

# Catalogue and Index

Editorial

June 2012, Issue 167

Those of us who don't work in academic institutions might find repositories a bit of an unknown entity. In this issue we hope to solve your dilemma – we give you everything you always wanted to know about repositories but were afraid to ask. We start with articles by Gareth Johnson and Jackie Wickham that provide some background information, and follow up with more personal articles from a variety of practitioners that give a flavour of what it's like managing a repository, dealing with classification problems, and moving from cataloguing into repository management. We hope these will answer most of your questions. Finally, we include a review of a seminar and a book review.

Cathy Broad

[library@ethicalsoc.org.uk](mailto:library@ethicalsoc.org.uk)

Heather Jardine

[heatherjardine402@hotmail.com](mailto:heatherjardine402@hotmail.com)



## Contents

2-4 UKCoRR – At the Heart of the UK Open Access Repository Landscape *by* Gareth J Johnson

5-8 The Repositories Support Project *by* Jackie Wickham

9-15 Managing a Creative Arts Institutional Repository *by* Anne Spalding

16-19 Repository metadata for diverse collections *by* Alex Clarke and Anna Lawson

20-23 Cataloguing, classification and the University of Northampton Institutional Repository *by* Janet Milner

24-27 Implementing a Current Research Information System (CRIS) with an existing Institutional Repository (IR) *by* Dominic Tate

28-32 Why not send a cataloguer? *by* Janet Aucock

33-34 Reviews

## Why not send a cataloguer?

Janet Aucock, Head of Cataloguing and Repository Services, University of St Andrews

The moment when the cataloguing world and the repository world converged for me was in 2003. The place was the eleventh floor of the Livingstone Tower in Glasgow, then home of the Centre for Digital Library Research (CDLR) at the University of Strathclyde. The reason was a JISC project called HalRST - *Harvesting Institutional Resources in Scotland Testbed*<sup>1</sup>, a project in which St Andrews was a partner. The amazing views of Glasgow will be forever associated with HalRST meetings and with my introduction to “repository speak”, to the new buzzwords of metadata, OAI-PMH, interoperability, harvesting and discovery.

In the beginning it was rather unclear who would represent St Andrews on the project. Perhaps this was indicative of the lack of awareness of what a “repository” could represent and what sort of data it could contain. As we now know, there is no hard and fast definition of this and content can be very diverse. Our participation had been initiated by a Vice Principal with a background in IT and there was preliminary talk among academic staff of setting up an eprints archive. The most likely project representative seemed to be our University Archivist. The project title appeared to match their job title, but our archivist begged to differ. Library managers thought about how to match existing skills and experience to this new project and came up with an answer “Why not send a cataloguer?” And so it began.

### So why is a cataloguer suited to repository work?

My cataloguing experience had started straight away after finishing a PG qualification in the 1980s. I hadn't been altogether sure if this was my real long-term choice of specialism within the profession. Cataloguers were traditionally hidden behind the scenes and that wasn't particularly appealing. However, what did appeal was the power to organise and control bibliographic data, especially in order to make that data available in a key library service, the Library catalogue. Universal Bibliographic Control was the mantra of the time, albeit only then in the context of print resources. Underlying all of this was the technological shift of paper and microfiche catalogues to online systems and the growth of bibliographic utilities, which could readily provide sources of catalogue records.

I spent a period of time simply cataloguing and learning the tools of the trade. New acquisitions at St Andrews were varied, with a significant percentage of foreign language material and a healthy amount of “grey” literature and older publications. AACR, MARC, LCC and LCSH were the cornerstones, and standards closely adhered to. By the 1990s I led the cataloguing team at St Andrews and had gained valuable experience in working closely on three catalogue data system migrations, moving us initially from microfiche and finally to a webpac interface. By the early 2000s that experience had extended even further with a migration from UKMARC to MARC21 and a period of time as the Library Systems manager. In the meantime, there were extensive retrospective cataloguing projects in progress. We also embarked on rare book projects which showed me how bibliographic descriptions could be augmented and extended into new areas: for example physical descriptions of bindings, provenance notes, genre headings and specialist indexes for rare book printers and publishers. By 2003 we had started to grapple with the new dimension of electronic publications, especially electronic journals. Working on the look and feel of webpacs and how we displayed, indexed and presented our catalogue data to our users taught me to always remember the end user and the importance of the public facing service we provide.

So how did this prepare me for the projects we embarked on from 2003 onwards? I understood the structure and organisation of bibliographic data. I knew about a variety of formats and was starting to understand the demands of cataloguing electronic resources. I had experienced detailed data migrations and covered the minutiae of mapping data fields and subfields, as well as understanding how that data could be indexed and displayed in library systems. I had become well acquainted with quality control, standards, productivity and project management with regard to data. Data should be fit for purpose and sometimes it needed to be sophisticated as shown by rare book and retro projects. Efficiency dictated that we should reuse data and source readymade data as much as possible. Better still, data upgraded at local institutions could be contributed back to large national and global data aggregations to improve discovery for users and to put that data back into the pot for reuse by other institutions.

Perhaps what I should particularly highlight are metadata skills. This is a fundamental transferable skill to support repository work and probably deserves a heading on its own. Identifying data elements and structures is key, whether this is in a MARC21 database or a Dublin Core repository.

In retrospect I can see that all of this was leading towards digital projects and new ways to present our data. But perhaps this wasn't immediately obvious at the time. In 2003 the cataloguer's role was still determined by being limited mainly to the print material that traditionally made up a library catalogue database. It required the creation of other silos of data to break this tradition. It required coming out of the catalogue box, taking on digital projects, transferring skills and understanding different approaches and solutions.

## **Repository development**

By the end of 2005 the HalRST project had reported its results<sup>2</sup> and we had produced a local project report about our St Andrews experiences<sup>3</sup>. We were convinced that developing a repository solution for institutional research publications and especially for electronic theses was a strategic priority<sup>4</sup>. We moved forward with a DSpace repository using Dublin Core metadata and hosted by the Scottish Digital Library Consortium<sup>5</sup>. Parallel to this we worked on the institutional policy and infrastructure to support an e-theses mandate.

## **Does a cataloguer still have a lot to learn?**

The answer to this is an emphatic yes. As described earlier, existing skills and experiences were invaluable for the setup of a new database and implementing enhanced structures and metadata content. The EThOS toolkit<sup>6</sup> provided valuable help with information about UKETD\_DC: the metadata core set of recommended Dublin Core fields for theses and dissertations and the UKETD\_DC Application Profile. I could see that previous database management skills would expedite the DSpace setup. Knowledge of subject indexing schemes such as LCSH and LC would add value to repository metadata.

But there is still a lot to learn in order to set up and manage a repository service. One key element is to learn to accept and embrace user-generated metadata. The repository is a good exemplar of this and although quality control is still very important it is helpful to develop a healthy and pragmatic attitude to data quality in a user-driven environment. A certain amount of compromise is required when implementing what are accepted and given standards in the catalogue world. Authority control is a good example of something that requires development and patience.

These are just the metadata-related issues. On top of this is a whole raft of skills and knowledge to develop. A repository manager needs to understand open access in all its forms and to understand the research publication process. They need to acquire new knowledge about publisher policies, licensing and copyright issues. They need to understand their institutional environment, especially when implementing new policies to make digital material available. This was certainly our experience in providing an e-theses service<sup>7</sup>. Add to this an awareness of preservation issues, an understanding of open source software, confidentiality issues, the design of effective workflows to meet the needs of users and of other central services within the institution. A crucial element is to understand the many aspects of the academic mind, how research is carried out, how it is transformed into publications and how those publications support career development. It is difficult to prepare for this learning curve but it would be much tougher without having a background in the transferable skills already described.

### **The pay off and the benefits**

There is always something to compensate for the hours spent trying to understand the Scholarly Works Application Profile<sup>8</sup> and how this might relate to your repository metadata. Or to make up for the work to map data elements from a Research information system into DC with crosswalks involving schemas like MODS<sup>9</sup>. New skills promote professional development. Repository work has extended my contacts and communication into a whole new community beyond the cataloguing community. It has led me to a variety of workshops, meetings and conferences across both the library sector and, more recently, the research management sector and has proved to be a source of genuine satisfaction and professional development.

This is now reflected in my own job title and also the way my role fits into the Library organisational structure. My Head of Cataloguing role is part of the Collections Management Division of the Library service, but my Head of Repository Services role is part of the Academic Liaison Team, in recognition of the user based services and liaison opportunities for research support which these services provide. It is also mirrored in another member of the Cataloguing/Liaison team who originally catalogued print theses for the Library catalogue and is now a key member of staff in the e-theses service. One element of that workflow is to link thesis records in the Library catalogue to their corresponding digital entries in the repository and to ensure that both sets of metadata are suitably enhanced. It's a fitting reminder of the close relationship between the services and the skills needed in both systems and one also described in a previous issue of Catalogue and Index<sup>10</sup>.

### **Momentum and new services**

Repository services<sup>11</sup> at St Andrews have recently expanded and this has been reflected in the development of a repository team. From its inception in 2007 the repository Research@StAndrews:Full text<sup>12</sup> has now become part of an integrated set of services. One primary emphasis is on integration with the University's research information system, a CRIS using PURE software and this aims to capture the research output of the University, providing open access full text for publications wherever possible<sup>13</sup>. In addition, the Library offers a journal hosting service using Open Journals Software<sup>14</sup>.

Interestingly structures and workflows in research information systems parallel those of cataloguing and repository systems. Publications data is sourced and reused from external sources wherever possible. Data is moved in and moved out and repurposed into aggregations for research groups or research pools. Quality

control and workflows are established. Access is required through subject metadata. Public interfaces are essential to promote institutional research. New data models such as CERIF<sup>15</sup> are now becoming well established in the field and become another essential standard for the cataloguer and repository manager to absorb.

It is encouraging to see that cataloguing skills and repository skills can transfer yet again to this new service environment. Research information systems will be instrumental in providing accurate and precise bibliographic data for the upcoming REF research assessment exercise and this itself positions bibliographical skills centre stage. My own role seems to be travelling along a path of increased integration. As cataloguing, repository and research information systems develop and integrate so the interdependencies of these teams within the institution demonstrate increased integration. The associated skills and knowledge also moves around within these teams and is exchanged and consolidated. Thus in St Andrews there has been a very real cooperative approach to these services across the Library, IT services and the Research Policy Office.

## The future

Clearly the role of the cataloguer can transform in the repository and research information environment. So what of the library catalogue itself: has it moved with the times and out of its own silo? It was once the flagship and sole database for many libraries, but now it is one of many competing databases, many of which reflect digital content. Discovery services such as the SEEKER service at St Andrews, using the EBSCO discovery platform, provide a blend of institutional and external content and a centralised institutional entry point into information resources. In this sort of service, catalogue content and repository content is served up in the same interface. Behind this of course is a spot of catalogue data mapping and some OAI harvesting. But the data stands up remarkably well to this sort of interface, perhaps testament to the involvement of cataloguers in repositories? In this environment the catalogue itself seems to have undergone the same journey and joined an integrated service in the same way as the staff creating it.

I would encourage cataloguers to get involved in repository and related services. You have more in common with these services than you may think. Get out of your comfort zone. Your knowledge and skills are building blocks for repository services. The current buzz is research data management. Perhaps we will send a cataloguer or a repository manager to the first project meeting on research data? What this is really about is transferable skills which our profession excels in. We can all adapt, although perhaps cataloguers are demonstrating their expertise in adapting more than most.

## References

<sup>1</sup><http://hairst.cdjr.strath.ac.uk/about.htm>

<sup>2</sup><http://hairst.cdjr.strath.ac.uk/documents/HaIRST-FAIR-FP.pdf>

<sup>3</sup><http://hdl.handle.net/10023/624>

<sup>4</sup>Aucock, J. 2009. Developing the Digital Research Repository at the University of St Andrews. *SCONUL Focus* 46. Available at <http://www.sconul.ac.uk/publications/newsletter/46/12.pdf>

<sup>5</sup><http://www.sdlc.ed.ac.uk/index.html>

<sup>6</sup><http://ethostoolkit.cranfield.ac.uk/tiki-index.php?page=toolkit>

<sup>7</sup>Conference poster at ETD Aberdeen 2008 <http://hdl.handle.net/10023/513>

<sup>8</sup>[http://www.ukoln.ac.uk/repositories/digirep/index/Scholarly\\_Works\\_Application\\_Profile](http://www.ukoln.ac.uk/repositories/digirep/index/Scholarly_Works_Application_Profile)

<sup>9</sup>Metadata Object Description Schema <http://www.loc.gov/standards/mods/>

<sup>10</sup>Care, Rachel. 2011. Ensuring our professionalism is necessary in years to come: the transfer of traditional skills to other domains. Catalogue and Index 162. Available at <http://www.cilip.org.uk/get-involved/special-interest-groups/cataloguing-indexing/Documents/CandI162.pdf>

<sup>11</sup>Library web pages <http://www.st-andrews.ac.uk/library/resources/electronic/fulltext/>

<sup>12</sup>Research@StAndrews:FullText: <http://research-repository.st-andrews.ac.uk/>

<sup>13</sup>Proven, J. and Aucock, J. 2011. Increasing uptake at St Andrews: Strategies for developing the research repository. ALISS Quarterly 6(3): pp6-9. Available at <http://hdl.handle.net/10023/1824>

<sup>14</sup><http://www.st-andrews.ac.uk/library/services/other/journalhosting/>

<sup>15</sup><http://www.eurocris.org/Index.php?page=CERIFintroduction&t=1>