



UNIVERSITY
OF WOLLONGONG
AUSTRALIA

University of Wollongong
Research Online

Deputy Vice-Chancellor (Academic) - Papers

Deputy Vice-Chancellor (Academic)

2010

Leveraging research quality assessment exercises to increase repository content - an Australian case study

Michael K. Organ

University of Wollongong, morgan@uow.edu.au

Publication Details

Micheal Organ, Leveraging research quality assessment exercises to increase repository content - an Australian case study, Paper presented at the CAIRSS (Council of Australian University Librarians Institutional Repository Support Service) Community Day, Royal Melbourne Institute of Technology University, Melbourne, 10 September 2010, 7p.

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au

Leveraging research quality assessment exercises to increase repository content - an Australian case study

Abstract

The legacy of Australia's national research quality assessment process – Excellence in Research for Australia (ERA) - at the University of Wollongong during 2010 was improved integration between the institutional repository and research management systems, and a move towards digitisation of the Higher Education Research Data Collection (HERDC) with consequent flow-on of metadata and digital objects to the institutional repository. Whilst ERA was a diversion from the task of securing open access content through faculty promotion and one-on-one contact with researchers, it nevertheless gave rise to a semi-automated process which promised improved rates of content acquisition.

Disciplines

Arts and Humanities | Social and Behavioral Sciences

Publication Details

Micheal Organ, Leveraging research quality assessment exercises to increase repository content - an Australian case study, Paper presented at the CAIRSS (Council of Australian University Librarians Institutional Repository Support Service) Community Day, Royal Melbourne Institute of Technology University, Melbourne, 10 September 2010, 7p.

Leveraging ERA to increase repository content

Michael Organ

Manager, Repository Services

Library, University of Wollongong

Content

Abstract

The long and winding road – RQF to ERA

The good, the bad and the ugly – HERDC and ERA

More than meets the eye – ERA transformed

It's a wonderful life – post ERA

Abstract: The legacy of Australia's national research quality assessment process – Excellence in Research for Australia (ERA) - at the University of Wollongong is improved integration between the institutional repository and research management systems, and a move towards digitisation of the Higher Education Research Data Collection (HERDC) process with consequent flow on of metadata and digital objects to the institutional repository. Whilst ERA was a diversion from the task of securing open access content through faculty promotion and one-on-one contact with researchers, it nevertheless gave rise to a semi-automated process which promises improved rates of content acquisition.

The long and winding road - RQF to ERA

Institutional submissions for the national Excellence in Research for Australia (ERA) quality assessment process closed on 31 July 2010 and substantially brought to a close a task which had occupied the minds of repository and research office managers since the Research Quality Framework (RQF) was first mooted in 2004.¹ Participation in both RQF and ERA was long and laborious, involving prioritised allocation of limited institutional repository resources, usually at the expense of core activities such as outreach and the pursuit of mandates (real or virtual) to achieve sustainable content supply. Promotion of open access nationally ground to a halt with the ERA green light early in 2008,² opening the door for publishers and copyright agencies to further tighten their grip on the research publication process.³ Repository staff were forced to focus on implementation of a set of complex, and often ambiguous, guidelines and technical specifications which began appearing from late 2008.

Throughout this period the RQF and ERA were both chore and challenge for repository personnel, addressed with a mixture of enthusiasm and dread. Initial excitement in 2005 - linked to the RQF's identification of open access repositories as key items of research management infrastructure - was dampened by its subsequent abandonment at the beginning of 2007 and replacement with a more

¹ The Australian Prime Minister John Howard announced in May 2004 his government's intention to develop 'Quality and Accessibility Frameworks' for publically funded research. At the end of that year he formed a 13 member Expert Advisory Group (EAG) to produce a Research Quality Framework (RQF) during 2005. Trials were subsequently held in 2006. The RQF was officially dumped by the Rudd Labor government in 2007 and replaced by ERA at the beginning of 2008. Though submissions closed at the end of July, repository managers were required to be "on call" through to the end of September if non-digitised material was required by the peer review panels.

² The international Open Access Day / Open Access Week initiative has had a very limited take up in Australia since its inception in 2008, with Charles Darwin University and Queensland University of Technology participating in 2008 and only the University of Wollongong Library during 2009 via the production of an Open Access guide (URL: <http://uow.libguides.com/openaccess>) and various on campus activities. For more information on global activities connected with this event refer the following URL: http://oad.simmons.edu/oadwiki/Events_celebrating_Open_Access_Day#2008.

³ One element of publishers addressing this issue had been the introduction of a "pay for open access" option whereby authors are charged a fee to make their papers freely available online upon publication of the journal. For example, the publisher Longwoods charges \$2,500 to make articles from its journals *Healthcare Policy*, *Nursing Leadership* and *World Health and Population* freely available. For further information refer the following URL: <http://www.longwoods.com/pages/open-access-policy>.

complex, data-driven process (ERA) which placed less emphasis on repositories. During this crossover period the allocation of funding from 2007 via the Australian Scheme for Higher Education Repositories (ASHER) and the creation of a national system of repositories served to maintain the open access momentum.⁴ The RQF revealed that repositories would solve some of the research quality assessment issues, and most especially those associated with outputs lacking meaningful citation counts or other bibliometric data and flagged for peer review.

The stalled RQF of 2005-7 prepared institutions for the task ahead, however there was no easy way around the manipulation of large volumes of data as called for by ERA and the fact that the process was a new one. Further complicating the task was the fact that though the ERA goal posts had been set in 2008, they did not stop moving until the final weeks before submissions were due at the end of July 2010. For example, the Australian Research Council (ARC) – the government department responsible for implementation of ERA - announced on 5 July 2010 that tildes (~) in repository urls, and streaming media files related to items for peer review, would not be accepted by the SEER online management system.⁵ No doubt this caused consternation amongst those repository managers who were unaware of such restrictions. Earlier, on 1 June, the ARC had released SEER business rules stating that 'Not available' or alternate forms were not permitted in the institutional submission XML file for optional terms, reversing a rule which had existed throughout the trial period. Such late changes resulted in increased workloads for repository and research office staff.

The impact of ERA was both positive and negative, depending on how one views the process and its many parts. On the negative side it took repository managers away from the task of securing openly accessible content and promoting open

⁴ Information on the Australian Scheme for Higher Education Repositories (ASHER) program is available here: URL:

<http://www.innovation.gov.au/Section/science/Pages/asherandiap.aspx> (accessed 20 July 2010). Funding was allocated for the period 2007-2010, specifically to support ERA.

⁵ "ILLEGAL CHARACTERS IN URLS - Illegal characters (spaces, square or curved brackets, tildes (~) or

quotation marks) within submitted URLs." ERA Team, 'Stage 2 Data Integrity Checks (Repositories)', (email), ERA Discussion Google Group, 5 July 2010 (URL: <http://groups.google.com/group/era-discussion>). In regards to the issue of streaming media, the *ERA-SEER 2010 Technical Specifications* (ARC, December 2009, 53p) document refers to supported formats, including "popular video/audio format (i.e. mp3, mp4, avi)" (pages 46-7) and notes that "SEER will not accept executable and script based files (i.e. .exe, .bat, .sh)". However the author could not find any specific reference to the exclusion of "streaming files" as per the 5 July 2010 email.

access within their institutions and externally. ERA was not about open access and did not lead to the deposit of a large collection of such material in IRs - material which would generate full text downloads and citations and thereby promote both the institution and its staff as per the traditional IR model. ERA was, for many repositories, a dead weight, requiring a large amount of work for what appeared to be no tangible outcome, apart from vague promises of funding in the future – funding that would possibly be a rehash of existing funding. ERA was something institutions had to do, due to the political imperative, and it had to be done well, to support and enhance institutional reputation. To perform badly in ERA, could result in the loss of ‘research’ status across the institution or within faculties – such was the threatening reality of this seemingly no-fault process.

Outreach activities by repository state were put on hold or scaled down, and many of the positive initiatives arising out of the ASHER funding were not taken up or were frittered away and lost in the haze of ERA data manipulation and guideline adherence. Government withdrew from active talk of influencing open access at the national level and handed responsibility for ERA over to the ARC. With ASHER funding ending in 2010, the long-term sustainability of repositories once again fell question, not helped by a perception that they only existed to serve an ERA support function. The negatives of the whole exercise were all too visible. So, indeed, were the positives.

ERA was a quality tool which, even during its development and submission phase, provided government and universities with a better understanding of research outputs, directions, strengths and weaknesses. It also helped those participating to better manage the research performance process locally – from individual academic through to senior administration. The University of Wollongong open access repository Research Online played an important role in ensuring the success of the institution’s ERA submission and was one element of the process which ran smoothly, both during the preparation phase and in the final making available of digital objects for peer review panels. As a result, kudos was earned by the University Library and repository staff for their role in ERA. It may also, ultimately, have helped improve the profile of repositories amongst government, academics and senior administrators, though this is yet to be fully assessed.

ERA offered the chance to improve various processes, both within the Library and the Research Services Office and out amongst the faculties. ERA was undeniably a distraction, however it forced institutions to push various technological boundaries, especially in regards to repository development, interoperability between repositories and research information systems, and the creation and ingestion of metadata and digital object storage and access. It was in these areas that the ERA legacy lay.

The good, the bad and the ugly – HERDC and ERA

ERA left repository managers around Australia with an increased skills base and expanded knowledge of institutional research management processes and infrastructure. Whilst the RQF and ERA were basically a hard slog, four years down the track there was light at the end of the tunnel. At the University of Wollongong that light is the realisation that the legacy of ERA may in fact be an influx of material into the repository. This was unexpected, but the opportunity has now arisen to leverage ERA to secure repository content for the long term. A local or national mandate may prove unnecessary.

A key to moving forward and making the most out of ERA is the symbiotic relationship between ERA and HERDC. Whilst Australia's research quality assessment process was initially modelled on the British equivalent, both the RQF and ERA owed a lot to the local Higher Education Research Data Collection (HERDC) process which had been in operation since 2002 and continues to exist post-ERA. HERDC development, especially the move towards digital-only submissions, largely went into stasis with the announcement of ERA, whilst the latter evolved beyond mere consideration of four types of published research outputs – as in HERDC – to encompass a wide variety of research outputs and creative works. All the while HERDC continued to operate in the background and could not fail to be noticed by those involved in the ERA process. For example, the backbone of the University of Wollongong's ERA submission was the in-house research information system (RIS) which had been built around a publications database developed in 2006 for HERDC and since that time expanded to include data on research groups and projects.

The HERDC data in RIS formed the core of the ERA submission publication data and continued to be expanded upon to accommodate the latter's guidelines and technical specifications. Unfortunately ERA and HERDC selection criteria and scope varied, and this caused problems as one set of data (HERDC, expanded to include local publication types) was squeezed into an ERA box which accepted most forms of legitimate research and creative work, required more detailed bibliographic information and was constrained by the controversial ranked journal list and subsequent data manipulation (e.g. apportionment) as institutional submissions were finalised.

Those staff working on ERA post 2008 brought a critical eye to RIS and found that a data set 'dirty' in regards to their ERA requirements i.e. it was not as comprehensive as the guidelines demanded, it was not consistent across the ERA 6 year survey period, and it therefore required a lot of 'cleaning' prior to final submission. Accordingly, many questions were asked as to why things were done in a certain

way, with the answer invariably that they were done according to the HERDC rules at the time. The imposition of a changing set of RQF and ERA rules post 2004 meant that a lot of data preparation had to be redone and any set would only be finalised at the point of deadline. However, as the dust settled from all these years of frantic activity, it became clear that various processes and technological innovations were being put in place which could be utilised by repository managers in a post-ERA world, or at least in the period of grace before the next manifestation of ERA was introduced.

More than meets the eye – ERA transformed

At the University of Wollongong the RQF and ERA drove change and left gains which will be used to secure metadata and digital objects for the institutional repository. These changes include:

1. **Interoperability:** Improved integration between the institutional repository Research Online and research information system RIS, enabling the exchange of metadata and digital objects between the two via XML and Excel.

For example, prior to ERA, there was no connectivity between Research Online and RIS. Post ERA, RIS is now the point of entry for bibliographic data and digital objects relating to University of Wollongong research outputs. It is from RIS that RO now, and into the future, will primarily acquire content. Previously this content was acquired directly for RO from academics and faculty staff, or by trawling the internet for open access material. A manual process was then necessary to ensure that a url link was made between the RIS and RO entries so that the faculty member's office University cv page would have live links to their RO entries. Systems are now in place to automatically, and in batch mode, achieve these linkages and exchange data and digital objects between the two systems.

2. **Scaling:** An expansion of the research information system RIS to accommodate both HERDC and ERA.

The in-house research information system RIS was extensively modified during 2008-2010 to accommodate the needs of ERA, whilst taking into account the ongoing obligations to HERDC. RIS can now manage both processes, though they remain distinct. The prominent role of the institutional repository in ERA also facilitated changes to RIS to improve interoperability. This was primarily in the form of enabling batch upload via XML from RIS into RO, and subsequent

retrieval of repository urls back into RIS via SRU. Both of these processes were used and refined during ERA implementation.

3. **IT Knowledge:** Increased knowledge of the research information system RIS by central Information Technology Services (ITS), and to a lesser extent of the simpler repository Research Online.

Research Online is a proprietary system managed by the University Library and off-site by Bepress (US). As such, local ITS support and knowledge was minimal. Collaboration between the Library, Bepress and ITS was increased in support of the RQF and ERA, resulting in an improved shared knowledge base.

4. **Processes:** The implementation of new processes by the Library in regards to dealing with research outputs in connection with HERDC, ERA and open access items for the repository.
5. **Faculty knowledge:** Increased knowledge within the faculty in regards to bibliographic information and related information such as digital objects and citation data
6. **Digital objects:** Increased accessibility to digital copies of research outputs, plus HERDC evidence is uploaded to RIS and can then be exported into the IR
7. **Quasi-mandate:** We have no mandate at Wollongong, and with the distraction of the RQF / ERA, any local momentum towards that end dissipated.