

Institutional Repositories: libraries coming to the help of their institutions

Naicheng Chang¹ and Alan Hopkinson²

¹General Education Center, Tatung University, Taipei, Taiwan

²Middlesex University, London, UK

Abstract: Librarians are increasingly being called upon to help academics justify their existence and bring in research income, with the digital repository. Librarians are being asked to do manage digital repositories of academic research and are best placed to do this. Managers of repositories are using tools available to them to visualise various activities relating to the repository to make it more visible and a better promotion for the activities of the institution whose research is being displayed. Repositories are being used in research assessment exercises like the UK Research Excellence Framework.

Keywords: Digital repositories, Research Assessment, research papers

1. Introduction

Librarians in academic libraries are accustomed to the libraries they manage existing as a repository of purchased and donated materials and have over the years specialised in supplying to students and researchers the materials they hold which are needed by their clientele. They are increasingly using quantitative methods to justify their existence as the people who pay their wages are often unaware of what they really do and threaten them with cuts as an easy option for economising when their host institution is in financial difficulties. Librarians have become quite adept at saying what they do for their readership. Recently librarians have seized upon a new opportunity, the possibility of benefiting their host institution in a way that they had never thought of before. This involved their entry into publishing though it is not usually seen in this way. This is the establishment of digital repositories specifically of the materials produced by their users, mainly the academics but also librarians themselves and the students particularly PhD students.

Not only are librarians justifying their existence by establishing these repositories; librarians are in turn helping academics to justify their existence and bring in research income. The digital repository is becoming a window for

each university to exhibit the areas of its excellence as made manifest through its academics' research papers or in the case of visual arts through reports of exhibitions, exhibition catalogues or even images of art works or buildings designed.

2. Why the Library?

When digital repositories began to take off, I for one did not think that this was a task for the library. Some of the Faculties of the University had already mounted papers on their websites and I saw this as the way forward, with each having its own collection of papers on its website. Middlesex University Library Service felt that the university needed an e-print repository but expected the university's academic faculties (known as Schools) or the university's computing service to take this in hand. However, at a conference for librarians on the problems of electronic materials, Stephen Pinfield of Nottingham University and Director of the SHERPA project, was certain in his belief that it should be the library that provides this service rather than say the university press or the university registry or anyone else for that matter. So Middlesex University Learning Resources service resolved that this was something that should be accomplished. In the event we realised that the repository needs to be supported by a central service to ensure homogeneity of use and provision across a university and to ensure that even the smallest and least technical faculties could use and reap the benefits of a repository. At Middlesex University the library staff included a team that was responsible for library automation and bibliographical services and they were the obvious people to manage a digital repository.

3. Implementing a digital repository

We decided to establish a repository at lowest possible cost (Middlesex University, 2005). This involved one of the library staff who was undertaking a Masters in Library and Information Science at City University and needed a repository for his project) in setting up eprints after having installed Linux on a spare computer which was set up as a server with internet access. We looked at DSpace as well as eprints but decided eprints was better with its support being available from a UK institution. Eventually the software having been set up was transferred to one of the university's virtual servers and supported by a maintenance contract with the eprints organisation but at the point of its installation it was not an official university activity.

When the repository was set up, this was only the beginning. The next part proved much more difficult, persuading staff to use it. In Middlesex University's School of Engineering and Information Science has in the past been very strongly involved in the usability aspect of digital libraries and staff have done some work partly to support research and partly as consultancies for people wanting digital libraries. They could probably have set up their own repository on their own server and of course it would not be impossible for some of the other Schools to do this, but it would be a pity not to have a general Middlesex University server. Many different members of library staff have an interest but

until one person was appointed to oversee the repository little progress was made. The big effort is getting academics interested. In the vacation we found that staff were away and could not contribute. In the term-time they were too busy. The few staff approached seemed a little reluctant to commit themselves to placing their articles in the repository, perhaps because they were not sure of the legalities relating to copyright.

We then turned back to the School of Engineering and Information Science because the Dean was interested politically in the concept of open archives. He asked the researchers and academics in his School to contribute and some did so, but there was not as great a take-up as expected. In two years we had just 44 abstracts and papers.

Eventually the Director of Research who is also Associate Vice Chancellor of the University reported back from a workshop that Middlesex should have a repository and it should be managed by the library. It then took off in a big way. In 4 years we have added over 8000 papers.

4. Other areas of quantitative evaluation

Having established a good repository we have been attempting to make it more attractive to users. These users include those academics self-archiving as well as the people often in other universities who wish to discover researchers at Middlesex with whom they would like to cooperate. We devised a new front end to make it more attractive and a system for making it easier to launch non text materials from the portal itself. These proved too difficult and costly to achieve. So we kept the standard eprints look and feel and added features to indicate usage, the top ten accesses.

The screenshot shows a web browser window displaying a table titled "TOP 10 FULL-TEXT ACCESSES". The table lists the top 10 most accessed full-text documents in the repository, including the number of accesses, the authors, the research division, the title, and the date deposited. The browser's address bar shows the URL "http://eprints.mdx.ac.uk/top10_full_text_accesses.html".

No. of Accesses	Authors	Research Division	Title	Date Deposited
409	Haim Mazar	Health and Social Sciences	An analysis of regulatory frameworks for wireless communications, societal concerns and risk, the case of radio frequency (RF) allocation and licensing.	22/10/2008
300	Kay Cakwell, Lynne Henshaw, Gina Taylor	Nursing, Midwifery and Health Sciences	Developing a framework for critiquing health research	4/11/2009
204	Michael Muller-Camen	Human Resource Management	Unitarism, pluralism, and human resource management in Germany.	24/7/2009
110	Michael Muller-Camen	Human Resource Management	Human resource and industrial relations practices of UK and US multinationals in Germany.	24/7/2009
75	Luke White	Arts and Education	Damien Hirst and the legacy of the sublime in contemporary art and culture.	2/3/2010
72	Marion Colledge, Sandra Hargreaves	Arts and Education	Helping Dyslexic Students to Negotiate the Challenges of Essay-Writing: A case study from a first year module in BA English Language.	1/12/2008
72	Stephen D. Phor, Siu-Tsen Shen, Mehmet Karamanoglu, Sid Odedra, Mehmet Ali Erbil, Chris Barlow, Darren Lewis	Engineering and Information Sciences	The future of battlefield micro air vehicle systems.	28/11/2010
67	James Vasilios Kritsis	Work Based Learning	Managing change in educational organizations: managing and leading higher education in an era of change in Greece. Case study: University of Indianapolis, Athens.	28/4/2009
66	Michael Muller-Camen, Phil Almond, Patrick Gunnigle, Javier Quintanilla, Anne Tempel	Human Resource Management	Between home and host country: multinationals and employment relations in Europe.	9/3/2009
60	Michael Muller-Camen, Wolfgang Mayrhofer, Johannes Ledolter, Guido Strunk, Christiane Erben, Buch	Human Resource Management	Unternehmenserfolg und Personalmanagement: eine international vergleichende empirische analyse.	9/3/2009

The top twenty downloads are recorded in a different way in University College London's eprints repository which is called UCL Discovery (University College London, 2007). They are incorporated into a blog which is kept for 2 months. When viewed, the blog included a commentary: 'The top two papers in January 2012 are reversed in their download order in February 2012, although both papers have proved very popular this month. The WCRF/AICR Expert Report [‘Food, Nutrition, Physical Activity, and the Prevention of Cancer: a Global Perspective’](#), leads the Top 20 downloads with over 1100 downloads, whilst the 2004 IFS briefing paper [‘The ‘fat tax’: economic incentives to reduce obesity](#) saw 970 downloads this month.’ UCL's website also includes a useful page on Open Access (University College London, 2012).

A program has been written by the eprints software developers to link in with Web of Science to provide citation statistics for papers in the repository and this is used to good effect to list the most cited articles in Web of Science citations.

Citations from Web of Science	Authors	Research Division	Title	Date Deposited
289	Torben Lund , Lorraine O'Reilly , Patricia Hutchings , Osami Kanagawa , Robert Gravely , Phillip Chandler , Julian Dyson , Jean K. Picard , Andrew Edwards , Dimitris Kioussis , Anne Cooke , Elizabeth Simpson	Biomedical and Molecular Science	Prevention of insulin-dependent diabetes mellitus in non-obese diabetic mice by transgenes encoding modified I-A-chain or normal I-E-chain.	25/1/2010
258	Peter J. Delves , Ivan Roitt	Biomedical and Molecular Science	The immune system: first of two parts.	18/6/2009
221	Eraldo Paulesu , Eamon McCrory , Ferruccio Fazio , L. Menoncello , Nicola Brunswick , Stefano F. Cappa , M. Cotelli , Giuseppe Cossu , F. Corte , M. L. Lousoo , S. Pesenti , Ann-Marie Gallagher , D. Perani , Cathy Price , C. D. Firth , Uta Frith	Psychology	A cultural effect on brain function.	27/10/2009
185	M. Ali , N. R. Lemoine , Christopher J. Ring	Biomedical and Molecular Science	The use of DNA viruses as vectors for gene therapy.	3/12/2009
183	Jeremy A. Garson , Christopher J. Ring , Philip W. Tuke , Richard S. Tedder	Biomedical and Molecular Science	Enhanced detection by PCR of hepatitis C virus RNA.	3/12/2009
153	Peter J. Delves , Ivan Roitt	Biomedical and Molecular Science	The immune system: second of two parts.	18/6/2009
152	Lucy Ghail , G. W. Neill , M. T. Teh , S. T. Wong	Biomedical and Molecular Science	FOXM1 is a downstream target of Gli1 in basal cell carcinomas.	30/10/2008
134	Bahman Baluch , Derek Besner ,	Psychology	Visual word recognition: evidence for strategic control of lexical and nonlexical routines in oral reading.	18/4/2011
118	Michael Quinlan , Claire Mayhew , Philip Bohle	Management	The global expansion of precarious employment, work disorganization, and consequences for occupational health: a review of recent research.	17/2/2011
106	M. Groudine , Colin M. Casimir	Biomedical and Molecular Science	Post-transcriptional regulation of the chicken thymidine kinase gene.	2/12/2009
100	Frank Uhlir , C. Acosta-Esteban , Dhivya Aravamudan , David Meek	Biomedical and	Limited evolutionary conservation of imprinting in	24/11/2009

This was a little controversial as it does not include much of the university's output, for example that from the Art and Design departments, which is outside the scope of Web of Science.

The number of full text accesses over the last few months is represented here in a table available for all to see.

Number of Full-Text Accesses

Year

Month

16530

2012

February

16931

2012

January

11438

2011

December

22435

2011

November

14661

2011

October

11738

2011

September

10820

2011

August

Additionally, academics who put their output into the repository can see along with anyone else how much they have submitted by year. They can export their data in a variety of reference and other data formats.

5. Additional data sources

Additionally there are external sources of data on repositories. ROAR is the Registry of Open Access Repositories which is hosted by eprints.org (University of Southampton, 2012a). They have set up a directory of repositories which includes information on the type of software used by the repository as well as its date of starting and the number of records currently. As of April 2012 there are 2,730 repositories listed. The National Taiwan University has the most records of any university, at 259,013.

The eprints.org website has much other information explaining and proclaiming the value of repositories and the Open Access movement.

6. Librarians helping to quantify research

In the UK, the Research Excellence Framework exercise is due to start soon, with pilots already beginning. The Research Excellence Framework (REF), previously the 'RAE', is an assessment undertaken on behalf of the UK government to gauge the quality of research in universities. Plug-ins are being developed for repositories under JISC funding which can be used to generate data in standard CERIF4REF and/or native CERIF format. CERIF is an XML schema developed by the Centre for e-Research at Kings College London. A number of benefits result from the ability to extract data in these standard

formats, including the ability to import into a commercially available Research Information System, options to exchange data with other institutions with CERIF-based systems, or use as an interchange format for moving data from one repository system to another. These plug-ins will convert repository data into a format which can be directly uploaded to the REF data collection system, by-passing the need for spreadsheets and/or specialised local REF databases.

The EPrints team is developing innovative features that align closely with the needs of the EPrints user community (University of Southampton, 2102b) . EPrints repositories have a potentially important role to play in institutions' responses to the requirements of the REF. A new REF2014 plugin has been specifically designed to help institutions manage two types of data, their staff details and research outputs.

The REF2014 plugin has been developed in concert with the University of Glasgow. After successful testing with early adopters the plugin is now available on files.eprints.org and via the EPrints Bazaar for one-click installation for those using the latest version of EPrints.

EPrints Services will be offering a package of options to include REF2014-specific technical support, training and expert consultancy for those institutions wishing to automatically populate REF-related fields from other systems.

Thus librarians managing these systems will be at the forefront of universities' research assessment and will surely gain from this exposure to one of the university's core activities

7. Comparison of systems and their value: future work

Eprints and Dspace, the major systems used for repositories, are similar in their characteristics and there is little to choose between them, the main difference being for the UK the features developed for the REF. There are also other possibilities for establishing a repository by having it supported by a commercial organization. In the time available for this paper there is no possibility to report on a comparison but the authors hope that this will be the subject of a future study. Points which need to be taken into account are the tools provided by the packages to encourage a high quality of data to be entered which will enable the alignment of the data in the repository with that in the library catalogue. For example, the data model used by the standard eprints implementation is based on Dublin Core with a large number of additional fields. In the DSpace City University of Hong Kong repository the data elements relating to a thesis are actually labelled with Dublin Core qualified, if “?mode=full” is added to the URL of a retrieved document. Library catalogues tend to be based on Anglo-American Cataloguing Rules currently being superseded by Resource Description and Access. We intend to survey repository managers to see if their database definitions are sufficient to provide the quality of data needed and how they align records from the repository and the catalogue if at all. In the case of theses because they are usually catalogued in library catalogues and now are held as full text in repositories with associated metadata, it will be interesting to see how they deal with data in different formats, and how different the formats are in practice. It will also be interesting to investigate

the extent to which the repositories are becoming used as tools by academics in finding partners in other institutions, and the extent to which the papers in the repository are achieving higher numbers of citations, both questions which are being asked by research managers and repository managers in academia.

References

- Middlesex University (2004). Research repository. London: Middlesex University, 2004. <http://eprints.mdx.ac.uk>
- University College London (2007). UCL Discovery. London: UCL, 2007. Accessed 2 April 2012. <http://discovery.ucl.ac.uk>.
- University College London. (2012). UCL Discovery : What is Open Access? London: UCL, accessed 2 April 2012. http://discovery.ucl.ac.uk/open_access.html
- University of Southampton. School of Electronics and Computer Science.(2012a) Registry of Open Access repositories. Southampton: The University, 2012. accessed 2 April 2012. <http://roar.eprints.org>
- University of Southampton. School of Electronics and Computer Science. (2012b)E-prints: are you ready for REF2014? . Southampton: The University, 2012, accessed 2 April 2012. <http://www.eprints.org/ref2014/>