Building an institutional repository at Loughborough University: some experiences

This item was submitted to Loughborough University's Institutional Repository by the/an author.


Additional Information:

- This article is published in the journal, Program [© Emerald] and is also available at: http://www.emeraldinsight.com/info/journals/prog/prog.jsp

Metadata Record: https://dspace.lboro.ac.uk/2134/2640

Publisher: © Emerald

Please cite the published version.
Building an Institutional Repository at Loughborough University: some experiences
Author: Joanna Barwick is the Institutional Repository Manager and a Support Services Librarian at the Pilkington Library, Loughborough University, UK. E-mail: J.P.L.Barwick@lboro.ac.uk

Abstract

Purpose
This paper provides a summary of our experiences of setting up an institutional repository at Loughborough University and focuses on some of the key issues we have had to consider, the choices we have made and the challenges we have overcome.

Methodology
The paper outlines the various decision processes involved during the 12-month pilot phase. These include: choosing appropriate software; customising DSpace; implementing licences and gathering content for the repository.

Findings
Our experiences highlight some of the challenges involved in setting an institutional repository.

Value
This paper gives a direct insight into the different types of work involved in the setting up of an institutional repository and is an example of a system set up outside the boundaries of project funding.

Keywords:
Institutional repositories; Open access; Research; University libraries; Project development
1. Background

Loughborough University in the UK has an “international reputation for excellence in teaching and research, strong links with industry, and unrivalled sporting achievement. Loughborough is research-intensive. Its ethos from its earliest days has been to produce highly relevant research that solves real world problems to assist both business and industry, and improve people's quality of life… ….Nearly half of the University's income is for research, and its research partnerships and links with business and industry are the envy of other institutions. The quality of teaching at Loughborough is verified by exceptionally high scores awarded under the Government's independent assessment scheme, which have consistently put the University towards the very top of the Teaching Quality league tables”.

(http://www.lboro.ac.uk/about/profile.html).

The university has 24 departments and over 30 research centres and institutes within the three faculties of Engineering, Science and Social Sciences and Humanities and has about 14,000 students.

This paper provides a summary of our experiences of setting up an institutional repository (IR) at Loughborough University (http://magpie.lboro.ac.uk/dspace ). Figure 1 shows the opening screen of the IR.

Take in Figure 1

Figure 1. Opening screen of Loughborough’s IR
The first year of development (May 2005-April 2006) was seen as a pilot project when members of the IR Steering Committee worked closely with different stakeholders throughout the University to raise interest and gain support for the development of this new service. A librarian’s guide to IRs was written by staff at Loughborough (Barwick and Pickton, 2006) and this is available via the IR. By late 2006 the repository contained over 2000 items, including published articles, theses, pre-prints and conference papers and is being indexed by Google, Google Scholar and the Open Archive Initiative harvester and search engine, OAIster. OAIster (http://oaister.umdl.umich.edu/o/oaister/) is hosted by the University of Michigan and contains almost 10 million records from IRs in over 700 institutions. Loughborough is also registered on the Directory of Open Access Repositories (DOAR at http://www.opendoar.org/) which is hosted by the University of Nottingham and Lund University, and the Register of Open Access Repositories (ROAR at http://archives.eprints.org/) which is hosted by the University of Southampton. As evidenced by the entries in these directories many other institutions are involved in setting up IRs. For example, Simpson and Hey (2006) describe developments at Southampton University, Jones and Andrew (2005) outline the work in developing the
Edinburgh Research Archive and Bevan (2005) describes the collection of e-theses at Cranfield University.

Levels of support for the IR within our institution have varied but at the end of the pilot phase, which was marked with an official launch by the Vice Chancellor, we had built strong foundations which should help us to ensure the sustainability of the service. This article is a case study of a year in the life of an IR and focuses on some of the key issues we have had to consider; the choices we have made and the challenges we have overcome.

2. Starting out
Conscious of the developments in open access publishing, the Library at Loughborough had been interested in developing an IR for some time and had set up a Working Group in 2003 to investigate the funding and staffing implications as well as identifying the key advantages for the University. This group (which included the Support Services Manager, the Systems Manager and the Academic Manager (Engineering)) produced a report of findings which was taken to University senior management and initial support was received. The Library was given a small amount of money (£15,000) and a vacancy for an IR Manager was advertised as part of a new Support Services Librarian position and I was appointed in April 2005. An IR Steering Committee was set up to oversee the pilot phase and to draw upon the skills and expertise of people within the Library and University as a whole. The development of this service has involved collaboration between the University Library (with representatives including the University Librarian, the Support Services Manager, the Systems Manager and the Support Services Librarian) as well as Professor Charles Oppenheim from the Department of Information Science and two staff members from the Engineering Centre for Teaching and Learning (engCETL). This group has met regularly throughout this first year to discuss policy and direction.

3. Working with DSpace

There are a number of software options available when setting up an IR and different institutions have chosen different options in relation to the function that their IR will perform. At Loughborough, we wanted an open source solution and DSpace, the
open source repository software developed by the Massachusetts Institute of Technology (MIT) and Hewlett-Packard (http://dspace.org/), was chosen above EPrints (http://www.eprints.org/) and Fedora (http://www.fedora.info/) because it had a good web interface and the ability to manage various file formats: this related well to the fact that we were looking at developing a ‘blended’ repository.

Having chosen to work with DSpace, this was downloaded and installed on the repository server in May 2005. It has proved to be an excellent choice in terms of the flexibility and functionality it offers and maintaining the software has required minimal staff time. Over the course of the year, the Library’s Systems Team have overseen the successful implementation of OAI-PMH; the installation of the ‘Handle’ server and an upgrade to version 1.3, and then 1.4 in Summer 2006, of the software. The Handle System® is provided by the Corporation of National Research Initiatives (CNRI) and is “a general purpose distributed information system that provides efficient, extensible, and secure identifier and resolution services for use on networks such as the Internet” (http://www.handle.net/). The system provides Uniform Resource Identifiers (URIs) for each record within DSpace for participating institutions. We have been able to customise the software, using images from the University’s Media Services’ image bank and departmental logos and this has enabled us to brand the service clearly as the Loughborough University Institutional Repository as was seen in Figure 1.

Another advantage of DSpace is its structure: it gave us the flexibility to organise our repository collections according to the University’s faculty/departmental structure. This has meant that each department has its own area on the repository, which has given departments a sense of ownership of the system. This has proved to be a very important factor when demonstrating the service. It also means that departments or research institutes can link directly from their websites to their staff publications on the repository. Figure 2 shows an example of this link from the website at the Ergonomics and Safety Research Institute (ESRI) and Figure 3 shows the ESRI area within the IR.

Take in Figure 2

Figure 2. Link to the central IR from an institute website
Take in Figure 3

Figure 3 Area for an individual institute within the IR
There are various opinions about what the role of an IR is within an organisation. From the outset, we have focussed on Lynch’s (2003) definition:

*a university-based institutional repository is a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members. It is most essentially an organizational commitment to the stewardship of these digital materials, including long-term preservation where appropriate, as well as organization and access or distribution. .....*

This was relevant to Loughborough as one of the key aspects we have had to contend with is that, unlike many other institutions, Loughborough already has a publications database which serves as a record of the research output of each member of staff and performs many of the functions necessary for the Research Assessment Exercise. This system is an embedded part of departmental workflows: all members of staff are expected to manage their research output within this system and there are departmental administrators responsible for updating the information. For this reason, any repository development at Loughborough was going to be slightly different and the focus has always been full-text digital materials. The initial Working Group recommendations had been that the material within the repository would be limited to
published material: the existence of the publications database has led us to consider a more ‘blended’ approach; whereby, the IR includes different types of material including theses, pre-prints and official reports.

The publications database has been a useful tool for gathering content as we have been able to target specific members of staff and the Steering Committee has worked hard to develop links between the repository and this service. Throughout the first year, discussions were held with staff in Corporate Information Services (CIS), who manage the publications database, about the possibilities of sharing metadata between the two systems and linking via OpenURLs to material within the IR. This is a powerful selling-point for departments which we are keen to exploit, however some work is still needed to convince the relevant people that this is a worthwhile investment of time. In the interim period, we have managed to link the two systems via our ‘SFX’ service (using the MetaLib software from Ex Libris which is used to provide access to a range of other e-resources within the Library) which means that users can find the full text of items in the publications database from wherever they appear in the IR.

4. Metadata and licences

At the beginning of the pilot phase in May 2005, the Institutional Repository Steering Committee identified two key decisions which needed to be made before we could start gathering content. These were related to policies on metadata and the licensing of material within the repository: two different sub-groups were set up to look at these areas.

The first sub-group was tasked with investigating the metadata options for the repository. DSpace relies on a qualified version of the Dublin Core schema and the group identified which fields were necessary for the different types of material expected to be hosted on the service. There has been a lot of debate within the external repository community about metadata within repositories and whether additional subject metadata is necessary for these systems. The group decided that there was value in adding Library of Congress Subject Headings (LCSH) to the records. It was felt that, although DSpace uses free-text searching, LCSH were a
useful addition for when records were exported to external systems such as OAIster as the keywords are clearly displayed in the record. These were chosen as they were a familiar tool to library staff who would be uploading the material. LCSH are used as a controlled language with between four to six headings added to each record, in addition to any author-supplied keywords. Dublin Core, like MARC, has certain failings in relation to digital rights and at present there is still an issue about displaying rights to users. Solutions to these issues are expected to be an outcome of the current Joint Information Systems Committee (JISC) funded Rights and Rewards in Blended Repositories project at Loughborough University (http://rightsandrewards.lboro.ac.uk/) and will involve further enhancements of the DSpace software (Bates et al., 2006). Figure 4 shows an example of the metadata for a record in the IR.

Take in Figure 4
Figure 4 Metadata for item in Loughborough’s IR

The second sub-group was tasked to draw up a suitable deposit licence for the repository. This licence is based on the ‘Creative Commons’ model
(http://creativecommons.org/) and the SHERPA model licence and outlines the rights and responsibilities of author/repository. SHERPA (which stands for Securing a Hybrid Environment for Research Preservation and Access) is:

>a large and varied consortium and has within its membership a range of examples of repository environments and institutional structures. The partners are all research-led institutions, all with practical experience of building and populating eprint repositories”. SHERPA receives funding from a number of organisations including JISC and is involved in “investigating issues in the future of scholarly communication. It is developing open-access institutional repositories in a number of research universities to facilitate the rapid and efficient worldwide dissemination of research. (http://www.sherpa.ac.uk/)

This licence is embedded into the DSpace software, whereby any submitter of material must agree to its terms and conditions or an item is rejected. The licence has caused some problems: some people have felt unable to sign the licence as we ask them to agree that they are the copyright holder and/or have the right to grant the licence – this has been overcome by offering advice and support on copyright policies. There has also been some debate about whether we need an individual licence for each item submitted. The original licence was specific to a particularly ‘work’ but, after seeking expert advice, this has been settled by a minor change of the licence’s wording.

5. Early adopters

Six heads of departments volunteered their departments to take part in the pilot phase after a presentation by library staff at an Information Services Committee meeting. In the first months we met with representatives from each of these ‘early adopters’ to discuss the best way forward. It became clear very early on that different departments would have different reactions to the service: departmental motivation is not necessarily the same as institutional. One department already had an established pre-print archive on their departmental page which they were happy for us to add to the repository. Similarly, another department already had a working paper series on their departmental website and were at the time looking at using the Research Papers in
Economic (RePEc) subject-specific repository for depositing their publications in an open access manner (http://repec.org/). Some members of staff were very concerned about copyright issues and were hostile to the idea of electronic theses being submitted to the IR, but there were often some keen individuals. For some disciplines, theses were seen as a key tool for launching the academic careers of their authors and they were not interested in making these available for all in an open access repository. One member of staff questioned how making her work available would affect any royalties she received. In both cases, the potential benefits of open access were used to persuade these individuals that the advantages outweighed the disadvantages.

These initial meetings gave us an insight into how best to approach other departments. Different disciplines have very different workflows and attitudes to the research process and it became very clear that the departments would have their own views of what and how they wanted to use the service: part of the success of the repository has been our ability to be responsive to these differences. Working with these volunteer departments helped us solve the ‘chicken and egg’ problem: in the period June 2005 – January 2006, we had amassed approximately 600 items which we could use in demonstrations to new stakeholders. It also gave us an opportunity to test the software and our policies before widening the service to other departments in the University.

6. Advocacy

In September 2005, the Library’s academic librarians were encouraged to become involved with the advocacy for the IR and were asked to arrange meetings to introduce their departments to this new service. This has ranged from a five-minute brief at a departmental meeting to an hour-long seminar to a research centre. In the literature on the development of IRs, gathering content and selling the service to individuals is cited as the most difficult aspect: Loughborough has been no different. As Grieg (2005) states:
the cultural change required to build a successful repository does not happen overnight. An advocacy campaign will require a significant amount of repetition and reinforcement of the message.

Grieg’s report on advocacy was one of the outcomes of the JISC-funded DAEDALUS project at Glasgow University (Grieg and Nixon, 2005).

The academic librarians have played a crucial part in helping us to say the ‘right things’ and to direct presentations to individual groups but the response of academics to the IR has still varied from department to department. In fact, reactions have varied greatly within departments too. Some are very enthusiastic; some hostile; others merely sceptical or indifferent. There is often concern that this is another administrative demand on their already limited time and there is never a ‘good time’ to approach people: the demands on academics seem to be constant. At present, we are not asking academics to self-archive and all the work is being carried out within the library. It was hoped this approach, also used for the development of our Reading List System, would encourage them to participate more freely and seems to have been effective in many cases. Some departments were facing internal decisions about how to manage their publications and offering them the repository as a solution has been met with enthusiasm. This is the main reason why we have been flexible in our acceptance of different material: we have worked closely with departments, listening to what they want from the service and have developed their collections in relation to this. Over the course of this first year, we gathered over 1800 items and Figure 5 shows the growth of the IR during this period. From July 2005 –January 2006, it shows steady growth as we worked with our ‘early adopters’: the period from February 2006 onwards, and the steep increase, correlates to the phase at which we widened the service to other departments. There are a growing number of different types of material which includes pre-prints; working papers; published articles; conference papers and theses.

Take in Figure 5
Figure 5 Graph showing growth of Loughborough University’s IR: July 2005- May 2006
7. Challenges – copyright and versions

The main obstacle to the development and growth of IRs is restrictive publisher copyright policies. When an academic submits an article to a journal, most publishers will expect them to sign a Copyright Transfer Agreement. This agreement outlines what rights an author has to re-use their work after its publication. In most cases, the publisher will allow some form of archiving in repository systems (such as ours): this varies from whether this is a pre-print or permission to deposit the final version/post-print. According to the SHERPA website, 64% of publishers will allow post-print archiving which equates to around 90% of journals (http://www.sherpa.ac.uk/documents/15concerns.html). However, despite these permissions, there are usually certain restrictions imposed: the main restriction is that the publisher PDF cannot be used and others have embargo periods: for example, Taylor and Francis has an 18-month embargo on social science articles (see http://www.tandf.co.uk/journals/copyright.asp). This has led to problems because there seems to be a general misunderstanding amongst academics about copyright and the rights they have to re-use their publications after they have signed copyright transfer agreements. As part of the submission process, we are offering to do any copyright checking on behalf of authors. This is done using the SHERPA/RoMEO
database (see http://www.sherpa.ac.uk/romeo.php) or by contacting publishers directly with a standard e-mail. RoMEO (Rights Metadata for Open Archiving) was another JISC-funded project which investigated the rights issues surrounding the 'self-archiving' of research in the UK academic community (Gadd et al., 2003). The process of copyright checking can be quite slow but on the whole permission is given. The difficult part is obtaining the correct version of the paper as submitted by the author(s) for publication. Having to use the author final version has raised questions about quality control and is often thwarted by the fact that academics do not keep these versions or, for some, that their final version is very different to the version that is eventually published. Explaining these policies and talking through these issues is often the most difficult and frustrating part of encouraging people to deposit material.

These difficulties have highlighted the importance of having an institutional statement of copyright, which supports and encourages authors to retain control over their work. This was a recommendation made to the University’s Information Services Committee in May 2006: this was generally supported and Mary Morley (University Librarian) and Professor Charles Oppenheim (Department of Information Science) were asked to draft a document for circulation to all members of staff. This paper includes two main points:

- [the University]…strongly encourages all staff to submit copies of their research output… to the Loughborough University Institutional Repository
- Loughborough University recommends that its staff do NOT assign copyright to publishers.

After agreement from the committee and with support from our University’s Vice Chancellor, this statement will soon be sent, with information about the repository, to all academic and research staff - an important step towards embedding the Institutional Repository into academics’ research workflows and a clear sign of senior management’s support for the service.

8. Sustainability and the future
This level of support is an essential factor in ensuring the sustainability of the IR at Loughborough. Throughout the first year, we have engaged with various stakeholders to win support and to raise the profile of the service. We have met with representatives from the Research Office to discuss the possibility of the electronic submission of theses, a development we are keen to encourage. We have attended Information Services Committee and a Programme Development and Quality Team meeting to give presentations and we have made contacts with Professional Development to discuss the inclusion of the IR on various courses. The repository should now feature in the ‘How to get published’ course as well as the new staff induction. Building on these relationships will be key to the future of the service.

Having met with all our departments, approaches are now being made to the University’s Research Centres and the academic librarians will continue to play an important role in the advocacy of the service. The workload will begin to be incorporated into other workflows within the Library with training and support offered by the Institutional Repository Steering Committee. The committee will continue to meet as the ‘Institutional Repository Advisory Group’ and will oversee the implementation of many of the areas covered in this report as well as continuing to monitor developments.

External developments will have a significant impact on the development of the Institutional Repository. The Research Councils UK’s statement, released on 28th June 2006 (http://www.rcuk.ac.uk/access/default.htm) was a signal from the Research Councils about the importance of open access and may have direct effects on Loughborough’s research community. The Advisory Group will have a role in maintaining awareness of developments such as these, as well as any legal and technical issues in relation to IR development and the future of scholarly communication.

References (All URLs were checked 11th January 2007)


http://www.dlib.org/dlib/september03/gadd/09gadd.html

Grieg, Morag (2005), Institutional Advocacy Campaign: Guidelines and Practical Advice. DAEDALUS project: Workpackage 3 – Advocacy, available at:
https://dspace.gla.ac.uk/handle/1905/377

http://www.ariadne.ac.uk/issue45/greig-nixon/


http://www.arl.org/newsltr/226/ir.html