Working on an instructional repository: cataloguer, metadata specialist or something else?

Louise Harrington, Repository Administrator, University Library Service, Cardiff University
Tracey Andrews, Repository Manager, University Library Service, Cardiff University

The Cardiff University repository team is currently part of a wider research support team that includes Open Access funding and advocacy, research analytics, and the recently-established Cardiff University Press. The team consist of librarians and library staff who come from a wide variety of backgrounds: while the Repository Manager does come from a cataloguing background, the other team members originally worked as library and IT assistants or as library service managers, while several of us have subject librarian experience. Certainly, at Cardiff, we have found that the institutional repository role has widened out from a branch of cataloguing and metadata to one that involves more advocacy, training and direct customer interaction. One thing all of us have in common though is that we enjoy the detailed work involved in creating a perfect metadata record, while a subject specialist background is also useful as you get to know the subject resources and the researchers well, so will have some idea where to find elusive information to create or edit that perfect record.

The history of ORCA 2003-2017: where we got to where we are today

Cardiff University's institutional repository, ORCA (the acronym stands for Online Research @ Cardiff) started life as Cardiff EPrints in 2003, and in 2008 was upgraded and relaunched as ORCA, becoming a fully supported library service that contained, amongst other research outputs, all the University's Research Assessment Exercise (RAE) data. ORCA becoming embedded into researcher practice was given impetus by the University's Research Committee to approve the required deposit of new publications in ORCA, "subject to detailed consideration as to the means by which staff up-load data to the repository".[1] Between 2009 and 2011, ORCA received JISC funding to create a tool that would streamline the deposit process, minimise the deposit effort and simultaneously ensure capture of all required information without duplication of effort; in short create an easy, user-friendly workflow that would encourage authors to use the repository.[2] Web feeds were also created by an IT colleague so that schools could use the metadata in ORCA to feed their publications profiles on the school web pages. All new deposits are reviewed by repository staff before submission to the live database; the publication data is checked and, where available, information such as DOI, URL, abstracts, volume, issue and page numbers, keywords and subject fields are added. Because minimal data is captured from the user, there is more emphasis on the role of the library professional in adding data and making the record as complete as possible. This is where a cataloguing or metadata professional is invaluable in institutional repository work.

[1.] Harrington, Louise and Scott Hill. 2011. The I-WIRE Project final report. Available at: http://orca.cf.ac.uk/74670/
[2.] Harrington, Louise. 2011. Making a splash with ORCA: Cardiff University's Institutional Repository. SCONUL Focus 52, pp. 17-20.

The new deposit tool, the ability of the web feeds to update web pages within 24 hours, and the requirement from the University Research Committee for researchers to deposit bibliographic details of their work into the repository, combined with the decision to use ORCA for the 2014 Research Excellence Framework (REF) meant that researchers or administrative staff from all the schools began to deposit into ORCA in rapidly increasing numbers. Using ORCA as the repository for all the research outputs that could potentially be submitted to REF2014 meant a great deal of work for the central repository team, especially as many of the outputs were not already on ORCA, which resulted in enormous backlogs of outputs to be deposited, edited and reviewed. Getting the research into the repository also became very time-critical due to internal 'mock' REF programmes: a lot of schools would have internal deadlines that we did not know about which lead to a lot of pressure on the small central team. As well as the repository team and cataloguers working on ORCA, there was invaluable support from subject librarians and library assistants, who did a lot of adding and editing of records, while the repository team would do the final checks and submissions. The decision to involve more library staff in the process of course meant widespread training and ongoing support – another skill that ORCA team members have had to learn and develop!

No sooner had we submitted our return for REF 2014, then we had to learn new processes for the next round of the Research Excellence Framework (the phrase 'moving the goalposts' springs to mind!). The next REF requires journal articles and published conference papers accepted after 1 April 2016 to be made freely available as open access documents, which is a major adjustment for both academics and library staff, requires us all to learn new rules about copyright and embargo lengths, and also means a change to our editing processes. Because outputs have to be deposited at acceptance, our role is no longer so much checking and adding publications details, it is instead ensuring that deposits are added to the depositing in a critical time-frame and that publisher copyright rules are checked and embargo lengths are calculated correctly. It also involves a lot of liaison with our colleagues in the library Open Access team regarding payment of gold open access fees and correct license statements. While metadata skills are still needed, particular for items such as books, book chapters or reports which are not affected by the REF open access policy, knowledge of copyright and publisher requirements and advocacy skills are increasingly needed for institutional repository work. Institutional repositories were of course initially created as open access repositories in an attempt to circumvent the 'serials crisis' of the 1990s and 2000s where academic journal articles were hidden behind paywalls and library subscription charges skyrocketed. Due to many factors, however (for example, being unaware of copyright issues, suspicion of, or lack of understanding of, open access, simply not keeping the correct version of the file, using subject repositories such as ArXiv, the Physics pre-print server), institutional repositories have tended to be metadata archives rather than full text databases; at Cardiff full text files account for between 10-20% of the records. While there are always individual authors who support open access, and some disciplines which are open to it, it has been a slow process to educate and engage academics in the positive aspects of open access to research, one that is still ongoing.

Tips, hints and pitfalls

Institutional repositories have become increasingly central to university administration and research processes, especially with the REF and funder requirements. Good (or poor) quality institutional repository data affect as all audiences in the University – from academic staff to administrators to PhD students who, as a result of a change in University regulations, have been required to make their theses available online through ORCA since 2011. Positive aspects of the role are that you do get a lot of feedback, the vast majority of it being pleasant – something that you don't always tend to get as a cataloguer! You also do get a lot more contact with colleagues across the University, whether via email, phone or in person, and there are a lot more opportunities for training and advocacy, if this is an area which interests you; it is good for continuing professional development.

As cataloguing colleagues will understand, the quality of the data that is uploaded to ORCA depends increasingly on the quality of the data that we are given or can source; put simply, if rubbish data goes in, then in most cases, rubbish data comes out! Because of the requirement to make deposit as easy as possible for researchers in order to engage them, our system allows bulk imports from external databases that we have found to be poor sources for good publication data. By making things easy for depositors, we have made things more difficult for ourselves which can be frustrating. The other reason why data is suffering is due to REF requirements: because articles and published conference papers now have to deposit at acceptance, rather than publication, authors have very little information about the paper, and there is nothing online that we can access to improve the record quality. It is also fair to say that unless you come from a cataloguing background, you can miss the face-to-face role of the library site, as institutional repository and scholar communication teams tend to be office-based. However, we have found that many colleagues in the library sites wish to help with ORCA so we do get to visit libraries to help train and engage people.

Conclusion

A background in cataloguing is a huge asset when working on an institutional repository because of the detailed data quality work involved, and the emphasis on findability, but as we have shown, staff from any background can work in this area. While perhaps in its initial incarnation, institutional repository work was seen as mostly for metadata professionals, the move towards open access and the subsequent need for training and advocacy in that areas means that we are developing lots of different and transferable skills. There are lots of opportunities in this area with universities moving towards using Current Research Information Systems (CRIS) and increasing interest in open Research Data Management (RDM) which could provide further opportunities to transfer skills used in metadata work across to these systems and wider research support areas. Cardiff University's institutional repository's main aim is ultimately to promote and make our research (and where possible, our research outputs) available to the widest possible audience across the internet in the true spirit of open access. Our role in enabling this with good quality metadata gives us a lot of pride and job satisfaction, whatever our library background, in a job well done.

Bibliography

Giesecke, J. 2011. Institutional repositories: keys to success. *Journal of Library Administration* 51 (5-6), pp. 529-542. DOI: 10.1080/01930826.2011.589340

Harrington, L. 2011. Making a splash with ORCA: Cardiff University's Institutional Repository. *SCONUL Focus* 52, pp. 17-20.

Harrington, L. and Hill, S. 2011. The I-WIRE Project final report. Available at: http://orca.cf.ac.uk/74670/

Jain, P. 2011. New trends and further applications/directions of institutional repositories in academic institutions. *Library Review* 60 (2), pp. 125-141. DOI: 10.1108/00242531111113078