have lead the way, with one senior member of faculty self-submitting approximately 50% of the 100 papers under her name loaded during the first 9 months of operation (ro.uow.edu.au/jseberry/).

The manner in which Digital Commons records are expeditiously made available to Google, and presented therein with clean, precise entries to enhance useability, is a plus and must be one of the primary selling points of the software. The ease with which the package also interacts with search engines and compilers such as Oaister, Google Scholar, OpenDoar, ARROW and Scirus has enhanced the visibility of University of Wollongong research on the internet and given rise to the positive download statistics cited above. Research Online’s acceptance for harvesting by the Web Citation Index within ISI’s Web of Knowledge points to the improving status of such repositories.

How is the success of Research Online to be assessed? This is an ongoing and developing process, however when the first two years of the repository project are reviewed at the end of 2007, measurable elements to be taken into account may include the amount of content available, interaction with major harvesters and search engines, take-up by academic staff and their feedback, impact on university reputation, visibility and impact on citation rates. For the Research Online staff, the Digital Commons software solution is a practical means of attaining such data as the University of Wollongong takes its first tentative steps into the world of institutional repositories.

References


http://adt.caul.edu.au/about/aimsoverview/

http://www.arl.org/IR_agenda.html


http://www.bepress.com/index.html

http://ir.uow.edu.au/team.html

Cochrane, T, ‘How can a Repository Contribute to University Success?’ Presentation to The Successful Repository conference, Brisbane, 29 June 2006. [Powerpoint presentation, available on web site, viewed 8 July 2006].


Proquest, ‘Digital Commons@Proquest’ [web site], accessed 25 October 2006.
http://digitalcommons.proquest.com/

Gibbons, S, ‘Making a Repository a Success with your Academic Staff’. Presentation to The Successful Repository conference, Brisbane, 29 June 2006. [Powerpoint presentation, available on web site, viewed 8 July 2006].


About the authors

Michael Organ is Project Manager, Research Online, at the University of Wollongong, Australia. He has a background in archives administration and has published on areas such as library space utilisation, internet access to archival resources and institutional
repository download statistics. Michael Organ is the corresponding author and can be contacted at morgan@uow.edu.au.

Helen Mandl is the Associate Librarian, Planning and Development at the University of Wollongong, Australia. She has a background in teaching and information literacy and is involved in a range of digitisation and repository projects supporting the research aims of the university.