Catalogue and Index

Editorial

December 2012, Issue 169

Contents

See p. 2 for full contents list

Welcome to a bumper issue 169. In it you will find most of the papers and presentations given at our very successful conference in Sheffield in September. Taking "The value of cataloguing" as its theme, it demonstrated that cataloguers are a vigorous and robust breed, engaging with change of all kinds and remaining relevant to the delivery of every type of library and information service.

On a sadder note, we have to say goodbye to my co-editor, Cathy Broad, whose valedictory issue this is. She has helped enormously with the last half-dozen issues and with re-establishing a regular schedule of publication, and her keen proof-reader's eye will be much missed. Many thanks for all your help, Cathy, and very best wishes for the future!

Heather Jardine

(heatherjardine402@hotmail.com)

Cathy Broad

(library@ethicalsoc.org.uk)

Contents

CILIP CIG conference: The value of cataloguing

Katie Flanagan **Special Collections Librarian, Brunel University**

This year's conference was the first time I'd been able to attend since joining the group, although I've been to a few CIG library visits before. It was a very intense couple of days, which I blogged about on my return and was then asked to write up for the newsletter from the perspective of a conference newbie. As others have written accounts of it elsewhere, and the conference presentations are avail-able on the website, these are just some of the thoughts and reflections I brought away with me. They are very much from the perspective of a special collections librarian (I was Rare Books and Special Collections Librarian at the Royal College of Physicians when I attended the conference, although I have since started a new job at Brunel University).

First of all, and I really should know this by now, unlike the stereotype, librarians always seem keen to share their knowledge and get to know new people. The Cataloguing and Indexing Group was no exception. I knew only a couple of other attendees in real life, although I "knew" some of the others from Twitter and blogging, yet I quickly met a whole crowd of people all exchanging information, opinions and advice. Apart from anything else, it was enormous fun (well, have you ever been in a pub with about 30 cataloguers before?!). The university accommodation in Sheffield was very com-fortable, with excellent catering. Despite this there was a table at the back of the conference room for people to share the cake some delegates had made and brought with them. Before the conference I was worried that a two-day conference featuring 27 different speakers was going to be overwhelmingly intense, yet this was not at all the case. I found it tremendously interesting hearing so many different voices and experiences and learnt a great deal. The decision to split the conference into themes and, within that, into 30-minute talks and shorter lightning-round talks and case studies, meant the time never dragged.

Second, don't be scared about the future. Yes, there are big changes coming – not only RDA, but also that the world we work in is changing, especially in the current economic climate. Yet we heard from some brilliant speakers about how their libraries had weathered storms in the past (Heather Jardine, from City of London libraries, pointed out that everyone had survived the arrival of AACR2). There was such a range of experience present: some libraries were still using UKMARC, others were about to jump to RDA. There was controversy: is MARC really about to die? (probably). Many people talked about the advantages of collaborative working, which offers so many advantages. My favourite was Deborah Lee, from the Courtauld Institute, on NACO funnels, a means of both collaborating, re-ceiving training and increasing the usability of name authority headings for British users. I'd love to join a NACO funnel, but I doubt it would be possible for me at the moment. And I really enjoyed Ian Fairclough talking about collaborating to improve bibliographic data quality – the blog, Typo of the Day for Librarians, is well worth checking out. There was a definite feeling of "we're all in this to-gether" from the whole conference.

Third, librarians can really make a difference. Well, we knew that, but does the rest of the world? Dave Pattern's (University of Huddersfield) keynote address reported on research that shows how library use can predict what grade students will get. Amusingly, there is even a 2 a.m. rise in library use for students who get lower grades – demonstrating that they've left it to the last minute. Ap-parently the highest achievers are in the library by 9 a.m. How does knowing that affect how we pro-mote our library services?

Other key points, which I've made into a list of things to think about doing:

- Make the most of the functionality in your catalogue and what it can tell you about how people are searching it. Then use that data to enhance catalogue records.

- Cataloguing consistency AND workflow are important – how to catalogue consistently, speed-ily but still providing what users need. This means making some strategic decisions about outsourcing work, prioritisation, using tools that can speed things up and making use of controlled vocabularies. These were areas all touched on repeatedly in presentations by Lucy Bell (UK Data Archive), Michael Emly (University of Leeds) and Gary Green (Surrey County Council Library Services). Gary should also get an award for most amusing present-ation title for, “The incredible shrinking cataloguer meets the spaghetti junction automation robot”.

- Don’t ignore the history of your catalogue. Keep a record of the cataloguing decisions you take, so that future generations of people working with your catalogue will know what they’re dealing with (Anne Welsh and Katharine Whaite, both University College London).

- Start thinking in terms of data in networks, rather than hierarchical trees (Simon Barron, Durham University Library). RDA is the perfect opportunity to think about how we can do this. I think this has particularly strong application in the world of special collections librar-ianship.

- Find out more about RDA. Céline Carty (Cambridge University Library) updated us on RDA, having been to the ALA annual conference in Anaheim. She provided a handout with helpful links to the RDA Essentials webinars, amongst other resources. Stuart Hunt provided practi-cal tips on implementing RDA in your ILS. Both recommended exploring the RDA Toolkit, if you have access to it.

Look into ways to promote the value of cataloguing, for example, through the High Visibility Cataloguing blog and try to make the most of networking opportunities. (Céline Carty, Karen Pierce, Cardiff University, and Rachel Playforth, British Library for Development Studies).

A massive thank you to the organisers and speakers – I had a wonderful few days up in Sheffield and learnt an enormous amount. I also attended the FRBR training after the conference, which again, I found very useful, and I appreciated the addition of this at the end of the conference.

Our Hybrid History and its Lessons for Today

Anne Welsh, Lecturer in Library and Information Studies, University College London with Katharine Whaite, AHRC PhD candidate, Department of Information Studies, University College London

In the discussions surrounding new international cataloguing standard RDA, the phrases “hybrid catalogue” and “hybrid catalogue record” have come into vogue. To professional cataloguers, whose watchwords are ‘consistency’ and ‘predictability of search terms’, hybridization of the database has the potential to strike fear. Indeed, both RDA-L and AUTOCAT, the main listservs for professional cataloguers, have received mails concerned about the backwards compatibility of RDA, and particularly the number of changes that would require manual updating of records, should libraries desire to convert existing records to RDA.

The good news for concerned cataloguers is that the hybrid catalogue has been with us, we might argue, since the first librarian decided to do something differently from his predecessor without going back and changing all his predecessor’s records to match his own. Certainly it was a concept with which Cutter was familiar, his stated attitude being “If one already has a catalog with a large number of cards, and merely inserts in it as many of the Library of Congress cards as possible, I see no reason for altering one’s own style, either on the past accumulations or on the new cards that one is to write. The two kinds of cards can stand together in the drawers and the public will never notice the difference.” (1)

More recently, in its 2011 *Discussion Paper on RDA Implementation Alternatives* the Library of Congress Program for Cooperative Cataloging Policy observed that “The cataloging environment is already hybrid. OCLC WorldCat includes records created under AACR1, AACR2, RDA and a variety of other international rules. As OCLC continues to pursue global participation, particularly from national libraries, the environment will grow increasingly more diverse.” (2)

Even the hybrid record is not a new phenomenon – a stroke of the pen allowed for amendments to card catalogues, and a quick filing operation could move the card from one drawer to another where heading changes were required. The growth of computer cataloguing and in particular the now common facility to ‘batch modify’ records means that there are few (if any) online catalogues whose every record is exactly as it was entered by the original cataloguer under the rules in vogue at the time: modern technology allows for changes to be made that are less obvious to the end-user. In fact, data sharing itself leads naturally to hybrid records: for modern items, it is most common for only the first cataloguer in a consortium to create an entire record – everyone else downloads and amends this.

Of course, while accuracy must be assumed on the part of cataloguers, there are occasions when not every local practice is stripped out of a consortium record when it is uploaded to the pool, and when not every local amendment is made when it is downloaded.

At the EURIG meeting on RDA in 2010, the British Library’s Alan Danskin himself highlighted the BL as a hybrid catalogue par excellence. In a recent article on UK training needs for RDA, the record for Hosking’s *Some observations upon the recent addition of a Reading Room to the British Museum*, is highlighted as a favourite to use with students, since it includes Panizzi’s own manuscript notes. (3) In it we can see the pre-1968 format, pagination and publication details have been massaged through various data upgrades into the modern MARC catalogue as

> **260 |a**pp. 34. Pl. A-E. Edward Stanford: London, 1858.
>
> **300 |a**fol.

We can see the drawbacks inherent in this record were it to be shared with another database using machine-readable cataloguing with no human intervention (a *caveat* for the semantic web), but we can also see that for the human catalogue user, there is no difficulty in understanding this hybrid record in our modern environment (AACR2 in MARC). This record is not only typical of many legacy records on the British Library system but in comparable large libraries all over the world. In such examples, we are encountering the impact of cataloguer workflow – a concept much neglected in the mainstream literature, but which affects and in some cases *in*fects discussions of cataloguing principles and theory.

Clearly, in the interests of 'predictability of search' for users, consistency of entry is a primary aim for cataloguers. As Cutter puts it in his 'Reasons for Choice',

> Other things being equal, choose the entry
>
> That will be first looked under by the class of people who use the library
>
> That is consistent with other entries, so that one principle can cover all
>
> That will mass entries least in places where it is difficult so to arrange them than they can readily be found, as under names of nations and cities. (4)

Predictability, consistency and disambiguation have been our cornerstones, in theory at least. A lack in any one of these three attributes shifts the burden of intellect and time from the cataloguing agency at the point of documentation to the library user at the point of searching.

Hybrid records occur most commonly where the library has been able to batch modify some of the record, but not all of it. Where we see hybrid records, it can be argued that we are seeing evidence that a policy decision has been made not to manually update every field in every catalogue record affected by changes in the cataloguing standard. Someone within the organization has made the difficult decision that the benefit of such an update is not worth the expense.

The expense of the catalogue is an argument that has grown in the general library literature since the introduction of computers in the 1980s, to the point where, in 2008, Ingrid Hsieh-Yee felt able to state quite baldly that "What has become apparent is that cataloguing as currently practiced in most libraries and other information settings is not a cost-effective solution for managing digital resources." (5) This statement can trace its pedigree back to Karen Calhoun's report for the Library of Congress in 2006 on *The Changing Nature of the Catalog and its Integration with other Discovery Tools*. (6) In her 'Key Findings from Interviews and Literature Review (Appendix C)', Calhoun writes "One interviewee … mused, 'It is difficult to imagine the costs of converting millions and millions of MARC records in thousands of databases around the world to new metadata structures," (7) and later, in a subsection on 'Cataloging Practice' Calhoun asserts "There was some consensus around the position that cataloguing needs to be simpler, faster and less expensive. There is 'too much hand-wringing and worrying about each record: this is extreme and wasteful.'" (8)

Within the main body of the report, this attitude surfaces with a more positive spin: "Today, the online catalog is losing appeal for students and many scholars. Catalog usage, drifting downward compared to other discovery tools, may soon plummet. Fortunately there are ways to use the knowledge that today's catalog has reached the end of its life-cycle." (9) She suggests "An organization's strategic choice [for catalogue renewal] will depend on [its] position with respect to others who supply or produce catalogs, its financial position, its perception of the likelihood and rate of revitalization or decline of the catalog, the actual strength and nature of remaining demand for the existing catalog, the availability of practical alternatives, and the level of difficulty the organization will have diverting its capacity to new uses." (10)

Of course, the role of the catalogue as inventory cannot be overlooked. In fact, from a management perspective, Calhoun argues that "The typical research library catalog's strongest suit is its support for inventory control and as 'last mile' technology to enable delivery of the library's assets to the hands of local users." (11) She calls for the expansion of "the service model of the catalog to cover more of the scholarly information universe – metasearch" (12) and, indeed, we can observe that since the Calhoun report was published in 2008, much of our time and energy as a community has been used in implementing – or supporting our systems librarians to implement – discovery engines to pull together data from our catalogues and the journal, theses and abstracting services created elsewhere.

In reviewing the literature around changes in catalogue standards, two clear distinctions can be made between the current change on which we are embarking and previous changes. The first has been, and continues to be, well-documented, and that is the Internet's impact on bibliographic data production and curation. WorldCat, the Open Knowledge Foundation, and the Open Library (13) are just three examples of initiatives to extract from the deep web and make it accessible to search engine crawlers and, therefore, people using a simple web search.

The second shift is less frequently discussed – and that is the dialogue of expense. Notwithstanding the excellent work of NACO, and the free version of the name authority file made available by the Library of Congress, we cannot help but agree with Clack's observation, made as long ago as 1990, that many libraries find themselves unable to justify the expenditure on authority control, and, in her words, "merely paid lip-service to the concept." (14) Many more have outsourced their authority control to commercial vendors, with positive results, according to Tsui and Hinders, in terms of currency of the authority file, and overall efficiency gains. (15) However, we cannot help but observe that the driver towards outsourcing, in cataloguing tasks as in other areas of library management, is one of cost-effectiveness. Or, put another way, the reduction of expense.

Cost cannot, and, indeed, should not be ignored in any public service, but it should be clear that such considerations are not about general cataloguing principles but about the management of workflows. All of the great writers on cataloguing standards are careful to balance these two things: the need for a clear, concise statement of principles, and a pragmatic set of rules that can be applied day-to-day. In fact, in tracing the history of the creation of our cataloguing standards, we can argue, as William Denton has, that Cutter broke away in style from previous rules, by giving us "the first set of axioms made in cataloging." (16) Unlike Panizzi's work, which consolidated growth at the British Museum and gave it some structure, Cutter began with 'objects' which are then carried out by his 'means.' Denton points out that while Cutter's 'objects' remain, and we can draw a straight line from them to those of our new governing set of principles, FRBR, we can find different means by which to achieve them. FRBR, in effect builds on Cutter's means to form its 'user tasks' and then implements a new data model to satisfy these tasks in the digital age.

We might go further than Denton, in observing the evolution of axiom-based cataloguing. In fact, a survey of the literature covering the transitions from ALA and Library Association Rules to AACR and from AACR to AACR2, and finally from AACR2 to RDA, we observe in fact a revolution or cycle, as firstly easy-to-remember axioms are proposed, then built upon with specific cases and examples and then reworked and pared back to easy-to-remember axioms at the start of the next iteration of cataloguing standards, which themselves over time become padded out with cases and examples before the next iteration of cataloguing standards begins again, with axioms. In fact, each of the major revisions of cataloguing codes can be seen to be a response to a call from cataloguers for clearer, more concise rules. Most famously, we have Lubetzky's *Cataloging Rules and Principles: a Critique of the ALA Rules for Entry and a Proposed Design for their Revision,* (17) which opens with the chapter heading in large font "Is This Rule Really Necessary?" Lubetzky, of course, is largely credited with the first iteration of the Paris Principles, but Michael Gorman's recent autobiography makes for interesting reading, as he credits Lubetzky as his own cataloguing hero, inscribing the copy of AACR2 he gave him "*il miglior fabbro* ('the better craftsman')." (18) The autobiography also describes in some detail the streamlining approach he took to the rules produced by the Joint Steering Committee for AACR. It is, of course, early days, but we can only assume (and hope) that Chris Oliver, the new editor for RDA will adopt a similar editorial line.

Cataloguing workflows are a complex matrix. Leaving behind the consideration of a particular amount of money, and bypassing the wider political debates around library funding in general, it remains the case that even in well-funded times; even in times in which library administrators prioritise cataloguing, there was always, there *is* always, a cataloguing backlog. From Panizzi's pleas to the Parliamentary Committee for more time (which equals money), through to Gorman's account of the backlog at the Department of Printed Books in the 1970s ("about five years' worth of acquisitions"), (19) it is difficult not to empathise with Michael Winship's plea, "I, for one, would prefer short, accurate, and clear records of a library's entire holdings over long, elaborated ones that conceal, and to some extent cause, a tremendous cataloging backlog of inaccessible, and thus generally useless, materials." (20)

Indeed, it was the speed of example-led ALA rules, and the claimed impact on cataloguing backlogs that

'The Crisis in Cataloging': "According to [the ALA Rules's methods], there must be rules and definitions to govern every point that arises; there must be an authority to settle questions at issue. So the reviser sits in judgment on the cataloger, and the head cataloger is the supreme court for his particular library . . . Debate, discussion, and decision eat up a surprising amount of time. Hence the demand in some quarters for a cataloging code that will define or rule on all debatable points." (21)

After reading Osborn's paper, which documents the complicated maze of rules and precedents cataloguers negotiated, it is easy to see the appeal of Lubetzky's opening question "Is This Rule Necessary?" in his *Cataloging Rules and Principles,* and to understand its positive reception.

It is difficult, however, to judge cataloguer response to the early standard changes in any meaningful way. What we can see in the literature is a series of calls for clarification followed, at varying intervals by a response.

So where does this leave us in terms of action points for today? We conclude with an explicit statement of some of the areas discussed in this article. Firstly, *count your costs*. Whether activities are done inhouse, outsourced or abandoned, there is a cost attached. Moving to RDA, sticking with AACR2, there is a cost. And while we talk about cataloguing practices and workflows, the dialogue that talks to power talks money.

There's a lot of acceptance involved in cataloguing today. We need to *accept, firstly, the apathy of much of our user group*. As Martha Yee puts it in her cheerfully-titled and undated thinkpiece for SLC 'Will the Response of the Library Profession to the Internet be Self-immolation?', "Undergraduates [and here we might widen out from Higher Education and just say 'people'] have always tended to over-use ready reference sources until they are taught by both librarians and professors how to do effective research and critical thinking." (22) In evaluating changes in cataloguing standards, we have not in the past encountered frequent, vocal and outraged complaints from library users.

Secondly, *accept that there never was a better time*. As Rebecca L. Lubas has asserted, "Even when we thought that we called the shots, we really did not. Users still failed to understand the difference between finding articles in one place and books in another. The best researchers were promoted to point that they never entered the library themselves but rather had graduate students toil over their searches themselves."

We are at a stage in the cycle of the evolution of cataloguing rules in which *we must all be encouraged to focus on the principles first and foremost*. The details of the rules have changed, but the principles remain constant. The current dialogue has been honest about the emphasis it places on 'cataloguer judgment'. Judgment has always been a core activity within cataloguing. There is a tipping point – and further research would be required to establish when this occurs – in which there are too many specific case studies worked out in our standards – too many examples, and we come to rely on this case study approach as much as we come to despise it. We need to allow ourselves time to build up the confidence to employ cataloguer judgment: our predecessors needed it in the move to AACR.

Finally, if the history of cataloguing standards implementation tells us anything, it tells us that it is actually hard to get right, and certainly impossible to get right first time. *We need to question everything and encourage our staff to question everything*. Materials evolve, cataloguing standards evolve, and each revolution of the cycle begins and ends with the question, "Is this rule really necessary?" (23)

References

1. Cutter, C. A., *Rules for a Dictionary Catalog.* 4th ed., rewritten. Washington DC: Government Printing Office, 1904.
2. Library of Congress Program for Cooperative Cataloging Policy Committee (2011) *PoCo discussion paper on RDA implementation alternatives*. Original release: 4/05/2011, revised 4/15/2011, http://www.loc.gov/catdir/pcc/PoCo-RDA-Discussion-Paper040511.pdf, p. 2.

3. Welsh, Anne, Carty, Celine and Williams, Helen. '"Mind the [Trans-Atlantic] Gap, Please": Awareness and Training Needs of UK Catalogers.' *Journal of Library Metadata* 12(2-3), 2012: 253.
4. Cutter, *Rules for a Dictionary Catalog*.
5. Hsieh-Yee, Martha, 'Educating Cataloguing Professionals in a Changing Information Environment,' *Journal of Education for Library and Information Science* 49(2), 2008: 93.
6. Calhoun, Karen, *The Changing Nature of the Catalog and its Integration with other Discovery Tools. Final Report.* Washington DC: Library of Congress, 2006, http://www.loc.gov/catdir/calhoun-report-final.pdf
7. Calhoun, p.32.
8. Calhoun, p.33.
9. Calhoun, p.10.
10. Calhoun, p.12.
11. Calhoun, p.9.
12. Calhoun, p.9.
13. The Open Library, http://openlibrary.org/
14. Clack, Doris H., *Authority Control: Principles, Applications, and Instructions.* Chicago: ALA, 1990, p. 1.
15. Tsui, S.L. & Hinders, C.F., 'Cost-Effectiveness and Benefits of Outsourcing Authority Control.' *Cataloging & Classification Quarterly*, 26(4), 1999: 43–61.
16. Denton, William, 'FRBR and the History of Cataloging' in A.G. Taylor (ed.), *Understanding FRBR: What it is and how it Will Affect our Retrieval Tools.* Libraries Unlimited, 2007.
17. Lubetzky, Seymour. *Cataloging Rules and Principles: a Critique of the ALA Rules for Entry and a Proposed Design for their Revision*. Washington DC: Library of Congress, 1953.
18. Gorman, Michael, *Broken Pieces: a Library Life, 1941-1978*. Chicago: ALA, 2011, p. 201.
19. Gorman, p. 147.
20. Winship, Michael, 'What the Bibliographer Says to the Cataloger'. *RBMS: a Journal of Rare Books, Manuscripts, and Cultural Heritage* 7(2), 1992: 102.
21. Osborn, Andrew, 'The Crisis in Cataloging' in Michael Carpenter and Elaine Svenonius (eds.) *Foundations of Cataloging: a Sourcebook*. Libraries Unlimited, 1985: 92-103.
22. Yee, Martha. 'Will the Response of the Library Profession to the Internet be Self-immolation?' SLC, n.d., http://slc.bc.ca/response.htm
23. Lubas, Rebecca L., 'Managing the Challenge of the Twenty-First Century Cataloging Department' in Rebecca L. Lubas (ed.), *Practical Strategies for Cataloging Departments*. Libraries Unlimited: xiii.

Getting metadata to work harder: re-use, standardisation and streamlining, a data archive perspective

Lucy Bell **UK Data Archive, University of Essex**

Introduction

The UK Data Archive (the Archive) has been based at the University of Essex since 1967. It curates the largest collection of digital data in the social sciences and humanities in the UK, holding several thousand datasets relating to society, both historical and contemporary. It makes these available via its services: UK Data Service (1) (from October 2012); and, previously, the Economic and Social Data Service (2). It is also a place of national deposit for The National Archives.

The Archive provides access to over 5000 digital data collections, with more being added daily. All of these items are catalogued at study level, and many at variable level using the *de facto* standard data cataloguing schema, DDI (3) (Data Documentation Initiative). Currently, the Archive uses DDI 2.1 (now known as DDI-C, for *codebook*) but it is investigating a move to DDI 3.1 (or DDI-L, for *lifecycle*). The Archive also uses controlled vocabulary tools in its cataloguing activities, in particular the Humanities and Social Science Electronic Thesaurus (4) (HASSET) and a series of internally-controlled authority lists and CVs.

This paper describes recent changes to the way in which metadata are managed within the UK Data Archive, in the context of its long history.

HASSET

HASSET is a multidisciplinary thesaurus, primarily developed to support and index the Archive's collection. Its coverage is in the core social science subject areas and it uses standard hierarchical relationships: TT (top term); BT (broader term); NT (narrower term); RT (related term) etc. HASSET's role in the Archive is twofold:

- it is used internally for indexing studies and series (at variable level);
- it is also a separate product which is licensed to others.

Significant recent metadata/indexing developments

The last 2.5 years have seen a series of significant developments in the Archive's resource discovery activities:

- May – October 2010: a review was carried out of the Archive's resource discovery tools;
- 2011: a project was started to apply the review's results to the existing resource discovery applications;
- 2011 onwards: work was started to investigate the move from the DDI-C to DDI-L;
- June 2012 – March 2013: SKOS-HASSET, a JISC-funded project was undertaken to apply SKOS to HASSET and to test its automated indexing capacity.

It soon became clear that, despite working in different, if over-lapping areas, all of these initiatives were all pointing at one thing: the need for more controlled – and harder-working – metadata.

Resource discovery review

The resource discovery review found that trends in information-seeking behaviour show that academic users prefer simple, Google-like interfaces, but which still return acutely-focused and highly-relevant results. The look and feel of the interfaces should be simple but the results must achieve academic rigour.

The world of resource discovery is in a state of transition. Ten years ago, to interrogate an online database effectively, users would have been encouraged to consult a thesaurus in advance, collating all relevant keywords and constructing a complex search string. Today, most users are familiar with approaching a single search box and entering any terms of their own choosing. Convenience has risen in importance in the information-seeking process:

‘Ease of discovery and access in getting to the information resources relevant to their needs, and in keeping themselves informed of events and publications in their fields, is critically important for researchers.’ (5)

This quotation from the RIN report on researchers’ approaches to Web 2.0 highlights the need for speed and simplicity in gaining access to data and other information; however, the report also indicates the scepticism with which Web 2.0 tools, such as blogs and wikis, are viewed within the academic community. The expectation is still there that academically-professional resource discovery tools will contain quality-assured, sustained and, in the case of journals, peer-reviewed content.

Internet search engines are simultaneously liked and mistrusted. The 2006 RePAH report highlights academics’ views:

‘The internet search engine emerges from this study as an immensely useful digital resource-discovery tool … [Its] simplicity and speed appealed to our users, for whom a key determinant in their cost-benefit analysis of resource-discovery tools was whether it saved, rather than cost them, time. At the same time, users were also aware of the limitations of their internet search-engine of choice. Our users told us of their frustration at its lack of sophistication. They were suspicious of its ranking of hits returned. They were overwhelmed by the information redundancy which often accompanies its results.’ (6)

A worry lurks that something key will be missed. Although academics are good at scanning through large lists of references, they do not wish to resort to non-academic tools to obtain these lists. A dichotomy exists then: many users from the academic community would like, indeed, expect, fast and easy search interfaces, resembling Google; however, they also insist on their discovery systems being of a very high standard. This emphasises the need for extremely high-performing metadata. The JISC-commissioned meta-analysis of user behaviour projects supports this:

‘Users are beginning to desire enhanced functionality in library systems. They also desire enhanced content to assist them in evaluating resources. They seem generally confident in their own ability to use information discovery tools. However, it seems that information literacy has not necessarily improved. High-quality metadata is thus becoming even more important for the discovery process.’ (7)

Creators of resource discovery systems are at a turning point in development. Their systems must satisfy the users’ needs for the delivery of high-quality, academically-rigorous content, while also providing the most straight-forward interfaces possible. The role of metadata in this scenario becomes paramount as users enter a world which seems as serendipitous as possible, while still maintaining academic rigour, relevance and comprehensiveness. For data services to produce simple interfaces, which still return highly-relevant results, metadata are required which are both extremely powerful and, ironically, increasingly invisible.

The vision: use controlled vocabularies to enhance the user’s experience

Following on from the results of the review, it was decided that the Archive should develop a single search interface which allows the user to move seamlessly from one type of resource to another both via faceted browsing and directly from within each resource type. This required cross-referencing data collections with publications, research outputs, support guides and case studies, all using metadata. It also entailed the use of many controlled vocabularies. The resulting application was a faceted search/browse interface which was released in beta on 1 October 2012. (8)

Although the functionality of the new system was complex, it turned out that one of the biggest obstacles to overcome was the requirement to populate the facets with meaningful and consistent terms. Some facets were already in a fit state, using metadata elements which were already well-structured and controlled; however, because of the way in which cataloguing principles and procedures had changed internally over 40 years, several were populated with freetext (observation unit <anlyUnit>, spatial unit <geogUnit>, kind of data <dataKind>, time dimensions <timeMeth>). It was decided that the first three of these facets were essential for the functionality of the search application. Their content, built up over 45 years of cataloguing, contained a wide variety of uncontrolled entries; the first job was to map these to a series of controlled vocabularies.

The same procedures were used for all four elements:

- a dump of the metadata was obtained from SQL and imported into Excel;
- the CV to be used was identified;
- the metadata were exported into Google Refine, taking care to retain all the UIDs;
- Google Refine was used to identify existing, similar, freetext entries and, where appropriate, align them;
- the clean metadata were re-imported into Excel from Google Refine and mapped to the existing freetext entries (sometimes at item level, using the UIDs or, wherever possible, at CV value level).

The work to clean these four elements took place in the summer of 2012 and took two staff members, working c.0.4 FTE, four months to clean three elements.

In all three of these cases an existing CV was identified which could be used. A previous, JISC-funded, UK Data Archive project entitled U.Geo (9) had created a spatial unit CV, containing 194 entries, covering spatial units as small as Census Output Areas and as broad as Countries. Previously, the Archive's catalogue boasted 653 unique values in its spatial unit field. These have now been mapped to the list of 194.

The DDI's Controlled Vocabularies Group (10) had also already established a CV for Unit of observation. The 11 values this contains are as follows:

- Individuals
- Organizations
- Families/households
- Housing Units
- Events/Processes
- Geographic Units
- Time Units
- Text units
- Groups
- Objects
- Other

The Archive's 183 unique values have now been mapped to these 11.

Finally, the kind of data element has also been mapped to the DDI CVG's existing list:

- Alpha-numeric
- Audio
- GIS
- Image
- Numeric
- Textual
- Video

A total of 294 unique values have been mapped to these seven. It is expected that more work will take place on the kind of data typology, however, looking less at data format and more at data types (longitudinal, qualitative etc.).

Taking CVs further for organisational efficiency: automated indexing via SKOS-HASSET

The UK Data Archive is also working on a JISC-funded project entitled SKOS-HASSET (11) (June 2012 – March 2013), funded under the JISC Research Tools Programme.

SKOS-HASSET has three aims:

- apply SKOS to HASSET – making the thesaurus more flexible
- improve its online presence
- test its automated indexing capabilities using four quite distinct corpora:
    - ◊ questions
    - ◊ questionnaires
    - ◊ abstracts
    - ◊ publications

The project is taking a single controlled vocabulary, HASSET, and enhancing it via SKOS to make way for more efficient cataloguing practices, organisationally. It will also improve the users' experience by providing a new and flexible way to interact with HASSET as a product.

SKOS is a language designed to represent thesauri and other classification resources. It encodes these products in a standardised way using RDF to make their structures comparable and to facilitate interaction. Applying SKOS to HASSET will mean both that its terms will be more easily available to other services and also that, potentially, HASSET itself will be in a better position to be updated via the inclusion of new hierarchies.

This project will also test SKOS-HASSET's automated indexing capacity. SKOS-HASSET has been taken as the terminology source for an automatic indexing tool and applied to question text, abstracts and publications from the Archive's collection. The results have been compared to the gold standard of humanly-undertaken indexing.

SKOS has been applied to HASSET and the work to evaluate the automated indexing is underway. The project is due to complete in Spring 2013 and more information will be made available then on its results.

Conclusion

The past two years have seen the roll-out of a series of developments within the Archive which have, themselves, highlighted further issues, the resolution of which have, in turn, led to wider improvements in the services that the Archive maintains.

From the users' perspective, the faceted search/browse interface exposed a lack of standardisation in the underlying metadata. Freetext terms had been used for over 45 years; these are now being standardised in order to apply consistency to the indexes and to create efficient browsing opportunities for the users.

The application of more CVs will, additionally, provide organisational efficiencies. It will allow the Archive to streamline its deposit process by including more controlled terms in any communications with depositors. More CVs for elements other than just subjects/keywords, such as 'Kind of data' and 'Unit of observation', also result in fewer decisions being made when cataloguing. SKOS also provides the opportunity to work more flexibly with the thesaurus: automated indexing using CVs may result in suggested terms being offered to cataloguers.

which could speed up the process.

In the world of information retrieval, metadata is still essential in order to undertake efficient searches. Today, however, there is often too much material for information professionals to create all the metadata themselves in time. Coupled with this, users often expect applications to perform as well as they would have done ten years but without having compiled a search strategy in advance.

Information providers need to work within these constraints and make clever use of technology to point users to suitable items. The JISC Intrallect report contains a quotation from Vic Lyte, which sums this up nicely; Vic compares the use of a search engine search box with a human conversation:

'a new researcher wishing to approach scholarly inquiry to determine the impact of global warming on penguin populations in South Antarctica doesn't walk up to a Librarian and shout 'Penguins'.' (12)

What the developers of digital information resources need to do today is to ensure that, on hearing that cry, their systems respond to the users with intelligent, metadata-driven suggestions to satisfy their searches.

References

1. http://www.data-service.ac.uk
2. http://www.esds.ac.uk
3. http://www.ddialliance.org/
4. https://www.esds.ac.uk/search/hassetSearch.asp
5. Proctor R, Williams R and Stewart J. *If you build it, will they come? How researchers perceive and use Web 2.0: a Research Information Network report*. Research Information Network: London, 2010.
6. Brown S, et al. *RePAH: a user requirements analysis for portals in the arts and humanities*. Leicester: De Montfort University, 2006.
7. Connaway L S and Dickey T J. *The digital information seeker: report of findings from selected OCLC, RIN and JISC user behaviour projects*. HEFCE: London, 2010.
8. http://beta-discovery.data-service.ac.uk/
9. http://www.data-archive.ac.uk/about/projects/ugeo
10. http://www.ddialliance.org/controlled-vocabularies
11. http://hassetukda.wordpress.com/
12. Duncan, C. & Douglas, P., (2009). *Automatic metadata generation: use cases and tools/priorities*. Intrallect (for JISC): 2009

The fundamental interconnectedness of all things: the impact of networked knowledge systems on cataloguing

Simon Barron **E-Resources Co-ordinator, Durham University**

What is cataloguing? What is it that we cataloguers do?

This is *"the dawn of a new era in cataloguing"* (1). In these times of change, it's good to get back to first principles and ask the philosophical questions. What are we doing here? What is it that we do? What is our value? What is cataloguing? (2)

If you asked a hundred cataloguers, you'd get a hundred answers. One view has it that cataloguing and classification is the process of describing the bibliographic universe. Librarians and information professionals are familiar with the world of books, journals, information, data. We know about the secret web of connections and links and citations and themes and genres that connect the millions of information sources that we deal with every day. We have chosen to live in this bibliographic universe: to wander its streets together; to climb its mountains; to congregate in its squares; to see what lies down every dark alleyway; to explore every inch of it. And among information professionals, cataloguers and classifiers are the ones who have chosen to map the bibliographic universe. We are cartographers of the abstract. We chart the world of books, journals, and information: describing what we see, encoding it in a usable form, and sharing the results with our users and with each other. This abstract mapping is the true value of cataloguing.

Nowadays we have new tools at our disposal to do this. FRBR – Functional Requirements for Bibliographic Records – is one such tool to describe the bibliographic universe. It's one attempt to define the arrangement of the abstract entities that information professionals work with. It defines the relationship between these things: between the 'work' as envisioned by the author all the way down to the physical 'item' which a user can hold and touch. As well as defining the relationships between different versions of the same 'work' – between Group 1 entities – it defines the relationships and the links that works have with people and corporate bodies – Group 2 entities – and through this, defines their relationships with each other. FRBR is a framework on which to base our maps of the bibliographic universe. That is its abiding value.

RDA is built on a foundation of FRBR and will be another useful tool. It places a new emphasis in cataloguing on *"clustering of bibliographic records"* and using metadata to define the relationships between works. Previously the relationships between one edition of a book and a later edition of the same book or between the print version of a book and the ebook version have been somewhat ill-defined if not totally unexplained by a catalogue record. FRBR and RDA are tools to help accurately describe the universe of information and so they're both heavily informed by epistemology and ontology: two separate but linked branches of philosophy.

Epistemology is the study of knowledge systems: what knowledge is, how it's arranged, and how we can have it. It's been studied from the time of the Ancient Greek philosophers – from Plato and his pupil Aristotle – through the Enlightenment philosophers – Descartes, Locke, Hume, Schopenhauer – to the present day when the debate continues and has been renewed by new scientific discoveries and what seems to be an ever-expanding world of knowledge. The World Wide Web has emerged as a quasi-physical embodiment of our abstract world of information – our realm of knowledge – and so the debate is more physical, more real, and more important than ever. It's said that Socrates was in constant communication with a 'daemon' who supplied him with all his ideas and inspiration: today, we can all communicate with hundreds of people every day who give us fresh ideas and invite us to interesting events. The Web and these changing paradigms of communication have changed our view of epistemology. Ontology is the study of being and existence. It's relevant to considering the bibliographic universe in terms of trying to define that universe's metaphysical status. 'What kind of entity is information?' and 'What kind of thing is knowledge?' Ontology tries to define what things are: is the text of a book a purely mental construct or does it have some kind of physical reality? Can a 'work' be said to exist in the same way as a chair?

Partly because of the introduction of FRBR and RDA, epistemology and the ontology of knowledge are of central importance in modern cataloguing, indexing, and classification. Because FRBR assumes the existence of a bibliographic universe with some ontological status and it's the predominant intellectual trend in cataloguing, we go along with that. We need to consider what shape knowledge has, how it's arranged, and how we can accurately describe and represent this for our users.

For centuries, knowledge has been represented as a hierarchy and this has informed the traditional classification systems that are in use in librarianship and bibliography today. Dewey, Library of Congress, LCSH: they're based on ideas of hierarchy and taxonomy; of dividing and subdividing subjects like the branches of a tree. The conceptualisation of knowledge, in particular the 'tree' metaphor, has a long history.

One of the first, if not **the** first, representations of knowledge is in the Book of Genesis: God provides the first humans, Adam and Eve, with the 'tree of knowledge'. After that, one of the first real articulations of the concept of hierarchical knowledge comes from a library – from someone who was trying to work out what knowledge looks like so that he could organise his books. Aristotle, the great philosopher, had the largest personal library in Athens and to organise his collection accurately he envisioned knowledge in his work the *Organon* as a hierarchy based on the now-familiar principles of taxonomy and categorisation. His 'tree of knowledge' concept became codified as information theory developed and there are numerous examples stretching from Ancient Greece to the 20th Century. Linnaeus' classification of the natural world in his *Systema Naturae* divides things by genus and species and subdivides into nested groups. In 1605, Francis Bacon published *The Proficience and Advancement of Learning, Divine and Human* which divides all knowledge into History, Poetry, and Philosophy which were then subdivided and so on into different branches. In 1783, Thomas Jefferson catalogued his collection of books – a collection that would go on to start the Library of Congress. Jefferson divided the world of knowledge, similarly to Bacon, into Memory, Reason, and Imagination broadly corresponding to History, Philosophy, and Fine Arts. There are hundreds of other examples – Diderot's system for the *Encyclopédie*; John Wilkins' 40 Universal Categories – and we have other examples in practice in classification.

The classification schemes that we still use in libraries are heavily influenced by hierarchical thinking. Enumerative classification schemes – Dewey, Library of Congress, Cutter's Classification – explicitly *"treat knowledge as if it were a unity which can be subdivided into smaller and smaller units. At the top of the tree is the whole universe, which is divided and subdivided to arrive at all the different entities, events and activities represented in the subjects of books."* (3) Faceted classifications and analytico-synthetic classifications, though more flexible, also exhibit an essentially hierarchical structure with the small building up to form the large. The tree of knowledge – our centuries-old conception – continues to inform our epistemological systems and our thoughts on the ontological status of knowledge. Broadly speaking, our current maps of the bibliographic universe look like trees.

Now we've rethought this conception of knowledge as a tree and are starting to think of different knowledge systems. A new model – a new intellectual paradigm – is emerging. It's the idea of knowledge as a network rather than a tree: a web of interconnections between ideas, concepts, theories, data.

A network can be defined as a system of interrelations: *"individuals function as autonomous nodes, negotiating their own relationships, forging ties, coalescing into clusters. There is no "top" in a network; each node is equal and self-directed."* (4) As science and philosophy have advanced and the universe of human knowledge has grown, we've discovered connections and interrelations between things that seemed totally unrelated. It turns out that the branches of the tree of knowledge are all connected in different ways. Everything is connected. The universe appears to be holistic in that everything depends on everything else. We're beginning to see, in the words of the great detective, Mr. Dirk Gently, *"the fundamental interconnectedness of all things"* (5). The abstract world of knowledge turns out to be more complex – far more complex – than a tree shape and the more appropriate visualisation is something like a web or, better yet, a rhizome seed.

In a 2010 paper, Lyn Robinson and Mike Maguire of City University adopt Deleuze and Guattari's image of a rhizome as the better metaphor for information organisation (6). A rhizome is essentially a root: an underground mass of shoots and stems that grow in unpredictable ways in complex, laterally branching networks with different nodes shooting off in different directions. Deleuze and Guattari use it as an *"image of thought"* (7) which represents complex networked knowledge systems. Robinson and Maguire's paper is well worth reading as an excellent discussion of the changing concepts of knowledge structures.

Broadly speaking, we are moving from the tree to the rhizome. And we can see this shift towards networked systems in a range of subjects and different areas. In physics, chaos theory tells us that everything is linked: that one tiny imperceptible event can cascade to significant consequences in a seemingly random and impossible-to-predict way that is nonetheless based on cause and effect in a networked system. In social life, we readily talk about social networks, recognising that human relationships can be mapped onto a network with each person connected to every other person: Stanley Milgram's small world theory tells us that this can be done with a maximum of six degrees of separation. In technology, computer networks surround us, transferring data along connections between computers and servers and routers. They form the conceptual foundation for the Internet and the World Wide Web.

In academia, we're recognising the importance of the citation network – a network of references to and from various papers, journal articles, books. You may have heard of the mathematician Paul Erdős. His work was so prolific that any mathematician working today can be connected through citations to Erdős: an estimated 90% of mathematicians are connected to him through no more than eight links. (8)

Of these examples of networked systems, the citation network most closely relates to the networked systems of knowledge which are important for cataloguing and classification. We're recognising that knowledge can't be neatly divided into hierarchical categories and that in the bibliographic universe everything is connected in strange and sometimes complex ways.

For an example, consider Ludwig Wittgenstein's *Tractatus Logico-Philosophicus*. When we catalogue this book – here's the catalogue record for the book at Durham – off-hand we'd say it's a philosophy book – it's one of the cornerstones of modern formal logic – and at Durham, we classify it at 192 for Modern Western Philosophy of the British Isles but we could also stick it somewhere in 160 for logic, or, depending on how much you consider its implications, somewhere in 110 for Metaphysics.

That's straightforward hierarchical classification and it puts the book neatly into a distinct place on a shelf but it's not the whole truth. This doesn't represent the links that the book has with everything else in the bibliographic universe. What about its links to science, language, mathematics, and possible worlds theory? What about the links to Wittgenstein's other works? His other masterpiece, *Philosophical Investigations*, is a whole different genre of philosophy and refutes bits of the *Tractatus*: the two are nonetheless conceptually linked. What about the books written about this book: the different theories; the different interpretations; the books that owe their existence to this book? What about the *Prototractatus*: the original manuscript version written in the trenches of World War I? What's the relationship there: is it the same work or not? Whatever the answer, there is some kind of strange link. What about the different translations: the standard is the Pears and McGuinness translation but what about the German original, the versions without Bertrand Russell's introduction, the Ogden translation? What about the links to the fiction inspired by this book?

Even something as simple as this 80 page book is connected through a thousand interrelations to myriad other books and other nodes in the bibliographic universe. When we look at it closely and think about it, this book is a centre of a web – of a rhizome – connected to intensely different books, journal articles, people, and ideas. If it's our job as cataloguers to describe the bibliographic universe accurately and represent it as truthfully as possible, then we need to think about how to represent these connections. A MARC-encoded, AACR2-standard catalogue record doesn't do justice to the complex web of connections and interrelations that surround this book. Or any of the other books, journals, ebooks, ejournals, and other publications that exist in our libraries.

This is one of the central issues in cataloguing today. How do we represent networked knowledge systems and adjust our practices accordingly? Electronic resources are growing in importance in librarianship and are fundamentally arranged in a network. We're all going to be interacting with information arranged in networks and we should be thinking about mapping the digital world. Thinking about networked knowledge systems is an important consideration for doing this. So how do we catalogue in a network? It's an open question but broadly speaking, we need new practices, new technology, and new thinking.

In terms of cataloguing practice, RDA isn't necessarily the answer to all the riddles but it's a definite step forward. RDA is based on FRBR and therefore has a footing in ontology and serious thought about the bibliographic universe's structure. RDA as a new practice will help us to think about the connections between items, to look at things in a new way – for old and new professionals alike – and to better appreciate that information exists in a rich, complex, shifting epistemological network. How do we actually catalogue to reflect this? Do we use more access points? Do we index more fields? Do we add more fields in the 700s or do we need to more fully define relationships using 500 note fields? RDA is the biggest change to cataloguing in 30 years and so hopefully its implementation will give us the opportunity to consider some of these issues and perhaps rethink how we view our collections.

We also need new technology in cataloguing. Our modern epistemology – this vision of a networked universe with everything connected – is beyond the scope of our current technology for cataloguing and data representation. Though there are interesting things going on with e-resource management and linked data and things like that, these haven't really affected day-to-day cataloguing which is still based on flat, hierarchical MARC records. MARC needs to be replaced and the replacement needs to be able to show relationships more clearly, needs to help users to find information within a bibliographic network, and needs to make use of the links that integration with other software and other systems can provide.

The development of new linked data technology can help with this. OWL and other web ontology languages can help us to define domain-specific ontologies (9). RDF is a language that helps to define classes, sets, and relationships within an ontology and also has the potential to be utilised for accurate description of bibliographic systems. Linked and open data – the development of Web 3.0 – will help us to map the digital frontier and make it into a true mirror of our abstract knowledge systems.

And then there are data visualisation tools which can take metadata and turn it into something more visual and usable. The UK Institutional Repository Search produced by Mimas in Manchester can produce a basic visualisation of search results and the networked links between them. The results from a search term are grouped in different colours by subject – economics, technology, biology – and you can move them around and click on different nodes to produce more results similar to the ones you've clicked on. The more you click, the more complex the network becomes. This is a beta code powered by Autonomy software: it gives a demonstration of what can be done with data visualisation and how it can be used as part of digital humanities.

Most importantly – more importantly than practice or technology – we need new thinking in cataloguing. We need to think about networked knowledge systems and move on from the hierarchical bibliographic philosophy that has dominated librarianship and information management. Instead of Linnaeus and Dewey, we can look to d'Alembert, Paul Otlet, Vladimir Vernadsky: all of whom advocated networked knowledge systems of one form or another. Crucially, we need to think about what networked cataloguing can achieve. Cataloguing is a way to map the bibliographic universe and in the act of mapping, we can bring subjects together and see the intellectual landscape more clearly. The biologist, Edward O. Wilson, uses the term 'consilience' to refer to the unification of knowledge (10): the belief that different academic disciplines don't represent completely different domains but are part of a single ontology. One knowledge system. One network encompassing everything. *"...a maze of mazes, a sinuous, ever growing maze which will take in both past and future and will somehow involve the stars."* (11) Consilience encourages interdisciplinary research and bringing together seemingly disparate intellectual strands to form a single map of the world of knowledge. Having researched consilience for both my undergraduate and postgraduate dissertations, I conclude that this kind of synthesis will be a major intellectual trend in the 21st Century. Networked cataloguing is one way to achieve consilience and here modern librarians can make a real impact.

Cataloguing and indexing a networked knowledge system requires changes to our practice, our technology, and our thinking. RDA, FRBR, new ontological languages, linked data, digital humanities, and ever-developing software are helping to bring these changes but we as cataloguers need to embrace them. We need to encourage and accept the change. We need to start thinking in networks.

In this paper, I have argued that the most accurate – the most real – depiction of knowledge and the bibliographic universe is in the shape of a network. However I'm aware that I and the prevailing intellectual trend could be as wrong and misguided as we now believe the hierarchical theoreticians to be. The Argentinean poet and one-time librarian, Jorge Luis Borges, wrote that *"…obviously there is no classification of the universe that is not arbitrary and speculative. The reason is quite simple: we do not know what the universe is."* (12) His words remind us that all our human schemes for arranging knowledge are provisional and potentially deluded. Learning and discovery is a process of continuous development and who knows what we'll discover on the journey towards consilience and networked knowledge systems? In the words of Socrates, the only thing I know for certain is that I know nothing.

References

(1) Welsh, A., and Batley, S., 2012. *Practical cataloguing: AACR, RDA and MARC 21*. London: Facet Publishing, p. 172.

(2) Possible answers to the question 'What is cataloguing?' are either describing the bibliographic universe or, depending on your thoughts about the ontological status of knowledge, creating the bibliographic universe. In other words, applying order where none actually exists. That discussion gets into deeper questions about the ontological reality of the universe and is beyond the scope of this paper.

(3) Broughton, V., 2004. *Essential classification*. London: Facet Publishing, p. 32.

(4) Wright, A., 2007. *Glut: mastering information through the ages*. London: Cornell University Press, p. 7.

(5) Adams, D., 1988. *Dirk Gently's Holistic Detective Agency*. London: Pan Macmillan.

(6) Robinson, L., and Maguire, M., 2010 'The rhizome and the tree: changing metaphors for information organisation', *Journal of Documentation*, 66 (4), pp. 604 – 613.

(7) Deleuze, G. and Guattari, F., (B. Massumi, tr.) 1987. *A thousand plateaus: capitalism and schizophrenia.* Minneapolis: University of Minnesota Press.

(8) The source for that statistic is the Wikipedia page for Paul Erdős: http://en.wikipedia.org/wiki/Paul_Erd%C5%91s (accessed 11 October 2012).

(9) The word 'ontology' is used here in a slightly different but conceptually linked sense to the philosophical use.

(10) Wilson, E. O., 1998. *Consilience: the unity of knowledge*. London: Little, Brown and Company.

(11) Borges, J. L., 'The Garden of Forking Paths' in Borges, J. L., (A. Hurley, tr.) 1998. *Collected fictions*. London: Allen Lane, p. 122.

(12) Borges, J. L., 1942. 'John Wilkins' Analytical Language' in Borges, J. L., (E. Weinberg, ed.) 2001. *The total library: non-fiction 1922-1986*. London: Penguin, p. 231.

RDA at ALA, or, Attack of the acronyms

Céline Carty **University of Cambridge**

Earlier this year, I attended the American Library Association (ALA) Annual conference in California thanks to a conference/travel bursary from the John Campbell Trust. This is a time of such fast-paced and wide-ranging change in the field of cataloguing that I wanted the opportunity to hear all the latest developments in RDA (Resource Description and Access). Over five days, I went to hours of meetings, talks, presentations and discussions about RDA. This article, like my talk during the RDA Forum at the CIG Conference, aims to distil some of the information I learned at ALA.

The cataloguing landscape is changing constantly as RDA itself is in flux and so we are trying to move towards a new cataloguing standard on ever-shifting ground. In my CIG talk and in this article, I aim to point to places where the latest information and developments can be found as whatever I say is likely to be out of date quite quickly. For example, at ALA in June the JSC (Joint Steering Committee for the Development of RDA (1)) reported that the work on the re-wording of RDA was well underway and they were reviewing the first five chapters re-worded by the new editor, Chris Oliver. These five re-worded chapters should appear in the RDA Toolkit before the end of 2012. At the time of writing (November 2012), the JSC have just held their latest meeting, during which they have made decisions about a large number of proposals to modify or clarify RDA, including many from the British Library and EURIG (European RDA Interest Group). The decisions from the November meeting will not make their way in to the Toolkit until spring 2013 but an informal report is already available from the ALA representative John Attig (2).

Given all these changes, it is no surprise to learn that the RDA Toolkit has introduced a regular pattern of updates, with new releases about 8-10 times a year (on the second Tuesday of the month). For those people with access to the Toolkit, there is a virtual user group to introduce new features to subscribers and also “RDA Toolkit Essentials” webinars held 3-4 times a year which are very useful for training (3) For those without access to the online Toolkit, the updated version of the print RDA will also published in December 2012.

Over the last few years, MARBI (Machine-Readable Bibliographic Information Committee) has been developing MARC, with a swathe of new fields, additional subfields and adjustments to accommodate RDA (4). At ALA, MARBI reported that the pace of change had started to slow down, with just a few new fields and subfields in preparation for implementation, notably the new 264 field, Production, Publication, Distribution, Manufacture and Copyright Notice. However, all this activity could be seen as the death throes of MARC – shortly after the ALA conference, it was announced that MARBI would be disbanded after June 2013, to be replaced by a new Metadata Standards Committee at ALA, which tellingly said that MARC was “not expected to be its prevailing focus” (5).

The demise of MARBI – heralding the much-anticipated death of MARC – seems to be an inevitable outcome of the implementation of RDA. The Library of Congress (LC) has set up the Bibliographic Framework Transition Initiative (6) to look at a transition away from MARC. In June, Eric Miller of Zepheira, the consultants who had just been engaged to work on this, reported on their roadmap: they aim to review related initiatives, develop prototypes for testing and by the end of 2012 have convened advisory and test groups (7). The presentation was made available online in October.

At ALA, LC outlined their training programme, with online modules and exercises all made freely available on their new RDA website: http://www.loc.gov/aba/rda. The Program for Cooperative Cataloging (PCC) also reported on the vast amount of work done to prepare for RDA training and implementation by their many RDA Task Groups. This is all documented on their website but particularly useful is the Cataloger’s Learning Workshop (8) (hosted by LC) which hosts example RDA records for various formats (9) and also acts as a clearinghouse for PCC-approved free online training materials, which was all updated in early November 2012 and is invaluable for anyone wishing to teach themselves or others about RDA cataloguing.

A few final thoughts from my conversations at ALA in June and from discussions at CIG: there is already such a huge amount of training and information online (10), there is no need to reinvent the wheel, and in many cases something already exists that can be re-purposed. There is much we can learn from our colleagues, particularly since the US is still ahead of UK in terms of practical experience with RDA. There was agreement among those with experience of implementing RDA in the US that the FRBR terminology and philosophy were crucial to reading and understanding the RDA instructions. The feeling at ALA was very pragmatic, talking about “when” not “if” and this is something that carried over into discussions at CIG too. Things seem more certain than they did six months ago, even. However, the catchphrase at ALA seemed to be all about “evolution not revolution”, things are changing but gradually and building on existing practices in many cases. Finally, both ALA and CIG conferences this year really brought home the value of the wider cataloguing community in these unsettling times: the work that cataloguing colleagues are doing on RDA and making freely available is a huge benefit to all.

References

1. http://www.rda-jsc.org/
2. John Attig’s personal blog, http://www.personal.psu.edu/jxa16/blogs/resource_description_and_access_ala_rep_notes/
3. http://www.rdatoolkit.org/essentials
4. http://www.loc.gov/marc/RDAinMARC-7-12.html
5. http://connect.ala.org/node/186145
6. http://www.loc.gov/marc/transition/
7. Eric Miller repeated his ALA presentation to the staff of the Library of Congress later in the summer of 2012 and a recording of this is available online: http://www.loc.gov/today/cyberlc/feature_wdesc.php?rec=5605
8. http://www.loc.gov/catworkshop/RDA%20training%20materials/index.html
9. http://tinyurl.com/RDAexamplerecords
10. The webinars by ALCTS (Association of Library Collections & Technical Services) at ALA are made freely available online after 6 months, for example: http://www.ala.org/alcts/confevents/upcoming/webinar/cat/rda

RDA in your ILS

Stuart Hunt **University of Warwick**

Libraries are serious about the adoption of RDA. The decision has been made and it is no longer a matter of "if" but of "when". With the implementation date of 31st March 2013 clearly indicated by the British Library and the Library of Congress, catalogue managers, and cataloguers, need to be planning their move to this new standard. For the library this has to be done within the context of the integrated library system, or library management system. This article attempts to provide a checklist for factors to address for the implementation process.

Defining the context

There is a considerable gulf between what is, in potential, possible with RDA and in the reality of its implementation. As FRBR-based content standard, RDA accommodates the group one entities of work, expression, manifestation, and item. However, the RDA implementations that face libraries will not, contrary to RDA's hospitality, be FRBR-based. The integrated libraries systems, or library management systems, deployed in libraries continue to be bound by a flat record structure that has no accommodation for the hierarchical model of FRBR. The FRBR, or pseudo-FRBR, implementations that we currently see in libraries are concerned with discovery interfaces, where some attempt to accommodate the group one entities has been made, albeit at a relatively superficial and constrained level. Rather systems will be implementing RDA within a MARC-based environment, or into non-MARC systems.

As a consequence of this lack of hospitality to the FBRR model, library systems will not be implementing new record types (for example WEMI records) but will, instead, be limited to introducing new fields within existing data and data structures. This will apply both to MARC-based and non-MARC systems. Thus we see, for example, the introduction of the 33X block of fields, for example, within the MARC21 bibliographic format.

The lesson for libraries, therefore, is to make sure that their systems are ready to accommodate these new fields, rather than new record types.

Things to consider

It is possible to identify six fundamental areas for libraries that need to be considered when implementing RDA within their systems.

- Loading

The local library system needs to be able to load both bibliographic and authority records. The system needs to accept the new RDA fields that are present in the records. Typically library systems use load tables that specify what the system will do with records on loading. Latest generation systems are able to accommodate multiple load tables to cater for different record types and sources of data. The library needs to make sure that its load tables are adapted to deal with the new fields and sub-fields, so that they will load, rather than being rejected at the point of data loading. In some systems, if the new fields are added to these load tables, the records will load without the fields or, potentially, records will be rejected. The library must adjust all the load tables to accept the new fields for both bibliographic and authority records as appropriate. Similarly, non-MARC systems must be able to map new RDA MARC fields to the internal data structure that is used.

- Validation

Records within the integrated library system need to pass validation before they are added to the live database. Validation typically involves logical checks against the record structure and, possibly, parts of the data such as leader and 00X field data and values. If the validation rules in operation within the system do not recognise the new RDA fields any number of things can happen. For example, the record may be rejected or the cataloguer will not be able to move beyond the edit stage in their workflow. Thus, to avoid validation errors, and allow records to be added in their entirety to the database, the validation rules must be updated to accommodate the new RDA fields within the MARC format. Non-MARC systems may also require that new fields are able to pass validation.

- Indexing

Once RDA data has been loaded into the library system, the library will be faced with decisions about whether to index the data. This decision will need to be taken at both field and sub-field level. For example, the library has to decide whether to index new subfields in 1XX headings. Indexing decisions will not just be a decision whether to index new data or not, but also which index to add the data to. This may involve revising phrase indexes to incorporate additional subfields, or keyword indexes to include new field and subfield data. These indexing decisions will be influenced by the type of library and collections. For example, some libraries may decide to index parts of the new 3XX block in the bibliographic format, perhaps due to having multi-media collections, whilst others do not.

- Display

With the addition of RDA fields and subfields into the library catalogue, it is also necessary for the library to decide whether, or not, to display this data. Libraries will have control over the display of bibliographic data, either in results sets or as full records. Through the use of web templates and/or stylesheets it is possible to control the display of parts of the bibliographic record. Libraries will also have to make decisions about which captions to use against new fields. For example, deciding to display the 264 MARC field and use display labels, or captions, based on the indicator values assigned to this field.

- Discovery

Discovery platforms are widely adopted within libraries. These sit above the library catalogue, and a catalogue search is only one feature. The supplier of the discovery platform may be a third-party different to the supplier of the integrated library system. Even in those cases where the supplier of the discovery platform is the same as the library system, many issues may still need to be addressed. All these issues already addressed (loading, validation, indexing, display) will still be relevant as the data needs to be displayed both in the library catalogue and the discovery platform. This, effectively, doubles the work for the library in preparing for integrating RDA data into the local systems. This is particularly complicated if the library has to deal with two different system suppliers simultaneously.

- Exporting

Just as RDA data is loaded into the integrated library system, it is also exported. This may be in supplying data to a third-party union catalogue or database, or in loading bibliographic data to a discovery platform. As with loading data, the library system will have export tables that govern the export of data, specifying which fields will, or will not, be present within the exported record. If the library is exporting data to a third-party that will be in RDA, all the appropriate fields, subfields, and data elements need to be defined in the export tables. The library may have local control over this or may need to work with their system supplier to ensure that this requirement is in place.

Who to communicate with

Communication is at the heart of implementing RDA within the library system. The library needs to be in a position to make decisions based on dialogue with a number of potential third-parties.

- Record supply agencies

If the library is acquiring bibliographic data from record supply agencies such as OCLC or RLUK there is a need to know if the data follows RDA and when the supplier will be adopting the standard. A similar case applies for the supply of authority data either if it is acquired from OCLC, for example, or through the services of an automated authority control vendor. If the vendor will be supplying RDA data from a certain date, the library needs to be operating on the same timetable.

- Shelf-ready vendors

Many libraries use the services of shelf-ready vendors. Whilst the shelf-ready vendors are indicating that they will be adopting RDA, there is not necessarily agreement across vendors as to when this will happen. The experience of libraries to date is that few vendors are able to provide either sample or live RDA data at this time. Therefore, the decision to adopt RDA is impacted by the vendors and their ability to provide the data. Libraries may use separate load tables for loading the bibliographic records from shelf-ready vendors. Amending these tables will be contingent upon RDA uptake by the vendors.

- System suppliers

System suppliers are an obvious point of information and advice for libraries. It is essential, in the adoption planning stage, that libraries communicate with their system supplier. There may be many things that the supplier can do to assist and guide the library. Suppliers will be able to provide specific information about changes that need to be made to individual library systems. They will be able to either advise the library what it needs to do or be able to do the work on their system on their behalf. They may also offer informal networks between libraries, for example, enabling libraries with a system in common to share index rules, and so on.

- Discovery platform supplier

As has been noted, libraries increasingly use discovery platforms which may not be supplied by the same company as the integrated library system. For this reason the questions that are addressed to the system supplier will also need to be addressed to the discovery platform supplier. Even when the supplier of both systems is the same, the questions that the library asks may need to be posed separately for each platform.

- Partners

Where libraries belong to consortial databases or union catalogues, they will need to advise their partner institutions that they are migrating to RDA and when this will happen. Physical union catalogues will have all the issues of the integrated library system with respect to loading, validation, indexing, display, and exporting, and will need to be able to accept RDA data from member libraries. Without effective planning this is an area of potential strife for both the consortium and the member libraries.

- Colleagues

Within the library the move to RDA will result in bibliographic records that look ‘different’ to AACR2-compliant records in the local catalogue and discovery interface. This needs to be communicated with colleagues who work with catalogue data, such as systems staff, reference librarians, and support or enquiries staff. If colleagues within the library are not informed of changes, then there will be the potential for confusion. For example, the loss of the GMD for non-book materials may cause confusion, particularly in a catalogue with mixed AACR2 and RDA bibliographic records.

- Library management

Finally, it is essential that the adoption of RDA is communicated to library management. It is likely that a case for the adoption of this new standard will have to be made within the library. Explaining why and how the transition needs to take place should ensure the appropriate planning, and hopefully smooth running, of this change. Library management must not be kept in the dark; they need to be aware of the issues and the timetable that the library is working to in the transition to RDA.

Taking these factors into account should enable the transition to RDA to run smoothly. This only leaves the library with the challenge of adopting the new standard!

Finding Value in History: Gaining Knowledge by Examining Historical Practices

Katharine Whaite AHRC PhD candidate, Department of Information Studies, University College London

Since the idea of the modern catalogue coalesced with the work of Panizzi, Cutter and their prominent contemporaries, much has changed. Catalogues have seen their formats change from manuscript to slip to printed card to online public access catalogue, while their contents have expanded to include technologies that were not imagined when the catalogue was first constructed. The story of any library catalogue can be complicated, and these complications can lead to a rich source of investigation.

The catalogue of the library collections of the Natural History Museum (NHM) has evolved over the course of more than a century, from its origins as part of the British Museum, through its status as the British Museum (Natural History) [BM(NH)] on its move to South Kensington, and finally to its current state as the NHM. This catalogue is but one example of the way in which the catalogue, through its history of serving Cutter's Objects, can also be a history of the collection, of the way librarians perceive the collection, and the way librarians try to provide the widest and best access to the library's materials.

Far from being one choice among a variety of tools, the OPAC is now considered the primary source of information about library collections. While one of the strengths of the traditional catalogue is its capacity to help a user know when a search is complete (Oddy 1996 p. 35), for the user the OPAC has no physical structure, no obvious ending. It does not exist in a physical space, and so it can be difficult for a user to know when any search is complete. Initial interactions with collections are primarily search-driven, with the user entering the item or topic desired and waiting for search returns. But the way that OPACs function represents only part of our interactions with the collection – as though a collection is only a great mass of material, waiting for relevant items to be extracted.

By contrast, book and card catalogues physically set out the extent of the collection. The contrast is born out of the difference between the types of surrogates used for the collection. A book or card catalogue has a physical entry for each item, and using physical surrogates is much more like physical movement through the collection, more akin to browsing than how modern users experience search. An OPAC's surrogates are digital, stored on a remote server, and the retrieval of only certain records essentially blots out the existence of the rest of the collection.

This form of interaction with the collection is quite new. The first OPAC launched in the later decades of the last century (Bowman 2007 p. 317), but Dorothy Norris' survey, *A History of Cataloguing and Cataloguing Methods, 1100-1850*, begins in the 12th century. The centuries in between have moulded our interactions with books, seen the advent of accessible libraries, and expanded ideas of what a catalogue is and can do. This dense tapestry of actual practice is what makes the examination of historical catalogues such a rich vein of information.

While this piece will focus on records in the printed catalogues, my original research encompassed tracing a clutch of items through the NHM's historical records: using accession, purchase and donation books; and slip, card, printed, and OPAC forms of the catalogue. Over the lifetime of these books in this library, the cataloguing rules have changed eight times: through from Panizzi's "Rules for the Compilation of the Catalogue" (or the 91 Rules) first published in 1841 – used to catalogue books taken from the British Museum to establish a library for the natural history collections when they were moved to South Kensington – to AACR2, the standard they use today. The records for the items changed, and the items themselves changed. While it is the natural state of the catalogue to be constantly changing, more can be read in these changes than might at first be expected. Traces of professional practice, of individual judgement, represent not just these particular items, but can be looked at in light of being a small part of how the collections have been described, which inevitably alters how the collections are used.

In order to fully appreciate the nuances of the catalogue, it is necessary to describe, in some detail, the background of the items examined. Thomas Bell's *A Monograph of the Testudinata* was intended to be "an illustrated guide to all chelonia, both living and fossil species" (Sowerby, 1872 p. ii). Bell, a Fellow of the Royal

Society (Cleevely 2008), commissioned James de Carle Sowerby – from a very well known family of natural history artists – to draw the illustrations, and Edward Lear – the nonsense poet who started his career as a natural history illustrator – to lithograph them. The series was discontinued in 1842 before publication was completed. The plates, both published and unpublished, were later acquired by another publisher, and a new text written by J.E. Gray – another Fellow of the Royal Society and Keeper of the Department of Zoology at the BM(NH) (Cleevely 2008) – to accompany them.

This book was published in 1872 as *Tortoises, Terrapins and Turtles, drawn from life*. A proof copy of this second text, written by Gray, was bound with a copy of Bell's *Monograph*… held in the NHM library.

Evaluating the records for these items, different but so intimately connected, provided a window onto an interesting story, about the library, about the institution, and about the cataloguers who worked there. What is extremely useful in terms of equipping ourselves for changes that may come is to examine a few examples of evidence of individual cataloguer's choices found in these catalogues. These two records for Bell's *Monograph*… are taken from the *Catalogue of Books, Manuscripts, Maps & Drawings in the British Museum (Natural History)*.

The first was published in 1903 in the first volume – covering entries from A to D – and has pencil annotations, with the whole entry crossed out and 'cancel' written in the margin.

Bell (T.) *F.R.S.* A monograph of the Testudinata.

Pt. I-VIII.† *pp. xxiv* [*80*] *: 40 pls. col.*

4˚. *London* [1836-42].

Title from wrapper.

--------- [Another copy.]

Wanting the plates. There are bound up with this copy 11 additional pages (proof) of text, proof copies of four of the plates, and a proof of the text written by J.E. Gray to accompany the sets of plates which were issued in 1872 under the title "Tortoises, Terrapins and Turtles."

See SOWERBY (J. DE C.) & LEAR (E.)

--------- [Another copy.] *pp. xxiv: 46 foll text. And 61 proof pls. before letters with some copies as published.*

In this copy the text and the plates published and unpublished are systematically arranged. Wanting the wrappers, and text of *Emys Hamiltoni* and *Terrapene miniata.*

The second entry was published in the 1922 supplement – covering entries from A to I.

Bell (THOMAS) *F.R.S.* A Monograph of the Testudinata.

Pt. I-VIII.† *pp. xxiv* [*96*] *: 63 pls. (col.)*

4˚. *London*, [1836-42.]

The wrapper to Part VI has been used in lieu of the title-page ; all other wrappers are wanting.

This copy includes fourteen pages of text and twenty plates which do not appear in the list of contents issued with the parts, and may therefore never have been published. The majority of coloured plates in this copy were those that served as patterns. A "proof" of the text written by J. E. Gray to accompany the set of pates which were afterwards issued in 1872, under the title of : —"Tortoises, Terrapins & Turtles," by J. de C. Sowerby and E. Lear, has been inserted in this copy.

Wanting all the wrappers, and the plates of *Emys ornata* (adult) and *Chelonia mydas.*

In this copy the text and plates, published and unpublished, are arranged in systematic order.

--------- [Another copy.]

Wanting title-page : the wrapper to Part V has been used in lieu of title-page, all other wrappers are wanting, as also the following of the unpublished plates: -- *Testudo grœca, Emys elegans, E. ornata* (2 plates), *Hydraspis galeata* (2 plates), *Chelonia mydas,* and *C. imbricata.*

Notably, the first, cancelled, entry contains a cross-reference to the heading of Sowerby and Lear, but the second does not. Here, what we can interpret is that the cataloguer is choosing to include the information that Sowerby and Lear are related to Bell's *Monograph…* as notes, instead of using a cross-reference. As this option was available under the contemporary rules, it begs the question – why? We can see the cataloguers exercising their own judgment, and leaving traces of their decisions. The later cataloguer must have considered it to be more representative of the item, and more useful, to provide additional information to those consulting this entry, rather than to direct users to consult another entry in the catalogue.

The same holds true for the card catalogue entries for Bell's *Monograph*... The cataloguer chooses not to insert cross-references to or from the artists, although that option is available under contemporary rules. Another reason might consider the users of the collection: scientists working in the museum. A scientific community at that time might be more likely to be interested in what the text and illustrations as a whole can tell them about the object under study, and not in the illustrator as an important creator responsible for the work. This may not be a perfect interpretation, but taking account of knowledge about the library as an institution, the rules its library staff operated under at the time, and its user group, can help provide a reasonable educated guess as to why such decisions were implemented.

Another interesting factor to consider is the title pages of the two books. Below are quasi-facsimiles of the title pages (or what serve as title pages) in the bound volumes examined.

A Monograph of the Testudinata

A | MONOGRAPH | OF THE | TESTUDINATA. | BY | THOMAS BELL, F.R.S. | FELLOW OF THE LINNEAN, GEOLOGICAL, AND ZOOLOGICAL SOCIETIES ; CORRESPONDING MEMBER OF THE SOCIETY OF NATURAL | HISTORY OF PARIS; OF THE ACADEMY OF NATURAL SCIENCES OF PHILADELPHIA; OF THE NATURAL HISTORY SOCIETY | OF MANCHESTER; AND LECTURER ON COMPARATIVE ANATOMY AT GUY'S HOSPITAL | [double rule] | LONDON: | PRINTED FOR SAMUEL HIGHLEY, 32 FLEET STREET | BAR | *Price* 1*l.* 1*s*., *Coloured.* | RICHARD TAYLOR, PRINTER,] [RED LION COURT, FLEET STREET.

Tortoises, Terrapins, and Turtles

TORTOISES, | TERRAPINS, AND TURTLES | DRAWN FROM LIFE | BY | JAMES DE CARLE SOWERBY, F.L.S. | AND | EDWARD LEAR. | [double rule] | LONDON, PARIS, AND FRANKFORT: | HENRY SOTHERAN, JOSEPH BAER & CO. | 1872.

It is clear that despite having very similar contents, the main responsibility for the book is allocated quite differently. Accordingly, the presentation of the title page is different, and that difference influences the cataloguer's decision about what information about the book is important to record, especially as an access point.

There are many issues that influence our interpretations. Bell, for instance, was an important and influential figure at the time of the printing of his *Monograph…*, and was also responsible for its addition to the BM(NH) library, so we expect his contributions to be widely acknowledged. By the time of the reprint it was the pictures in this work that were deemed more important than the accompanying text, so instead of Gray's name replacing Bell's as the responsible creator, credit is focused on the artists, Sowerby and Lear.

These two texts are quite intimately related, and although declaring the nature of their relationship in the context of Functional Requirements for Bibliographic Representation (FRBR) is a tricky business, they seem to be two expressions of the same work. This relationship would be very difficult to glean from records that were based solely on the books' title pages. But cataloguer discretion allows for some correction of the lack of specific information that occurs on title pages, which can be seen in the examples that follow.

Consider this record in an earlier BM(NH) catalogue, that of the Zoology library, published in 1884.

SOWERBY (J. de C.) & LEAR (E.). Tortoises, Terrapins, and Turtles. The plates drawn under the superintendence of THOMAS BELL. Preface and notes by J. E. GRAY. 4˚. *London*, 1872.

Perhaps most interesting about this entry is that it goes against what is expressed as the purpose of the Zoology library catalogues in the preface of the first edition, published in 1880, as a "…List [that] has been printed chiefly with a view of serving as a temporary guide in making additions to the Zoological Library, and of preventing the acquisition of duplicates… no bibliographical completeness of the titles has been attempted; on the contrary, the titles have been as much abridged as seemed to be consistent with the recognition of the works catalogued." (BM(NH) 1880, p. i)

This makes the notes visible and rather lonely ones in a mostly note-less catalogue. Again the contents of the entry force the question – why include this information? It might not make much difference to anyone trying to find this specific book. But the cataloguer goes against the stated purpose of the catalogue to provide information that will be useful to the user group. Here, a descriptive note is used to explain the complex story of this book, relevant to those studying in this field, that preserves at least some of Bell's intellectual responsibility which might otherwise be bibliographically lost. This is a classic example of a cataloguer bending the rules for the benefit of the user, and for posterity.

Now we turn again to the *Catalogue of Books, Manuscripts, Maps & Drawings in the British Museum (Natural History)*. In volumes 2 and 3, there are two different kinds of references from Lear, the lithographer:

Lear (E.) & **Sowerby** (J. DE CARLE) Tortoises, Terrapins and Turtles drawn from Life. *See* SOWERBY (J. de C.) & LEAR (E.) fol. 1872

and Gray, the writer of the text:

GRAY (JOHN E.) Tortoises, Terrapins and Turtles drawn from life. [With introduction and explanatory text by J. E. Gray.] *See* SOWERBY (J. de C.) & LEAR (E.) fol. 1872.

with the main entry under Sowerby & Lear in volume 5:

Sowerby (James DE C.) & **Lear** (E.) Tortoises, Terrapins and Turtles, drawn from life. [With introduction and explanatory text by J. E. Gray.] *pp. xxiv, 16 : 60 pls. col.* fol. *London, &c.,* 1872.

This work consists of the plates originally drawn for T. Bell's "Monograph of the Testudinata" [*q.v.*]. Only forty of the plates had been published when that work was stopped. A revised proof of Gray's text is bound up with the copy of Bell's monograph.

The main entry also includes a link to Bell, and again draws an important connection not only for scientists, but also for the history of the institution.

Taken as a whole, these records show awareness of the needs of the user, and, resulting from that, a desire on the part of the cataloguer to tell the complete story of the work. "Doing so through the catalogue, linking these two items together in reference, both moving forward in time from *Monograph*… to *Tortoises*…, and tracing back through time from *Tortoises*… to *Monograph*…, should prevent future loss of the intellectual links between these two volumes" (Whaite 2010, p. 45).

This particular investigation raises a lot of questions. Having witnessed the importance of the change in catalogue entries over time, can the idea of cataloguing something once – rightly, and for all time – persist? Is it even desirable?

Can librarians rediscover the practices that aid them in being as prepared for maintaining catalogue records as they are for creating and expanding them? Can we encourage respect for the catalogue, which must include respect for its limitations? As we accept that the library is a growing organism, we must also accept that the catalogue is an evolving one, whose body must be added to but also adapted. How can we tackle the problem of the digital entry, which can be modified with a simple batch command and leave no trace of its former constituents?

What becomes very clear is that the catalogue does not just run alongside the history of the library, or the institution the library is part of, but is part of it and helps to shape it. Understanding the practical history of one's own library, and how its librarians have crafted their practice to serve its purpose, makes better catalogues and better cataloguers.

Investigating a catalogue as one would a text can be extremely rewarding, and provide information about users, librarians, collections, libraries and institutions, as well as shed light on how those entities interact with each other. While considering all of these factors does not lend itself to easy interpretation, it is absolutely worth the effort. In order for these investigations to remain interesting, and to produce anything like insight, it is vital that cataloguing librarians preserve the records of their own work – both at the scale of policies and practical documents, and full and thorough catalogue entries – as the raw documents that will help us to write our history. Records of why cataloguers do what they do must be preserved to prevent wasting the body of knowledge they have built. The future of cataloguing history relies on the generosity of cataloguing professionals, with their knowledge of and work on their own catalogues.

Works cited:

Bell, Thomas. 1836-1842. *A Monograph of the Testudinata*, Pts. I-VIII. London.

British Museum (Natural History). Department of Zoology. 1880. *Catalogue of the Books in the Department of Zoology*. London.

---. 1903-1940. *Catalogue of the Books, Manuscripts, Maps & Drawings in the British Museum (Natural History)*. 8 vols. British Museum (Natural History): London.

Cleevely, R. J. 2008. 'Bell, Thomas (1792–1880)', Oxford Dictionary of National Biography [online]. Oxford: Oxford University Press. Available from: http://www.oxforddnb.com/view/article/2029 [Accessed 27 Nov 2012].

Cleevely, R. J. 2008. 'Gray, John Edward (1800–1875)', Oxford Dictionary of National Biography [online]. Oxford: Oxford University Press. Available from: http://www.oxforddnb.com/view/article/11344 [Accessed 27 Nov 2012].

Cutter, Charles Ammi. 1904. *Rules for a Dictionary Catalog*, 4th ed. rewritten Washington, D.C.: GPO.

Norris, D.M. 1939. *A History of Cataloguing and Cataloguing Methods 1100-1850: With an Introductory Survey of Ancient Times*. London: Grafton.

Oddy, P. 1996. *Future libraries, future catalogues*. London: Library Association.

Sowerby, James de Carle and Edward Lear. 1872. *Tortoises, Terrapins, and Turtles: drawn from life.* London: Henry Sotheran, Joseph Baer.

Whaite, K. C. S. 2010. *Changing Records, Changing Expectations: The Evolution of the Catalogue at the Natural History Museum*. MA thesis, University College London.

Collaborative authorities: introducing the UK NACO funnel project

Deborah Lee Senior cataloguer, Courtauld Institute of Art

Introduction

Collaboration is a key part of the development of twenty-first century cataloguing practices. The LC/NACO authority file – under its various names and guises – is a proud example of how cataloguers have been collaborating for many years. However, UK participation in one particular international collaborative project, the Name Authority Cooperative Programme (NACO), is currently mostly limited to UK legal deposit libraries (Library of Congress, 2012a). (1) This paper describes the work being done by the CILIP Cataloguing and Indexing Group (CIG) to design a sustainable UK NACO funnel, which will enable all interested parties in the UK to participate in NACO in the future.

The first part of this article briefly outlines the concept of a NACO funnel. This is followed by a short discussion about the benefits of participating in a funnel. The main part of this paper summarises the structure and format of the proposed UK NACO funnel, which is followed by an update on the progress of the project plans. This article will not give any theoretical details about NACO, as it deals exclusively with the proposed UK NACO funnel project. In addition, limitations of space and confidentiality limit the level of detail and specificity that will be covered; this is particularly pertinent in the omission of the calculations about projected time commitments and financial details. These details and an in-depth description of the project are provided in the draft UK NACO funnel project proposal (Lee, 2012). As the project is only at the planning stages, this account does not include a formal project plan or timeline. However, even with all these provisos, it is hoped that the benefits of being part of NACO and the general ethos of the project are illuminated in this article.

What is a NACO funnel?

NACO is the name authorities branch of the Program for Collaborative Cataloguing (PCC), who define themselves as "…an international cooperative effort aimed at expanding access to library collections" (Library of Congress, 2012b). In the NACO programme, members add new names and edit existing names in the LC/NACO authority file. Traditional authority standards are used such as AACR2 (or RDA) and LCRIs, as well as the NACO guidelines. The training programme – which traditionally lasts for five days – is followed by a review period of three to twelve months. (2)

A NACO funnel is a way in which libraries work together to contribute records collectively. In the words of the PCC:

"A NACO funnel project is a group of libraries that have joined together to contribute name authority records to the national authority file" – (Library of Congress, 2012c)

Funnels enable libraries to share costs; furthermore, they allow libraries to participate who would not normally have enough new names in any given year to become NACO members. Funnels are usually formed around a particular geographic location or subject – for instance, "NACO Canada" or "NACO-Music" (Library of Congress, 2012d). For more information about NACO and NACO funnels, see for instance the PCC NACO funnel webpages (Library of Congress, 2012c) and Hugh Taylor's presentation for a CIG workshop (Taylor, 2009). The UK NACO funnel project aims to establish a geographic funnel for the UK.

Benefits of NACO funnel membership

There are many potential benefits of being a member of a NACO funnel, and a few are highlighted below.

"Collaboration"

Though membership of NACO is free there are other costs involved, for instance the expenses of the trainer; being part of a NACO funnel enables these costs to be shared amongst institutions. There is also the potential for sharing expertise between funnel members at different institutions, which is particularly attractive for libraries where there might be only one professional cataloguer. (3)

"Access to free, high-quality training"

Training is a significant part of NACO. Funnel membership involves attending the NACO training sessions; wherever in the world the training takes place, the opportunity to participate in these highly-regarded courses consisting of centralised training materials is invaluable (Library of Congress, 2010).

"Professional development"

Participating in the NACO funnel will provide cataloguers a chance to maintain or increase expertise and fluency in authority work. Moreover, the UK NACO funnel model is based upon peer training, so that participants will also get the opportunity to expand their training skills and experiences.

"Representation of authors and organisations important to home institutions"

NACO participation means that the people and organisations which are important to the library or the library's home institution can be represented in the LC/NACO authority file if they are not currently listed. (4) For instance, cataloguers could ensure that all names in faculty-authored books or staff- curated exhibition catalogues are included in the LC/NACO authority file if they so wished. Furthermore, not only would cataloguers have the opportunity to add locally-significant names, but will also have the chance to influence the form of that name; they could utilise their local knowledge and personal connections to ensure that the preferred form of the name is used or to obtain knowledge useful for disambiguation – obviously, as long as no NACO procedures or rules are violated in the process.

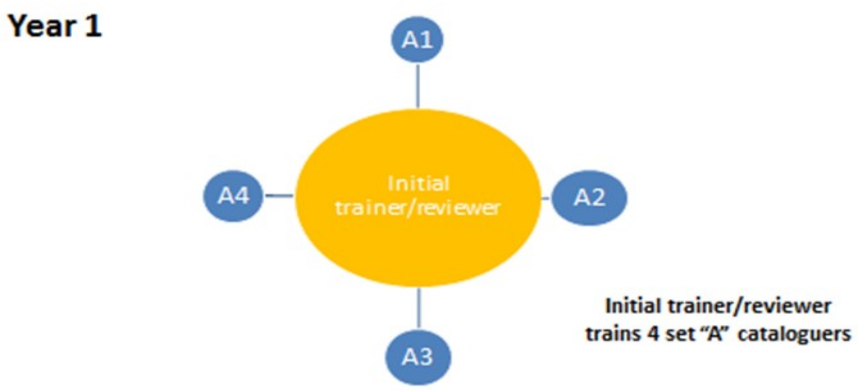
Structure of the proposed UK NACO funnel

Traditionally, a NACO funnel functions via a single coordinator. The role of the coordinator includes registering new members, distributing documentation and generally being the point of contact for the funnel. Often the coordinator is also responsible for the training and reviewing of all funnel members. (5) At present, there is not a UK-based NACO trainer, which presents a number of problems and potential extra expenses. For instance, as well as the potential costs of obtaining an external NACO trainer, there are issues concerning sustainability and hospitality to potential new members in subsequent years. On the other hand, the burden placed on a UK cataloguer who becomes the funnel's official NACO trainer would be heavy, and the pool of NACO-trained cataloguers in the UK – the ideal background for someone taking on this role – is relatively small. In short, the traditional NACO funnel model has some significant issues if applied to the UK cataloguing community. There is, however, a possible solution: cascade the training and reviewing through the majority of participants in the funnel.

The structure of the UK NACO funnel was designed using the information, advice and shared experiences of funnel coordinators and senior cataloguing staff at UK legal deposit libraries. Emails were sent to geographic funnel leaders and specific cataloguing staff at UK legal deposit libraries, with requests for information on specific topics. (6) The response rates were 48% from the geographic funnels and 67% from the UK legal deposit libraries. (7) The responses revealed a number of very useful ideas and pieces of advice. First, the training and reviewing part of the coordinator role is very time-consuming, yet the administrative burden is light. Second, as well as some funnels having a semi-distributed structure, a fully-cascaded funnel also exists, the NACO-Music funnel. Third, the time taken for someone to become independent – in other words, when a cataloguer is no longer being reviewed and is independently submitting authority records – is critical to the success of the funnel. Contacting the music funnel coordinator was the next step taken, which included asking about the funnel's structure and the logistics of maintaining a devolved funnel. Using the responses from the geographic funnel coordinators, the UK legal deposit libraries and the music funnel coordinator, the UK NACO funnel structure was designed.

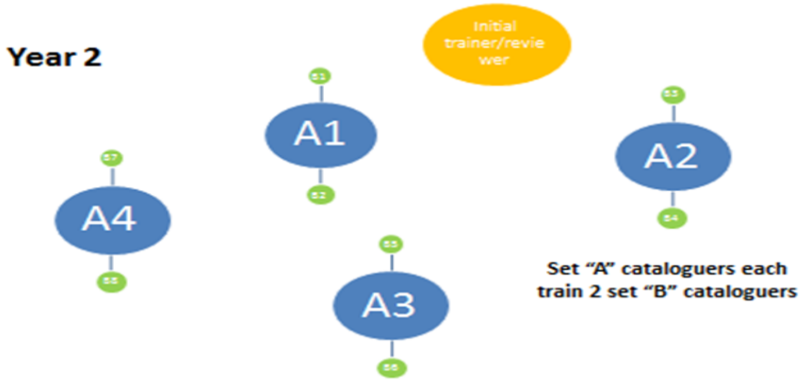
In the first year, the initial trainer/reviewer trains four people (set "A"), from four different libraries (see fig. 1). The initial trainer/reviewer is likely to already have substantial NACO experience, and the project has been designed to seek funding for this post for the first year as well as to provide funding for the initial training course.

Fig. 1: Funnel in year 1



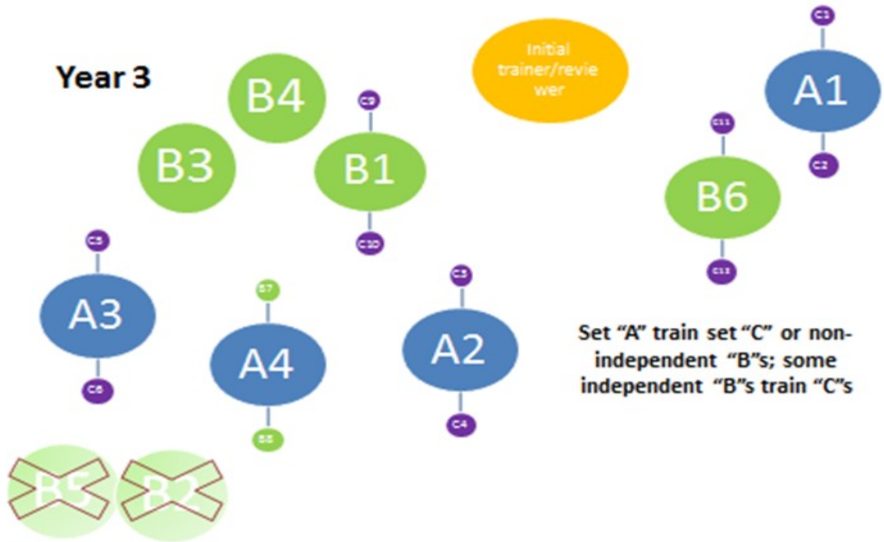
In the second year (see fig. 2), each of the cataloguers trained in the first year (set "A") trains and reviews another two cataloguers (set "B"). (8)

Fig. 2: Funnel in year 2



There are various possible expansion models of the funnel, but after looking at the prospective growth of the funnel compared with the respective workloads for individual participants, "two" was selected as the most ideal number of trainees-per-trainer. (9) From various responses and figures given by funnel leaders and UK legal deposit libraries, combined with Marais' research (2004) which itself draws upon earlier statistics (for example, Byrd and Sorury, 1993), it is possible to guestimate a worst-case scenario figure for the time commitment needed to participate, train and review. Though space does not permit a full scale analysis of the various figures and calculations, the time commitment of the set "A" participants in year two is still in the order of a small number of hours per week (Lee, 2012). In the second year, the initial trainer/reviewer becomes a senior cataloguer figure, helping and advising the new trainers and answering complex enquiries. (10) It is assumed that from the second year onwards there is no or little funding.

Fig. 3: Funnel in year 3



While the third year (see fig. 3) carries on in the same fashion, the model becomes more complicated as once we move into the regular funnel – as opposed to the initial years – dropout rates have to be considered. Three different percentages need to be taken into account: those participants who drop out completely, those who need more time to become independent but remain in the programme and those who are independent but cannot train others. (11) In the model, we assume that the set "A" participants are "perfect", at least for the first three years; in other words, they all stay members of the UK NACO funnel, they become independent within their first year and they are happy to train new members in their second year. It is impossible to guarantee any of these factors with any particular participant; however, as the modelling shows that the funnel is more likely to be sustainable and to grow if the set "A" participants fulfil these criteria, it is critical to select set "A" participants on the basis of those who are most likely to be "perfect" for the first three years.

The third year is represented in figure 3 and dropout rates have been assigned: a quarter of set "B" are independent and have become trainers of set "C", a quarter of set "B" are independent but not able to train others, a quarter of set "B" still require reviewing and a quarter of set "B" have completely dropped out. (12) These figures also show how quickly the funnel can grow, even with these modest trainee-per-trainer quantities

and high dropout rates. As there is not an infinite pool of potential NACO participants in the UK, there is no need to have infinite capacity for new members each year. The modelling suggests that the emphasis should be on creating a sustainable funnel rather than rapid expansion.

Therefore, the email responses and modelling demonstrate two key points for a successful UK NACO funnel, especially over the first few years. It is critical to get as many participants independent as soon as possible; non-independent cataloguers impede new members from being able to join the funnel. Also, the funnel needs an initial cohort who are, in theory, happy to train others after their first year. The key to success in the first few years is always having enough trainers; without this, stagnation occurs.

Future of the funnel, progress and concluding points

The potential impact of RDA has featured in discussions about the UK NACO funnel. By the time the project commences, the initial trainer/reviewer will be RDA authority-trained and experienced in RDA record contribution, as all NACO members will be NACO trained by the end of March 2013 – see the RDA NACO training timetable (Library of Congress, 2012e) – and the funnel is not expected to commence until at least late 2013 or 2014. However, while RDA implementation is occurring, it is unlikely that potential participants would have the opportunity to focus on NACO. RDA has both potential advantages and disadvantages for the UK NACO funnel, and the impact of RDA on this proto-funnel will only be known in time.

CILIP CIG started discussing the possibility of establishing a UK NACO funnel after the 2010 CIG conference, though substantial work only commenced from mid-2011 onwards. Tasks completed so far include the scoping of international funnel leaders and the existing UK NACO participants as described above, as well as discussions within CILIP CIG. Creating the structure of the funnel and the modelling has been a significant endeavour; this was due to many factors including the lack of existing data on dropout statistics and the

The next task is to determine interest amongst the UK cataloguing community. As well as conversations and emails resulting from the presentation at the CIG 2012 conference, a survey will be circulated to the UK cataloguing community within the next few months. Simultaneously, formal contact with Library of Congress will be made, to discuss the potential funnel and in particular, the structure of the proposed cascaded model. Once all this information has been collected, the formal process of seeking funding and national support will start. The information from both these tasks will be combined with the existing models, scoping and business case, which will be used to approach specific organisations for support and/or funding. Though only at the beginning of the path towards a UK NACO funnel, the CIG paper and this resulting article mark the start of what hopefully will be an exciting new chapter in UK collaborative cataloguing.

References

1. There are three exceptions to this: Wellcome Library, National Art Library and University of Strathclyde. However, private correspondence with colleagues in the first two of these suggests that these two programmes are currently hibernating or moribund, so Strathclyde is currently the single, non-legal-deposit library member in the UK (Library of Congress, 2012a).
2. Some respondents to questions about NACO funnels gave accounts of non-traditional training periods; these were shorter than the standard five days and/or did not follow the intensive structure.
3. The term "cataloguer" is used here and elsewhere in the article as an umbrella term for positions such as "metadata librarian" and "assistant librarian (cataloguing)".
4. To some funnel coordinators, one of the strongest motivations for establishing a geographic funnel was the opportunity to represent names and organisations specific to the location of the funnel. This view is echoed by Larmore (2006 p. 76), who suggests that a significant driver for establishing a South Dakota funnel was the ability to represent South Dakota people and organisations.
5. Not all funnels work in this way, as usually the trainer/reviewer is an official NACO trainer.
6. The majority of geographic funnel leaders were emailed; the only omissions were for funnels based in countries or regions which did not have English as one of their official languages, which was due to language-based, pragmatic reasons. So the majority of the population was initially used for the research, with a small amount of "pragmatic sampling". The contact names and emails were retrieved from the NACO site, as listed in May 2011 (Library of Congress, 2011). In some cases, the coordinator information was out of date so the emails were forwarded to the correct person. Problems with the Courtauld email system at the time of sending the NACO-related emails is acknowledged as a methodological issue with the collection of this data. Though emails were re-sent once the problem was revealed, it is possible that some emails or responses were not received. Only three out of five of the UK legal deposit libraries were contacted, as the other two had already offered advice and information by the time of the information-gathering emails.
7. Both percentages have been rounded up to the nearest whole number.
8. Set "B" cataloguers will review the new cohort, but details about how this training is carried out are under discussion. The advantages of a single, centralised course and its corresponding expenses, are juxtapositioned with the advantages of a de-centralised system, where ideally each trainer would carry out their training locally and to a convenient timetable for the trainee.
9. Any fewer than two, the modelling shows that the funnel could come to a standstill; conversely, more than two creates a major time commitment for the participants. For space reasons, the modelling which demonstrated what happened when different values for the number of trainers per trainees were used has not been reproduced here.
10. It would also be possible for the initial trainer/reviewer to train and/or review new cataloguers. However, this has not been included in the model.
11. In all cases, the question of how long each person will stay in any one of those dropout categories is an added complication. For instance, person "X" might not be happy to train new members a year after they start, but could be happy to do this three years later.

12. It was not possible to find any existing data on dropout rates, therefore they have been estimated. However, the responses from geographic funnel leaders gave a good account of some of the main reasons for funnel members not becoming independent or dropping out altogether; as well as lack of institutional support, the other main reason cited was not creating enough records per month, either through lack of time or unavailability of new names at the cataloguer's institution. It seems there is a critical mass of new authorities needed for successful NACO participation, even as part of a funnel.

Bibliography

Byrd, J. and Sorury, K., 1993. Cost analysis of NACO participation at Indiana University. *Cataloging & Classification Quarterly*, 16(2), pp. 107-121.

Larmore, D.P., 2006. A new kid on the block: the start of a NACO funnel project and what is needed to start your own. *Cataloging & Classification Quarterly*, 42(2), pp. 75-81.

Lee, D., 2012. *NACO funnel proposal: final draft*. [Unpublished and confidential document, produced on behalf of the CILIP CIG committee]

Library of Congress, 2010. NACO training program workshop. [online] Available at: <http://www.loc.gov/catworkshop/courses/naco/index.html> [Accessed: 29/11/2012]

Library of Congress, 2011. *PCC NACO projects*. [online] Available at: <http://www.loc.gov/catdir/pcc/naco/funnels.html> [Accessed: 24/05/2011]

Library of Congress, 2012a. *List of PCC NACO liaisons by institution or project*. [online] Available at: <http://www.loc.gov/aba/pcc/naco/liaisonsac.html> [Accessed: 2/12/2012].

*This link gives access to A-C, so from here use the hyperlinks to get to listings for the remainder of the alphabetical list. At the time of writing there did not appear to be a publicly available URL which directs the user to a page for the complete list.

Library of Congress, 2012b. *Program for Cooperative Cataloging*. [online] Available at: <http://www.loc.gov/aba/pcc/> [Accessed: 29/11/2012]

Library of Congress, 2012c. *PCC NACO funnel projects.* [online] Available at: <http://www.loc.gov/aba/pcc/naco/nacofunnel.html> [Accessed: 29/11/2012]

Library of Congress, 2012d. *[List of NACO funnels and coordinators]*. [online] Available at: <http://www.loc.gov/aba/pcc/naco/funnels.html> [Accessed: 29/11/2012]

Library of Congress, 2012e. *PCC/NACO Institutions and NACO Funnels Scheduled for RDA in NACO Training Modules* [online] Available at: <http://www.loc.gov/catworkshop/courses/rda_naco/Scheduling/NACO%20Institutions%20and%20Funnels_Schedule.pdf> [Accessed: 29/11/2012]

Marais, H., 2004. *Authority control in an academic library consortium: using a union catalogue maintained by a central office for authority control*. PhD. University of South Africa. Available at: <http://uir.unisa.ac.za/xmlui/bitstream/handle/10500/2546/thesis.pdf?sequence=1> [Accessed: 2/12/2012]

Taylor, H., 2009. *Cooperative name authority data - the LC/NACO Authority File. Paper given at CIG workshop: A. Rose, by any other name: a workshop on authority control. London, England, 23rd October 2009. [online] Available at:* <http://www.ukoln.ac.uk/events/CIG/2009/authorities/presentations/ppt-2000-html/h-taylor_files/v3_document.pdf> *[Accessed: 29/11/2012]*

The effective use of e-mail lists and other projects to improve bibliographic data quality

Ian Fairclough **George Mason University**

A user performs a keyword search. It fails. Why? Because the word was misspelled in the database. Somebody browsing the shelves of a library by Dewey class number doesn't find what they want because the work is misshelved, not because the shelver didn't pay attention, but because of a wrong number in the record. Someone looks for somebody by a personal name heading and they find not only works by the person they were looking for but also others with whom that person was confused because they share the same heading, or because a heading has actually been wrongly assigned. Or in another case, a new heading has been since created to which that person should have been moved. A scholar searching for works published in a bibliographic series finds some of the works in that series, concludes he's found them all, but is wrong: there are others he didn't know about because no access to those other works is provided by a series heading.

Many of the projects described herein were developed particularly to address some of these issues. All of them use e-mail for their development, and they are all cooperative, collaborative in nature. But, unlike the NACO program, which is very strictly monitored by Library of Congress, mostly they are independent of any official supervision. There are no qualification requirements to participate in any of these projects, although you may wish to check with your supervisor before you do.

Several interrelated projects began with Jeffrey Beall's Dirty Database test, published in American Libraries in 1991, when keyword searching was comparatively new. Beall has written many excellent scholarly works, but when he gave me permission to reproduce this test, he asked me to make clear that it's not scientific in nature.

Jeffrey Beall's Dirty Database Test

1. Febuary _______
2. Guatamala _______
3. Misssion _______
4. Goverment _______
5. Fransisco _______
6. Conditons _______
7. Recieve _______
8. Wensday _______
9. Seperate _______
10. Grammer _______

You can do the test in your local database by going through these misspelled words, scoring one for each truly misspelled word. Ignore ones that have *sic* in brackets, or *i.e.* followed by a correction. Add up the total, and subtract from 100. The score can then be used to compare databases of comparable size. Beware of situations where other languages are involved, and personal names: this last word *Grammer* is actually a correctly spelled personal name.

Terry Ballard, who was then the systems librarian in Adelphi University, did Jeffrey Beall's Dirty Database test on Adelphi's system, which scored 98 out of 100, and so did very well. But Ballard's curiosity was aroused, and he proceeded to create an alphabetical list of all words in the database. He perused it to investigate which ones represented misspellings, and found numerous typographical errors. This database lists all the words that have been discovered by Ballard and his associates working collaboratively over the last twenty-plus years, and has grown from an initial 500 to about 8,000. This is the first half dozen:

- **Aamer***
- **Aand**
- **Aassis***
- **Aasso***
- **Aaway**

- **Abandonded**

It is now compiled by Tina Gunther from activity on the LibTypos-L discussion group. If you would like to join in, please go to http://groups.google.com/group/libtypos and sign up**.**

Beware of running this database against your own and making all the corrections! You can't really make changes without checking even if the misspelled word is embedded in a larger character stream, because some are genuinely misspelled, or personal names, or correct in another language.

Perhaps some readers are familiar with the Typo of the Day (http://librarytypos.blogspot.com). Each day people post to the blog which Terry Ballard set up. You may not be able to correct everything in your local database, but at least on one particular day you can make your database free of this one typo. And, you will be working in concert with librarians around the world. Maybe at this very moment someone's correcting the exact same typo that you're correcting!

Terry Ballard's 2008 article (1) concludes like this.

- In the first 24 days of October, 2006, the web page experienced more than 1,800 hits, averaging 77 per day.
- Visitors came from the United States and 39 other countries, representing every continent except Antarctica.
- The work that began 15 years ago has established itself as a permanent effort in the library world.

Actually, the work has been going for six years or so, but it will only continue as long as people continue to contribute – without that, it will cease.

There are also a number of Listserv® lists for error reporting, such as:

- DEWEYERROR
- LCCERROR
- PERSNAME-L
- SERIES-L

(Note that DEWEYERROR and LCCERROR do not have the suffix –L).

"Listserv" should not be used as a generic word, partly because it is a trademarked phrase, but chiefly because not all electronic distribution lists are Listservs. There are rival software programs, which don't necessarily behave in the same way. They might have archives, you might be able to "set nomail" when you go away, and so forth. Listserv does have these features, and therefore I have set up the various lists described here, in conjunction with several other people.

The first list I set up was DEWEYERROR, after I started work in Marion Public Library in Ohio in 2001. One of my responsibilities was to look over the entire collection of nonfiction, and reconcile certain issues with Dewey Decimal Classification. In the process of doing so, I checked to see what the strictly correct number should be, then checked to see whether that pure number should be modified in accordance with local procedures. Working with Library of Congress cataloging, which is supposed to be so good that you don't even have to check it, I found numerous cases where I wrote and told them about numbers that I suspected were in error. Once in a while they'd kindly give me a correction, but usually they'd say "Thank you for telling us." LC always are very grateful to hear about errors they have in their own catalog.

It occurred to me that other people likewise were writing to the Library of Congress. And some agencies, particularly large public libraries in the United States, would take LC records and use them as is, ending up with misshelved items for the reason of a wrong Dewey number. I identified two groups of people: those who check the number and report, and those who don't check, but probably would be glad to hear. "I can bring those two groups of people together," I thought. With an announcement on the Autocat discussion list, I got about thirty people together in an e-mail group that I administered from my Outlook mail account on my workstation. Other people sent me copies of messages that they had sent to LC, and I forwarded them to the group, which grew to about seventy. After a couple of years this process became unwieldy for technical reasons such as people's e-mail addresses failing every time I'd send something out. So I thought, "It would be great if it could be a Listserv list." Not only would those problems largely go away, but also we would have archives of everything that had been posted, so people could go back later on, if they wanted, and find out what messages had been sent out, if they came in later to the project. But we didn't have Listserv locally. Fortunately, Margaret Maurer at Kent State University told me that she was going to get the Ohio Library Council's Technical Services Division (OLC) to sponsor the project, which they kindly did. They came up with the name DEWEYERROR, and it started on October 27, 2004 . You can subscribe via LISTSERV@LISTSERV.KENT.EDU.

Incidentally, I'm a member of OCLC Global Council and have had a close relationship with several OCLC staff members over the years, doing a number of presentations, on error handling and quality control, with Brenda Block, who is now retired as the head of OCLC's Quality Control Section. Coming up in January is a webinar, in preparation jointly with Cynthia Whitacre.

This is a post to DEWEYERROR:

> **Friday January 6, 2012**
> **LCCN: 2010033097**
> **Author: Phillips, Carl**
> **Title: Double Shadow**
> **Dewey number in record is 813.54. This is a book of poems, so the Dewey number should be 811.54.**
> **This message has also been submitted to LC.**
> **Barbara Thiesen (Bethel College)**

It has the three components that are required. First, identification of the bibliographic item. Second, a statement of the problem (most Dewey classifiers can immediately tell what is wrong, because Dewey is mnemonic, but other people won't know unless they're told). Thirdly – and this is a list requirement - only LC cataloging is in scope for the DEWEYERROR list. LC don't want to know that the record is wrong in OCLC, only that a record is wrong in LC's catalog. In this message the reporter fulfilled that requirement perfectly.

I started LCCERROR (LISTSERV @LISTSERV.KENT.EDU) on November 17, 2010 after moving into my new position at George Mason University. Now I was working with LCC, so it became feasible for me to set up this list. I got people at Kent State to help again as listowners, to spread the work around. With several lists, if one Listserv site goes down, this avoids having the entire suite of tools simultaneously becoming inoperable, and also helps to plan for the succession of list owners.

LCCERROR is different from DEWEYERROR in some respects. LCC numbers have cutter portions, with subfield b in field 050. Dewey numbers don't necessarily have those, and LC doesn't shelve by Dewey. Sometimes the LCC cutter number can be in error. Only LC records are in scope for DEWEYERROR. That's not the case for LCCERROR because LCC classification numbers are assigned by other agencies, and perhaps those LCC numbers are accepted as is without question.

Here's an example post from Jay Shorten of the University of Oklahoma, who has been very helpful in setting up these lists.

> **LCCN 2011006448 <u>The lesson of Carl Schmitt</u>, Expanded ed., is not complete yet, but has an 050 00 of JC263.S34 M44514 2011. This number is not correct; it should be JC263.S34 M44513 2011 to match its previous edition, LCCN 98023580.**
> **(LC has changed their call number.)**

In this example, the difference between the number that was given and the number as it should be may be considered a minor detail. Does it matter? LC didn't respond, "Please don't bother us about little things like this!" It mattered to LC sufficiently that they went ahead and made the change. And in many quality control concerns, the answer to "Does it matter?" is, "It matters if you make it matter."

PERSNAME-L (LISTSERV @LISTS.OU.EDU**)** is dedicated to concerns about personal name headings in bibliographic records. Jay Shorten was alerting Autocat about situations where personal name headings had changed, and I asked him if he would be willing to set up the PERSNAME list. This he did, in July 2007, with my help. Here's an example of how this differentiation takes place.

Persons named Joseph Seymour

First Author, of:
First troop Philadelphia City Cavalry, c2008
Second Author, of:
Lenin & the vanguard party, 1997

Joseph Seymour is a historian with the U.S. military. When I tried to assign the heading for him, I came up with this name authority record which has two authors on the same record.

First Name Authority Record As Originally Written

010 n 2008044954
040 DLC ǂb eng ǂc DLC
1001 Seymour, Joseph
670 Seymour, Joseph. First troop Philadelphia City Cavalry, c2008: ǂb t.p. (Joseph Seymour) cover p. 4 (Philadelphia native; US Army historian, Washington, D.C.)

This is an undifferentiated name authority record (NAR). The two authors are the author of the books *Philadelphia City Cavalry*, and *Lenin and the Vanguard Party*. It might be one and the same person, for they're both about history. I checked to see what had happened in the history of this name authority record, using OCLC's Show LC Superseded Records functionality. The NAR was originally written for the author of *First Troop Philadelphia City Cavalry*. And then, when it became known that another Seymour had written *Lenin and the Vanguard party*, the information was added in to the record, thus rendering the NAR undifferentiated.

First Name Authority Record Edited to Include Second Person

1001 Seymour, Joseph
670 [Author of First troop Philadelphia City Cavalry]
670 Seymour, Joseph. First troop Philadelphia City Cavalry, c2008: ǂb t.p. (Joseph Seymour) cover p. 4 (Philadelphia native; US Army historian, Washington, D.C.)
670 [Author of Lenin & the vanguard party]
670 Lenin & the vanguard party, 1997: ǂb p. 3 (Joseph Seymour; SL [Spartacist League]/U.S. Central Committee member)

Then I created a new NAR for the person for whom the existing record had been originally created.

Second Name Authority Record Created to Disambiguate First Author

010 no2012108857
040 ViFGM ǂb eng ǂc ViFGM
1001 Seymour, Joseph, ǂd 1968-
667 Formerly on undifferentiated name record: n 2008044954
670 The Pennsylvania Associators, 1747-1777, c2012: ǂb t.p. (Joseph Seymour) dust jacket (historian, US Army Center of Military History; lives in Baltimore)
670 Phone call from author, Aug. 14, 2012 ǂb (b. 1968; unused middle name Alan)

670 Seymour, Joseph. First troop Philadelphia City Cavalry, c2008: ǂb t.p. (Joseph Seymour) cover p. 4 (Philadelphia native; US Army historian, Washington, D.C.)

A 667 note documents this history, "Formerly on undifferentiated name record." The author told me that he did indeed write *First Philadelphia City Cavalry* but not *Lenin and the Vanguard Party*. That left the first NAR, after I'd removed the information pertaining to the first author, just representing the author of *Lenin and the Vanguard Party.*

First Name Authority Record After Disambiguation

010 n 2008044954
040 DLC ǂb eng ǂc DLC ǂd CU-A ǂd ViFGM
1001 Seymour, Joseph
670 Lenin & the vanguard party, 1997: ǂb p. 3 (Joseph Seymour; SL [Spartacist League]/U.S. Central Committee member)

It might appear that I've changed one person into another person by means of authority control! As with all tricks, however, there is a catch, which is:

Name Authority Records represent <u>name headings</u>
Name Headings sometimes do not represent <u>individuals</u> on a one-to-one basis
Headings and individuals can become confused
PERSNAME-L can help undo confusion

A NAR represents a name *heading*. But people think of them as representing persons rather than headings. Although in many cases there's a one-to-one correlation, in other cases there is not. We have situations with nicknames, pseudonyms, people who write under one name in some cases (like using Lewis Carroll for children's stories, and Charles Dodgson for other works on mathematics and logic). PERSNAME-L can help with getting the right headings associated with each author.

This work is co-operative – it benefits your own library, but if you share the information then lots of other libraries benefit too. There is another attitude out there, though, which I deprecate, where librarians are told that they should only spend time and money on work which directly benefits their own institution, and shouldn't spend time working for the larger cataloguing community.

This is how I go about getting disambiguation. I make an initial approach, something like this:

Hello [Person],
I am a librarian, working on your book [Title].
I also found your web page here
Please respond to confirm the e-mail connection. I have a question for you.
Sincerely - Ian

This gets people on the hook, and they're usually curious. They have no idea what I'm going to ask them about. If I don't hear back, I write again.

I did not receive a response to the message below, and would be grateful to hear from you.
Sincerely - Ian

Note that I haven't said "You didn't send one" - all I know is, I didn't get one. This will often get a response when maybe they've just set it aside to work on it later, and never got round to it, and leads on to the request for information, which also tells them why I'm doing what I'm doing:

I am working on improving the cataloging of your book by making it more likely that anyone looking for it under your name will not have to contend with other titles by namesakes of yours.

We catalog librarians attempt to provide unique headings for each person represented in catalogs worldwide. Since other persons with your name are already extant, I want to create a distinct heading for you.

To this end, please provide:

- **your full name in direct order (including any unused names) as ordinarily written**
- **on its own, your surname (if you have more than one surname, all of them)**
- **the year of your birth**

The work I do is part of an international project coordinated by the Library of Congress, the Name Authorities Cooperative Project, and will be copied by other librarians who find the record I prepare. Hence my concern to do the best possible job.

Thanks again for your help. Sincerely - Ian

Afterwards all is well. I've got a newly disambiguated person, and am able to tell people on PERSNAME-L so that libraries around the world can immediately adjust their headings. There are no expenses beyond the staff time to do so.

SERIES-L (listserv @po.missouri.edu), which sounds similar to PERSNAME-L, was started on 24th March 2009. It works to the benefit of public libraries, where children in particular tend to want books in series, as well as academic libraries, where multi-volume monographs can be issued as non-consecutive volumes of the larger series.

Why so many lists? They are notification lists, not simple announcement lists, in which you sign up just to receive information from LC, or the British Library such as a new update to the MARC formats. They're not lists in which people engage in lengthy discussion. If someone needs help in finding information to disambiguate somebody that they've worked on, they can ask. Maybe off list, people will work out the information needed. Later on, we tell the list how the situation was resolved.

There are other possible projects. Why isn't there a CORPNAME-L list, for example? Because no one has started one! Not knowing what the tolerance rate is, I started these lists one by one to avoid overwhelming people. People may be able to cope with personal names, but another list for corporate names might be too much. If anybody would like to start such a list, I'd be very glad to hear from you. But the issues with corporate names are different than those with personal names. They change their names for different reasons and at different rates, and are more likely to be associated with serials, which in some libraries is the domain of a specialist.

Resources such as these need to be nursed along carefully, lest they turn into: SILENCE-L. You don't want to set up a list and have a group of people subscribe, and then, nobody says anything. That list will be destined to fade away through inaction.

Please write for information about participating in each of these projects – and setting up new ones!

Address for correspondence:

ifairclo @gmu.edu

The presentation as given to the CIG Conference in September 2012 is available here: http://www.cilip.org.uk/get-involved/special-interest-groups/cataloguing-indexing/pages/presentations.aspx#2012

A related article describing these projects in more detail has been published as:

Ian Fairclough (2012): Collaborative Initiatives in Error Handling and Bibliographic Maintenance: Use of Electronic Distribution Lists and Related Resources, *Cataloging & Classification Quarterly*,

http://www.tandfonline.com/10.1080/01639374.2012.719074

A postprint version is available via George Mason University's Mason Archival Repository Service

http://digilib.gmu.edu/dspace/handle/1920/7982

Reference

1. 'Systematic Identification of Typographical Errors in Library Catalogs', *Cataloging & Classification Quarterly*, 46: 1, 27 — 33

Changing rules, changing roles: being a cataloguer in 2012

Heather Jardine **Bibliographical Access Manager, City of London**

The one thing that everyone is agreed about is that change is the new normal, which makes for anxious times.

The first sort of change that we are facing is changing rules. Rules have always changed – before AACR2 there was AACR1, and before AACR there were rules which have replaced each other down the centuries, all the way from the Library at Alexandria. The same is true of formats – we've seen UKMARC change to MARC21, and now we are anticipating the end of MARC altogether. So changes to rules and formats are nothing new, and although they are uncomfortable and unsettling, as professional cataloguers we can cope with them.

The second change is in the sorts of things we are cataloguing and, again, we've always had this kind of change. I remember when everyone was fearful about what we then called "audio-visual materials", when we were wondering what we were going to do with sound or (even worse) *video* cassettes. Today this sort of material is mainstream, and we don't think twice about cataloguing a Blu-Ray or a Playaway or any of a wide range of formats. Now we are facing a challenge in cataloguing everything "e-" – but we'll survive, we'll cope, as we did before.

The next kind of change has happened before as well, but it is different in that we can't control it because it's not happening within our world, but happening to our world as a result of influences from the wider world outside. Very often the immediate driver is a restructuring and the immediate effect is that we are required to play a different role within the organisation to the one we were used to.

Restructures almost always result in an overall loss of staff, which means that there are fewer people to do the same amount of work or, to be more precise, to achieve the same outcomes. However much it is possible to streamline and review workflows, it still means that everyone finds themselves doing tasks that they are not familiar with, and these can be regarded variously as an unfair demand, a challenge or an opportunity.

A very common outcome is that bib services are combined so that acquisitions work is integrated more closely with cataloguing, under the same manager and with the tasks spread more widely across the whole bib services team, so that cataloguers find themselves doing acquisitions. Now, if we had wanted to be acquisitions librarians, that is what we would have chosen to be, so there can be a certain amount of bewilderment and resentment. Of course there is a lot of sense in mixing acquisitions with cataloguing, because it isn't like the old days, when the acquisitions team created paper records and then passed the books across to the cataloguing team, who created a whole different set of paper records. We're all joined up now with an integrated LMS and probably accepted long ago that acquisitions staff have to learn a certain amount about cataloguing in order to create or download bib records. The real problem with cataloguers doing acquisitions, though, isn't that we don't all like it, but that when we are doing it, we are not doing cataloguing.

Of course it isn't just bib services that lose posts and people in a restructure. So, as we struggle to keep on top of a growing burden of work, and perhaps we don't do it very quickly because we're learning it as we go along, our colleagues need more than ever that we get it right, because they too have less time and fewer staff to pick up any errors. Unsurprisingly, this leads on to a loss of tolerance and a loss of engagement. Just as we are facing changes to rules (RDA in place of AACR2), changes to data structure (whatever follows after the end of MARC) and changing materials (e-everything), my colleagues would probably say that none of this matters to them as long as it doesn't affect the speed with which we get their books onto the shelf. We're not going to have a great long intellectual discussion about any of it. Even the best-intentioned people don't have the time or the energy to be interested.

After acquisitions, the next logical area for bib services to take on is management information, especially when library systems teams are also reduced or done away with altogether. And it makes sense, because we create the database, we put the information in, and we know how to get the information out. We also understand the business, in a way that centralised IS teams don't (and can't be expected to). If we run a query and find out that one of our libraries had 39 borrowers, we know that there is something wrong with the query, whereas a central IS team wouldn't necessarily recognise it as odd. I think we're well qualified to take over the management information, because we know what we're talking about.

Speaking personally, I have got the best team ever. They can and will turn their hand to almost anything, they rise to challenges and they enjoy doing new stuff. I could not have taken on all the new work that we have taken on, if I had been dragging a reluctant and resisting staff behind me. I also think that as managers, in the current climate, we have a duty to give our staff as many transferable skills as we can. The only problem is that when we're providing management information, we're not doing acquisitions and we're not doing cataloguing.

It would be nice to think that system management would be the next area to be passed to bib services because it was recognised that we had the best, the most adaptable, the cleverest, the hardest working, the most logical staff, with the most creative solutions, in the library service (which is true). More often it will be because, once everyone realises that handing it to over central IS isn't a good idea for the same reasons that apply to management information, they also realise that bib services staff will do it for nothing, whereas outsourcing it to the LMS supplier would cost several thousand pounds a year. While our colleagues in the libraries may not have time to care about RDA, they most certainly care if books are issued for the wrong period of time, or attract the wrong charges, or if overdues aren't sent out. Therefore managing the LMS will make a bib services team more crucial to the running of the service, and more involved with it, and makes us seem a bit more relevant to the day-to-day business, which is certainly a benefit for everyone. The disadvantage is, of course, that while we're doing system management, we're not doing management information, and we're not doing acquisitions, and we're certainly not doing any cataloguing.

Because any restructuring, or hint of it, or just the general economic situation and the fears of redundancy, make us all feel uneasy, we cannot afford to neglect advocacy and all it involves. We must explain what we do, how we fit in, what we contribute to the service and how we support it, because bib services can be an easy target when the axeman comes a-calling. We must promote ourselves to fellow professionals and the library community at large, and to our immediate colleagues, and as well as persuading them that we are essential we must persuade ourselves, because too often we believe the poor opinion that others have of us. Therefore we take on a range of different activities – blogs, Twitter accounts, behind-the-scenes tours, anything and everything that makes us more visible (including going to conferences and speaking at them). If we can make ourselves look like big beasts, we'll be harder to take down.

It also includes working more closely with our colleagues – getting ourselves on working groups, doing one-to-one training, going to their staff meetings. Every time someone sends in a complaint, whether justified or not, we to try to send back a polite and reasoned explanation. Put very simply, we've got to win hearts and minds. But when we're blogging, or tweeting, or writing those polite replies, and it all takes time, we're not doing any of the other tasks.

And then there's income generation, the besetting requirement of the public sector nowadays. As a result, along with everything else, we're trying to sell our skills either by providing training, or by cataloguing other people's collections. The money we can bring in does two things – it demonstrates that our skills really do have a market value (in other words, if other people are prepared to pay us, we must be worth something) and it offsets some of our staff costs. After all, there is really no way that you can reduce the cost of a bib services section without cutting staff, as we have no materials budget to take the hit.

Does it matter if we lose staff – aside from the human cost, that is? Yes, it does. Our bib services staff are hugely skilled and hugely knowledgeable about our collections, how to exploit them and how to get them into the hands of the people who need and want them. That skill, that knowledge, once lost can't quickly or easily be replaced. And those skills and that knowledge are exactly what we need in a time of cuts to minimise the effect on the service. We must do all we can to retain skilled staff, who can make the best of the lean times and be ready when the cycle turns (and the cycle will turn) to get back immediately into growth.

So here we are, in a period of reduction and restructuring, with bib services taking on all kind of new tasks and responsibilities. As well as all the hard work, anxiety and stress, there are a lot of good things coming out of the changes. Firstly, we have learned new skills and tried all sorts of things we've never tried before – and mostly we have enjoyed it. We have certainly taken on a more crucial role in delivering services and we are working more closely with our front-facing colleagues. We are demonstrating that our skills are relevant and even marketable. We have come out of the shadows and become visible. We have even embraced social media. These things have not destroyed us but made us stronger.

Of course there is a downside too. The workload has increased for all of us, the constant pressure makes us irritable and we can't often find the opportunity to think, to plan, to look forward. We spend our time fire-fighting, and not even putting the fires out, just damping them down before we have to rush off to the next one, knowing they will flare up again behind us. And we don't have enough time to spend cataloguing.

That is the real risk, because all the jobs that we have taken on, we have been able to take on because we have good and accurate data and because we have been able to exploit the skills that we have as cataloguers. If we neglect our cataloguing, if we cut corners and start to get slovenly, the whole house of cards will come tumbling down around our ears. If the data is wrong, then the management information will be wrong. If the data is inconsistent, then the LMS will not operate reliably. If we make mistakes, then when we get those books sent back to us with complaints, we will not be able to explain anything, we will just have to apologise. Our colleagues will lose confidence in us and we will lose confidence in ourselves. And we certainly will not be able to sell skills that we cannot demonstrate that we have. So whatever else we do – and for sure we will be doing many other varied tasks – we must continue to catalogue accurately, consistently and with good judgement and imagination. However tempting it may be to do so, we must not cut corners and we must never think that "good enough" is good enough, because it is not.

The Incredible Shrinking Cataloguer meets The Spaghetti Junction Automation Robot

Gary Green Technical Librarian, Surrey County Council Library Service

Introduction

This conference paper, presented at CILIP Cataloguing & Indexing Group Conference 2012**,** focuses on a project Surrey Library Service undertook to automate some of its cataloguing and classification work, following on from the need to reduce manual input in-house. As well as the shift from manual input to automation it covers the changes to our in-house processes and the need to maintain the accuracy and quality of the catalogue.

Project background

Surrey Library Service has a presence in 52 static public libraries, two specialist libraries (Performing Arts Library, Surrey History Centre) and four prison libraries. On average we add approximately 185,000 items into library stock per year – 180,000 of which are books, the remainder being audio visual material, such as CDs and DVDs. For at least 15 years the majority of our catalogue records have been sourced from an external cataloguing company, a lesser number from stock suppliers, and we also undertake some cataloguing in-house. Our library management system uses UKMARC as its cataloguing format, but the majority of our in-house cataloguing is carried out in non-MARC format and is converted by the system into UKMARC. The in-house cataloguing team currently consists of 0.75 full time equivalent staff. Other catalogue development work is not taken into account in this figure. There are further library service staff who undertake a smaller quantity of cataloguing for specialist areas of stock. eg local history.

Even though the majority of our records are sourced externally Surrey cataloguers have always been responsible for:

- Amending suppliers' records to match in-house cataloguing and classification practices.
- Manually creating a percentage of catalogue records that can't be sourced elsewhere.
- Dealing with cataloguing and classification inconsistencies.
- Undertaking retrospective catalogue maintenance work.

Over time the cataloguing team had reduced in size, but work undertaken in-house remained at a similar level. Other automated processes (separate from the project discussed here) helped the team ensure catalogue data was added onto the system as quickly and efficiently as possible. However, the team still needed to amend aspects of suppliers' records, including:

- Adding stock format codes into each record to indicate the type of stock. For example, Paperback books, DVDs, Talking books.
- Our local classification system had to be applied consistently to every catalogue record. Up until the introduction of Dewey edition 22 we were still using a classification system based upon Dewey 16, but with its own local interpretation. We also use a stock categorisation system known as Reader Interest Categories (RICs), which pulls together items based upon non-fiction subject, fiction genre or format, depending upon the type of stock it is. We have been using RICs since the 1980s.

Requirements that needed to be addressed

Even though we had already automated some cataloguing processes over the years the cataloguing team felt we needed to do more. So, with this in mind, we asked ourselves:

- Which other cataloguing and classification processes can be automated?
- How can we automate these processes whilst recognising that our data suppliers catalogue in slightly different ways and therefore provide data that isn't consistent with each other?
- Do we need to adapt our in-house processes to achieve our aims?
- How can we make these changes whilst ensuring that all our data flows are kept in place?

- How can we achieve this whilst ensuring that it does not impact upon the consistency and accuracy of the catalogue?

The data flows were very important, as information coming into the catalogue was in turn manipulated and fed into other modules of the library management system, including:

- Electronic ordering and acquisitions (including budgets and finance codes).
- Servicing and shelving requirements.
- Circulation – where item categories are used to identify, for example, if an item is an adult fiction book, children's talking book, or DVD, etc.

The catalogue record was often the starting point in this complex web of data.

Changes in processes

Our first change was to move away from Surrey's existing in-house Dewey classification scheme and accept the current Dewey edition classification numbers supplied in externally sourced catalogue records. The decision was made not to retrospectively re-classify our stock, as we continued to use RICs that would pull related stock (both old and new) together.

Secondly, by examining the data in our catalogue records we knew we could put together data mapping rules to identify which RIC or stock type should be assigned to each record. For example, if a record contained a reader interest level of "j", an age range of "7-11" and a subject of "Children's stories" we knew that it should be assigned a RIC of "Junior Fiction". However, we had to think about how the rules could be turned from a theoretical decision into a process that would add the correct information to each catalogue record.

With all of these aims in mind we came up with a couple of options:

1. Ask our record suppliers to provide us with bespoke catalogue records that aligned with Surrey's way of doing things.

2. Continue receiving the records as they were, but manipulate the data when it's downloaded into our catalogue, so that the correct format and RIC are added.

The first option meant that the data would be correct right at the beginning of the chain and wouldn't have to go through any alteration between leaving the record supplier and being downloaded into Surrey's library management system.

The second option meant that we had greater control over the data downloaded into our library system and we could amend any mapping rules we'd set up if necessary.

We decided that the second option was the most appropriate solution for us and we were able to set up a process that performed the necessary changes on the catalogue data as it was imported into our system. To achieve this the in-house cataloguing team put together data mapping rules and our library system supplier turned these rules into programming scripts that would be run against each batch of records downloaded into the catalogue. As it was so complex and there were so many scripts to put in place we set it up in stages. At each stage we went through a testing period and tweaked any scripts that hadn't worked as expected.

Our successes and lessons learned

The solution has been in place for a few years now, giving us a reasonable period of time to assess how successful it has been.

We were successful in a number of ways:

- It helped us reduce the amount of in-house cataloguing and classification work.

- It gave us greater capacity to keep up with the daily throughput of stock.
- It speeded up the cataloguing and ordering of new stock.
- It helped improve the consistency of our catalogue and classification.
- It helped improve the consistency of data that flowed from catalogue records into other areas of Surrey's library management system.

Lessons learned from the project included:

- If any of our catalogue record suppliers changed their cataloguing style or Surrey changed internal processes we needed to make sure that our mapping rules were changed to reflect this.
- If the data mapping failed it had a knock-on effect somewhere along the data flow. For example, incorrect budget codes might be assigned to a stock order.
- The more complex the mapping, the more difficult it is to untangle it and work out where things are going wrong.
- It wasn't 100% accurate. Sometimes this was because of the quirks of Surrey's cataloguing and classification style, but also because sometimes human flexibility is needed to make decisions about the cataloguing and classification.
- You still need someone to supervise the system to ensure it runs smoothly and you still need a level of manual input.

Plans for the future

The solution has been in place for a number of years now and we have delivered on our intended aims, particularly the need to increase the automation of our cataloguing and classification processes, whilst keeping our data flows in place and our catalogue data consistent.

Some of our in-house processes have changed over the last couple of years and we have had to change data mapping rules to reflect this. We are also currently assessing whether any other mapping rules need to be updated, particularly in light of the forthcoming introduction of RDA.

We are happy with the results we have achieved so far, but want to build on these successes. At this stage we continue to check a percentage of catalogue records, but don't need to amend as many records in-house as we previously have done. Ideally, we would want to reduce the amount of checking we undertake, and just be there to address problems as they arise – keeping a supervisory eye on what is happening. We are also considering whether we continue to use the same method of automation, or can we learn from this project and implement a different method that reduces the complexity of the mapping rules, whilst achieving better results?

Ultimately the expansion of this project will give us greater capacity to focus on other areas of the library catalogue that need attention.

The past is a foreign country: transforming a Bibliographic Services team from copy cataloguers to metadata creators

Helen Williams Assistant Librarian, Bibliographic Services, LSE Library Services

Fans of L.P. Hartley might recognize that the title of my article is taken from the famous opening line of 'The Go-Between', published in 1953. I chose this because the line 'The past is a foreign country: they do things differently there' (1) feels like an accurate description of how much both my role, and the work of my team, have changed in my time at LSE. So much so that the past really does look like a foreign country, and our work now looks quite different from how it did in December 2005 when I took up my post here.

Founded in 1895 LSE itself has a rich history as a specialist university for teaching and research across the broad spectrum of the social sciences. LSE Library Services began a year later in 1896. This paper fast-forwards just over 100 years, and focuses in on the second floor of the Library where more change was about to occur as the Acquisitions and Cataloguing teams merged to create the newly formed Bibliographic Services team.

This combined team meant that some in-house training was needed in order to help acquisitions staff with their cataloguing skills, and vice versa. Once this was complete, everyone was competent in copy cataloguing and in basic 'from scratch' cataloguing which was enabling the team to add 26,000 items to stock each year. The greater flexibility of the combined team meant that we were also able to carry out projects adding large collections of historic donations to stock.

While we enjoyed the challenges and successes of these changes, downstairs on the first floor, our Academic Services colleagues were also going through a period of change as the LSE Institutional Repository was developed. LSE Research Online, which we refer to as LSERO, began life in 2004 and 'aims to be a complete database of research created at LSE'. (2) Two of its key mission points are to 'include citations to the work of all LSE academic and support staff', and to 'provide Open Access to full text research where permitted by publishers and copyright law'. (3) The research is in the form of articles, working papers, datasets, books and book chapters, conference items, newspaper articles, research blog posts, etc, and at the time of writing, in July 2012, LSERO holds 23,494 citations, 6558 of which contain full text (though these figures increase every day).

All initial work on the repository was the responsibility of the E-Services librarian, based in Academic Services, and temporary project staff. When I joined LSE I inherited an RSLP (Research Support Libraries Programme) funded post working on LSERO, but at this stage, my line management responsibilities were really in a 'pay and rations' sense, while all repository related work was directed by the E-Services librarian. I had some input to development in terms of advising on metadata, but little hands-on practice with day-to-day workflows. As the repository continued to grow, more staff were appointed, including a repository manager in 2009, which took the total staffing level of the repository to 5.6 FTEs.

At this stage everyone involved in the repository was working on a bit of everything. Growth had been so rapid that it had not really allowed time to develop streamlined workflows across the increasing number of staff. Important decisions were noted down in a shared drive to ensure consistency, but there was not a stored body of instructions in the same way as we have with traditional cataloguing procedures. With LSERO continuing to grow there was a pressing need to scale up the processing of incoming material by streamlining the workflows so that a higher level of throughput could be consistently maintained.

Towards the end of 2009 the post I line-managed became vacant which gave us an excellent opportunity to think about making some changes. As a result we suggested that the workflows we had in place for our print materials should also apply to our digital materials, mirroring the existing responsibilities for print items between the Academic Services and Bibliographic Services teams. So academic liaison, content recruitment and advocacy would take place in Academic Services, while acquisition, processing and metadata creation would be done by the Bibliographic Services team. Not only would this mean that areas of expertise were exploited to best advantage, but for a service which had a cross-departmental workflow it would be clearer for line-managers to ascertain who was responsible for particular areas of work.

We also suggested that the vacant specialised repository post in my team was instead recruited as a general Bibliographic Services Library Assistant, and that then the entire team took on LSERO work. It would increase our flexibility to tackle changing workload priorities and also 'future-proof' the team by using their skills in following standards, checking data, and working with a high throughput of material while still maintaining accuracy and consistency.

It sounds a simple and straightforward proposal, but behind the scenes I had been doing a fair bit of number crunching to ensure that I thought we really could tackle this new dimension to our work, while not losing any of our efficiency in getting those 26,000 items per year onto the shelves.

Out of a 35 hour working week, one of my Library Assistants was spending an average of 12 hours a week on public service point duties, three hours on scheduled breaks and four hours on miscellaneous tasks (such as dealing with post and emails, or involved in training). So the full time LSERO post in Bibliographic Services had been contributing 15 hours a week, or three hours a day, to the work of the repository. With LSERO high on the agenda in the Library's strategic plan, the new proposal made an assumption that LSERO was a greater priority than cataloguing books which had been donated as part of big collections in the past. This did not mean we would completely abandon donations cataloguing, simply that we would spend less time on it. My statistics showed me that in the previous year my team had added 3098 donations to stock. Assuming those had taken an average of ten minutes each, that makes 516 hours spread across eight library assistants, so 64 and a half hours each over the course of a year. Divide that by 50 weeks (which takes into account closure days and bank holidays) makes 1.3 hours a week each, which multiplied by eight equals just over 10 hours a week as a team. Generously rounded up, that equates to two days. Two days, when we take into account the other tasks Library Assistants spend their time on, is equivalent to six hours of metadata entry. So, if we amalgamated the specific LSERO post into the Bibliographic Services, team, and added those 15 and six hours together, the team as a whole would be able to contribute 21 hours a week towards LSERO, while still keeping up with our acquisitions and cataloguing work. I would expect it to take a maximum of ten minutes to create a metadata-only record, so that would be a minimum of 126 records a week contributed to the repository, or a full text record would take 20 minutes, so that equates to 63 records. In practice, we are adding a combination of record types each week.

Having established where the time for LSERO work would come from, we needed to set up how it would be managed as a shared service between two teams. Academic Services and Bibliographic Services sit on two different floors of the building, so effective communication methods were going to be crucial, as well as setting out the work and responsibilities that each of the teams would have. The majority of our work arrives via a shared repository email inbox, to which academic staff email details of the research they want added to the repository. As multiple staff worked on this inbox, everybody had their own flag colour for clear identification of who was dealing with each request. An initial snag was that there were not enough flag colours to allocate one to each member of Bibliographic Services, though on reflection this turned out to be no bad thing. The Academic Services team would retain control of the inbox – they are the staff with the expertise in liaison, so they deal with all correspondence. That also takes into account that Bibliographic Services staff are in and out of the office around public service point duties, and they may additionally only spend one day a week on repository work, so if they took responsibility for replying to emails, there could be a lack of continuity for users. Instead emails are flagged green for us, by the Academic Services team, we work on the data entry and then flag an email purple if it is complete, or orange if there is an outstanding query. A purple flag tells the Academic Services staff that they can reply to say the work is added to the repository and then file the email, while an orange flag tells them some more liaison is needed before the work can be completed.

I had been mentioning the likelihood of Bibliographic Services getting involved in repository work for several months already, drip feeding the initial idea itself, the skills benefit it would bring to our team, and the opportunity it would give us to showcase the value of our skills to the rest of the Library, so by this stage the team were enthusiastic and ready to take on a new task. We revised the generic Bibliographic Services job description to incorporate the processing of digital materials, and we recruited someone to our vacant post. Now it was time to start the training. Having never done any LSERO data entry myself, I needed training first, along with our senior library assistant (SLA) for cataloguing, so that we would be able to roll out training to the rest of our team. The team had little previous exposure to LSERO and repositories at this point, so initially we introduced the basic concepts, alongside the impact we expected it to have on the way we worked as a team, the amount of time we would spend on it, and the expectations about what we would achieve in that time. We then undertook practical training for the team in pairs which meant training could be tailored to different learning styles. This was backed up by the comprehensive documentation we prepared to go on our Group wiki. Written workflows are always a helpful reminder, and particularly in this instance as I wanted to avoid seeing our new Dublin Core skills appearing in the OPAC, or our existing MARC21 methods appearing in the repository! Other differences it was important to explain related to name authority control and classification. All the name headings in our Library catalogue are authorised against the Library of Congress Name Authority Files, but without an established source of external authorities to follow for the repository, the team had to be careful to check internal consistency when adding creator names. The more recent addition of a name authority file for LSERO has helped us greatly with this. The team are also used to checking or applying Library of Congress Subject Headings and Library of Congress Classification to catalogue records. LSERO works rather differently, combining both of these through a classification letter with a definition attached to it. This is a very broad subject breakdown and it has been strange to get used to classifying something into a broad set, rather than specifically. The plus point, of course, has been that we save an awful lot of time in authorising subject headings!

Initially I, or the SLA, checked every record which was created by the team, giving verbal feedback, explaining any mistakes and asking people to correct those themselves. Although this sounds a bit 'big-brother-esque', and although it can be tempting to make minor edits yourself, at this 'bedding-in' stage I knew that people would learn better from their mistakes if they had to correct them. Once we could see someone was creating records of a consistently good quality we moved to spot checking, and once we were happy that people were both creating high quality records and felt confident in doing so, they were able to deposit records themselves straight into the live repository. Because at this stage the whole team were learning and we wanted to keep an eye on how everyone was getting on, we did not instigate peer review checking, though that has been used since when we have been training new or temporary project staff. Basic metadata was only the beginning of the training. Over time this extended to dealing with full text materials and more complex item types, each time requiring more training and the production of more documentation for the wiki.

Over the last year my statistics show that we have spent an average of 24 hours a week on LSERO, so not a great deal more than the 21 hours we had originally planned to spend. However, we have less control over the peaks and troughs of LSERO than we do over accessioning and cataloguing, so in our lowest month we averaged 16 hours a week on LSERO, while in our busiest months we averaged 42 hours a week, which is double what we had originally said we could provide while still maintaining the 'bread and butter' kind of work a Library expects to come out of Bibliographic Services. Despite all the juggling we have had to do behind the scenes, as a team we have been very keen to demonstrate our flexibility and our ability to take on new challenges while providing a high quality service. We had a particular opportunity to do this in December 2010 when all 22 academic departments were asked to submit their post-2008 research to LSERO as soon as possible as part of REF preparation. (REF stands for Research Excellence Framework and is the means by which the quality of research in UK higher education institutions is assessed in order to inform funding.) In the space of three months nearly 5000 new citations were created (by us and Academic Services staff) which was an increase of around 20% of the total size of LSERO at the time (4), and proof that the Library was worthy of its central role in REF activity at LSE.

As a result of our LSERO work new opportunities continue to come our way. In summer 2011 electronic submission of PhD theses was approved by our Research Degrees Unit and these are being stored in a similar, but separate, repository called LSE Theses Online. Expertise in cataloguing print theses lies with the Bibliographic Services team, so it was natural that we would volunteer to be involved in populating LSETO. Over the coming year we expect to contribute to the creation of metadata for two additional proposed repositories for official publications and exam papers respectively.

My reflections as I look back are on what a big change this was for our team, but also on how successful it has been. The long run-in period allowing staff to get used to the idea of their work changing was beneficial, and the whole team are now trained to deal with LSERO metadata. In the last couple of years we have seen a slight decline in the amount of print material we receive, alongside a dramatic increase in electronic materials. Furthermore, our print processes have speeded up thanks to actions implemented as a result of process reviews. By using their cataloguing skills to become competent metadata creators the team have learnt that their skills are transferrable, compatible with the electronic environment and make a vital contribution to the strategic priorities that are met through LSERO. It has been encouraging to see the team take pride in that, and know that their skills are relevant and valued. By learning now to adapt our skills to keep up with the fast-changing environment in which we are operating we have, I hope, laid down the foundations necessary to 'future-proof' the team in changing times.

References

1. Hartley, L.P, *The Go-Between*, Penguin Books, 2000, p. 5

2. London School of Economics and Political Science, *What is LSE Research Online,* http://eprints.lse.ac.uk/faq.html

3. London School of Economics and Political Science, *What is LSE Research Online,* http://eprints.lse.ac.uk/faq.html

4. Stewart, Neil, and Puplett, Dave, *A repository manager's experience: mini REF at LSE,* RSP – Readiness for REF (R4R) Workshop, 5th September 2011 http://www.rsp.ac.uk/events/readiness-for-ref/

The problems of cataloguing in higher education

Jennie-Claire Perry, Project Co-ordinator and Robin Armstrong Viner, Head of Collection Management, both at the Templeman Library, University of Kent

Introduction

The Higher Education (HE) sector in the UK is experiencing a period of fundamental and rapid change in all areas. Higher tuition fees for many students are a key driver, and successful HE institutions will need to be agile in their response to the new landscape to ensure they attract the high quality academic staff and student numbers they seek. This new marketplace demands that the whole organisation demonstrates both an improved student offer and best value. However this efficiency imperative represents an opportunity for cataloguers in HE libraries to examine our structures and practices to determine new ways of working that contribute to the needs of our wider institutions.

The new HE marketplace

Following the recommendation of the 2010 Browne review that the cap on university tuition fees be removed, the UK government began to shift the emphasis of HE funding from HEFCE and other central sources to income from student fees. Although the increase in tuition fees doesn't actually equate to extra income for HE institutions, fee-paying students may expect higher standards of teaching and services than before. This in turn will be reflected in the National Student Satisfaction survey and other league tables – the benchmark established by the UK government against which HE institutions are judged by their users and external commentators, including those applying to university.

If tuition fees are not generating extra income, then how can cataloguers achieve the demonstrable increase in services and standards demanded of us in the new marketplace? For those of us already working in the sector restructuring, budget cuts and efficiency drives are all too familiar and are likely to continue. In an increasingly competitive environment, HE institutions will need to focus on balancing what we offer our 'customers' and the cost of delivering that offer. As we have seen elsewhere HE libraries, with high staffing costs and practices that are little understood outside the profession, are an easy target for budget cuts. As such demonstrating the value of our services to our colleagues is as important as to our users.

The cost of cataloguing

With the emphasis on student experience there has been a focus on streamlining back office processes. As cataloguers we are increasingly being asked to justify our cost and quantify our contribution to the student experience. Anticipating this, the Metadata & Processing team at the University of Kent has undertaken a systematic revision of our processes in order to demonstrate that we deliver a cost-effective, sustainable and valuable service to both our service users and our wider institutional community.

Prior to the service review the cost of cataloguing and processing each new item at the University's Templeman Library was between £5 and £11 on top of the cost of the item itself and excluding processing materials. The backlog of new items awaiting cataloguing was such that an estimated £100,000 of new stock was sitting on trollies in our offices, where it could not easily be accessed by users and where it could be waiting for up to six months. With the throughput of the team roughly equivalent to the 17,000 print items purchased each year, our cataloguing backlog would probably never diminish.

It is difficult to justify these costs and delays to our users in terms of benefit to their academic outcomes or experience of the library. In Lean methodology value is defined by the user; would a user actually pay upwards of £5 just to have the item added to the catalogue with enriched metadata? How much added value was this metadata actually giving? Focusing on the user throughout, our review sought to address these issues.

As part of a review of our acquisitions procedure, we undertook a shelf-ready trial that included significant alterations to previous cataloguing practice. Prior to the shelf-ready trial, the majority of new stock had been purchased at the campus book shop and arrived with minimal processing and no catalogue record supplied.

Bibliographic records were downloaded from a variety of sources and then edited according to cataloguer's judgement to an 'acceptable standard'. There was little sharing of procedures and standards and much was left to the individual cataloguer to decide, a costly process and slowing throughput considerably. The move to shelf-ready purchasing meant that the team needed to define that acceptable standard, so that we could ensure that the bibliographic data we purchased from our suppliers them was of sufficient quality. This meant that the introduction of shelf-ready has increased the consistency of our cataloguing, rather than detracting from it as was initially assumed.

The bibliographic records supplied are batch loaded into the library management system on a daily basis, which means that cataloguers no longer see every single item that comes through the department. However, regular quality checks and reports will mean that they no longer need to and any editing of records needed that cannot be automated is now carried out by a team of library assistants trained to spot errors and able to refer any issues on if necessary. Throughput has increased significantly, giving our users faster access to the materials they need and demonstrating to our colleagues that we can respond to their needs effectively.

Our shelf-ready process is one of a number where we are striving to achieve best value and we are continuing to explore ways in which we can further streamline our operations. We are using process mapping to identify areas where procedures can be made more efficient or in some cases removed from our workflow altogether and plan to outsource some of our more routine quality control. This reflective approach will continue, and through it we are hoping to embed a sense of the need to drive for best value alongside best practice at the heart of everything we do.

Is cataloguing a dying art?

Of course, developments such as shelf-ready cataloguing inevitably lead to questions around the sustainability of cataloguing in HE libraries. If we can specify, check for and achieve sufficient metadata quality without having to select and manually edit records ourselves then do we need cataloguers at all? Challenging though this situation initially appears for us as cataloguers, it presents huge opportunities for the profession. By removing the need for highly-skilled metadata staff to intervene at an item level with the items common to many libraries, they can be freed up to spend more time working on those rare and unique items that make our collections distinctive and important to both our users and the wider HE community.

Many cataloguers in HE institutions are already involved in the description of special collections and research outputs in institutional repositories. However, there are many other high value areas where cataloguers' skills can add to their institutions' efficiency and effectiveness. The drive to make research data available on open access requires new metadata schemas and structures while data protection and freedom of information mean our knowledge is vital to ensuring both best practice and compliance in information governance and records management.

Of course, technological developments in library service delivery will also rely heavily on cataloguers' skills. Anyone that has ever moved from one version of MARC to another knows how important it is to have consistent metadata in order to crosswalk between standards, and the imminent introduction of RDA at several of the libraries that supply catalogue records to our vendors means that expertise in this area will become operationally vital once again.

What we have to be realistic about is the need for continual revision and adaptation of our work practices to fit into the new, competitive institutions in which we work. For many HE cataloguers the days of being able to agonise over the correct MARC field for a given situation are long gone, and the time for manually editing every record to fit (possibly unnecessary) local standards is long past. For some of us, this will mean looking to collaborate with colleagues in other institutions to cut down on duplication of tasks. We might want to consider the automation of our most routine tasks, and the harnessing of batch editing to move from micro- to macro-cataloguing. Perhaps we will need to standardise our practice to benefit from outsourcing possibilities, removing local adaptations to MARC and classification schemes to enable greater integration with external companies. We need to show that we are flexible and that we understand how our core business has changed in order to demonstrate that we are still relevant in an era of cost-cutting and efficiency drives.

Evolution vs. extinction

So the story isn't one of extinction of a once-vital library skill, it is more about the need to continually evolve, not just as cataloguers but as information professionals. Broadening our skill-sets and keeping an eye on developments in technology, educational politics and finance are becoming crucial to our marketability within the profession. Unless we understand these changes in HE and move forward we risk conforming to the anachronistic cataloguer stereotype and seeming ever more irrelevant and expendable to those inside and outside our teams with responsibility for budgets. For many HE cataloguers maintaining the status quo is simply no longer an option; we need to show our ability to step up to these new challenges and demonstrate that our teams are flexible, dynamic and relevant, moving into this era of change in higher education.

High visibility cataloguing

Céline Carty and Helen Stein

This was an introduction to the HVCats initiative and, in particular, the new Cat23 idea. As it is changing and developing so fast, see all the latest updates on the HVC blog!

Aiming for visibility: promoting a university cataloguing department

Karen F. Pierce **Cataloguing Librarian, Cardiff University**

Introduction

In this paper I am going to discuss several ways of (hopefully) ensuring your university cataloguing department is practising high visibility cataloguing as best it can. In order to do this I shall be focussing on my own department and experiences for the main examples. Before we get down to practicalities, however, we need to decide just what 'high visibility cataloguing' actually is. I'd like to think that it is a fairly self-explanatory concept, but for the purpose of this paper I am looking at the idea of cataloguers and cataloguing departments making certain they don't disappear under the radar in their own institutions, as well as in the wider world. If library colleagues don't know who you are or what you do you may find yourself getting sidelined or marginalised, and the importance of your job may diminish in the eyes of those who don't understand the complexities of your role.

We are all aware of how easily librarians in general get boxed into stereotypes, and this is probably even more true of cataloguers. In 2002 Richard A. Murray encapsulated this image in his post on the LIS Career blog, where he stated:

> *"The stereotype of the cataloger is, for many, the hermit hiding in the bowels of the library shackled to an OCLC terminal all day, counting pages of plates and measuring the heights of books. On the rare occasion he or she is let out of the dungeon, it's to be the one at meetings who speaks in unintelligible MARC-ese about "non-filing characters" and "second indicator blank" and "space colon space." The cataloger's role in the library is to enforce rules that nobody understands and to make things as difficult as possible for everyone involved."* Richard A. Murray (2002) http://www.liscareer.com/murray_cataloging.htm

Following on from that perhaps familiar image, last year (2011) on the LIS New Professionals Network blog, Theresa Schultz caused a stir in the world of cataloguing with her post on the role of a cataloguer which was written with a whimsical but negative viewpoint. She started the post with the comment *"Is there any position more dreaded than 'cataloguer'? Not because they're scary, but because none of us really want to do it?"* She went on to note that, *"I can't think of anyone I went to LIS school with who liked cataloguing. We all thought of it as a necessary evil."* http://lisnpn.spruz.com/pt/The-anatomy-of-a-cataloger/blog.htm. Plenty of people sprang to the defence of our profession but it is clear that if fellow librarians still have a negative idea of who or what we are, then we need to continue to promote ourselves in a positive way to colleagues as well as to the general public.

For this paper, as I am going to use examples from the department I currently work in, I'd like to give some background to our set-up. The cataloguing department, as well as the rest of Collection Management, currently resides on the first floor of a University administration block. The department used to be in the basement of our Arts and Social Studies Library, but was moved about eight years ago. Our area is open plan and fairly spacious, and close to the city centre, which is handy for lunchtime shopping trips. We are however not in a library, or next to a library, or anywhere particularly near a library; hence there are issues of invisibility for our team. When the department was located in the basement of the Arts library there were opportunities for the cataloguers to mingle with other library staff at break-times, to exchange news and gossip, and to be able to run upstairs to pull a book off the shelf; but most importantly there was the chance to be visible as part of the University Library Service.

In my previous role working for the University of Wales College of Medicine, prior to its merger with Cardiff University, the cataloguing and acquisitions staff also had to do a fair share of work on the issue/information desk, and thus were equally visible along with other staff members. I'm sure there are many cataloguers in institutions that also spend time on the reference desk, or have roles that are combined with that of subject librarian, which is helpful for maintaining visibility. (I should point out that one of the cataloguers in our department is a subject librarian part-time, and a cataloguer part-time, but she is unique in that role).

Being so removed from library spaces, it is imperative therefore that we make an effort to have a visible presence within the University Library Service so that library colleagues know who we are. We do this in a number of ways, most of them fairly obvious, but it always pays to get involved.

ULS Briefings

Every couple of months the librarians and technical specialists who work within the library service meet up for a morning of updates in sessions known as ULS Briefings. A programme is arranged in advance and we listen to various individuals informing us on what is going on in their project or library. It is important to us as cataloguers to attend these events, to be seen, and to talk to colleagues. Even better, we sometimes have the opportunity to present at the Briefings ourselves, and have utilised the gatherings to inform staff of changes to cataloguing practices that will impact their libraries.This is a much better way of sharing information rather than just sending out an email; and also ensures that the other librarians have the chance to discuss these changes or issues with us in an open forum.

Working groups

Our library service is full of working groups, and tries to have a cross section of staff involved where appropriate. By becoming a member of a working group as cataloguers we are providing a gentle reminder that we exist. Between us, my colleagues and I sit on a variety of groups, such as Staff Development and Engagement, Communication, PRIMO project group, Collection Development, Repository, and the SCOLAR forum; some are of obvious interest, but others are not directly related to our day-to-day work and so give us opportunities to work with colleagues we wouldn't necessarily come into contact with normally.

Staff development

We are fortunate in having a ULS Staff Development and Engagement Group and an annual programme of staff development events. I sit on the group, and am able to ensure that the training needs of the 'technical specialists' don't get forgotten. Twice a year we have a 'Do Something Different Day' (1) where staff are offered the chance to work in a different site or role for the day; and the collection management department is always a favourite, as we generally find that many staff are really curious about what goes on in this 'mystical' place! The last DSDD in May was also opened up to other libraries within Cardiff, so there were cross-sectoral opportunities available. We have also had 'tours' of our department arranged in the past, although we feel it's not very exciting just seeing people sitting at their desks; so the DSDD at least offers the chance for an individual to have a more in-depth look at what we do. This is not to say that tours aren't a good idea, because they do at least bring colleagues into your work space and demonstrate practices in situ.

As part of the Staff development programme this year instead of a tour we offered a session entitled 'From pink slip to black box' which encapsulated the journey of a book from being ordered to being catalogued, and then arriving in its home library. This was done in conjunction with the acquisitions librarian and proved to be very popular. It was designed to help put into context the procedures involved, and for staff to understand just how a book gets to their library. We had envisaged the main audience to be at library assistant level as we had assumed most subject librarians understood the process; however it was the subject and site librarians who formed most of our audience, which demonstrated that even those we assumed knew what went on, actually needed, or wanted, more information. It is likely the session will be repeated at some point, although we did reach our target audience in the end, as we were given the opportunity to run the session again at the Library Assistants' Away Day in the summer.

Conferences

Much of what I am touching upon involves being visible in your own institution, but alongside that aim is the aspiration to become visible to the wider world. The first step should be to attend a cataloguing conference, and as was ably demonstrated at the Cataloguing and Indexing Group's conference in September, cataloguers do get let out of the basement on occasion. Being able to attend conferences or training events has a multitude of benefits, not least because you are learning from your peers and from experts; but you are also being seen, you are making contacts, and talking to people who might be able to offer help and advice at some future point. You are promoting your own visibility within the cataloguing world.

The next step is to present at a conference. Obviously this is a little more scary than just turning up and listening to others, but it does help towards promoting yourself further as an individual, and in promoting the work you may be doing, or the project you may be working on. As a presenter you are also more likely to have people come up and engage with you, and enter into a discussion about the topic you have been talking about, thus widening your circle of contacts.

A further step is to attend and/or present at a non-cataloguing conference. It's all very well talking to other cataloguers (and it is fantastic being at a conference where practically everyone is a cataloguer); but we also need to step outside of the 'echo-chamber' and away from our enclosed metadata world, and try to become visible elsewhere. Attending non-cataloguing conferences is not always possible for a cataloguer, as you may not get time or money to do so. On a personal level I was lucky this year and managed to attend five conferences, and through a combination of organizing, helping out, and presenting, I did so at a minimum cost level too. Three of these events were not about cataloguing; and I didn't learn anything that would directly help me in my cataloguing work, but I did gain a lot that helped me as an information professional in general. I may well have been the only cataloguer who attended these events, I don't know for definite, but in any case I was certainly in the minority. This also meant that I felt a bit like I was representing the world of cataloguing as a whole, as my very presence demonstrated that we don't just sit chained to our desks but are just as keen as other librarians to interact and connect on a professional level.

Once you've attended a conference it is also a good idea to share your experience, partly because you probably have information you wish to disseminate amongst colleagues, but also because it's a great opportunity to let non-cataloguing staff know that you have been engaging with people on a professional level outside of your institution. Our Staff Development and Engagement Group brings out a newsletter four times a year; this is an ideal place to write up attendance at conferences, but if you don't have a conduit like this, consider emailing a piece about your exploits to all your colleagues instead.

If there isn't a conference that is relevant to you, or near enough to attend, why not consider organising your own? This may sound like a step too far, but can be incredibly rewarding. In March I organised a cataloguing conference entitled 'Conversations with cataloguers in Wales' – aimed primarily at those who catalogue or work with metadata in Wales, although anyone was welcome to attend. I did this for a number of reasons.

1. Location – for many people who live/work in Wales getting to events in England can be a bit longwinded. Not so much for those of us who are in South East Wales, but the further West and North you go, there are transport and time issues.

2. It can be good to have a regional focus, and get together with people who work in the same area, and who are possible useful contacts.

Lack of opportunities – There aren't that many cataloguing conferences being held in Britain, so there was certainly room for another one.

I deemed the event a success with over 40 attendees, more cataloguers than I originally thought I would attract to a one day event in Wales. I was also aware, through Twitter and blogs, of support from virtual contacts, and that people the other side of Britain would have liked to attend, but couldn't due to distance (it's a two-way problem!). However, it was only a few weeks later that I discovered I hadn't advertised the event enough internally. Obviously all the cataloguers in my department attended; but after re-presenting my joint paper at the following ULS briefing, some of the subject librarians said they would have liked to have attended the event had they known about it. I was fairly confident I had contacted all library staff, but I had obviously not done so in the right way to attract their attention. While originally feeling pleased that people had travelled all the way from North Wales to attend, I was now frustrated that potential attendees on my own door step had missed out. Maintaining visibility is an ongoing process, and I learnt that internal publicity for a cataloguing event may need to be delivered in a different manner to external.

Being seen virtually

While you may or may not get the opportunity to promote your presence in the physical sense, these days we must also be aware of our virtual presence. Although engaging with social media and web 2.0 may not be everyone's cup of tea, there are many benefits in doing so; and if you are aiming for high visibility then it is imperative that you get involved. I'm not going to tell you all the places you might want a virtual presence, it's best to explore for yourself and work out where you want to be, and what tools you want to use and why, therefore I just want to highlight a couple of the main ones.

The most obvious example is Twitter which houses a fantastic community of cataloguers. Librarians in general have embraced Twitter and utilised it in a professional capacity way beyond many other professions. As a cataloguer it is a great place to have queries answered, to canvass opinion about certain practices, and to engage in debates on topics of interest. It's also a place where cats, knitting, and cake feature fairly prominently too! Whether you can access it at work or at home, Twitter can be a productive tool if you let it and can widen your access points to information, bringing you into potential contact with blog posts and online articles that will enhance your professional life. And of course, it allows you the space to become more visible as a cataloguer too. You could also have a Twitter account for your Cataloguing department in general; this is not a route that we have gone down so far, but many of the libraries within our Library Service have their own accounts, and are able to communicate with different user groups this way.

If 140 characters are not enough for what you have to say, then blogging is the way forward. If you haven't set one up before, check out the cataloguing blogs that are already out there. I'd recommend going to Planet Cataloguing which is an aggregation of cataloguing blogs, as well of course as CIG's blog and the High Visibility Cataloguing blog. In addition to a personal blog, for some time now myself and colleagues have been mulling over the idea of setting up a blog for our cataloguing department, as a way of making ourselves more visible. Several of our site libraries have blogs as well as Twitter accounts, and we are already involved in the Special Collections (SCOLAR) blog – although it seems a lot easier to blog out about rare books than new books, especially when there are so many lovely illustrations one can include. The first decision we really need to make though, is about who we are hoping our chief audience will be. Are we going to be talking to fellow cataloguers or librarians? Or are we aiming at communicating with the staff and student of the university, like the site libraries are doing? While pondering this question I was interested in seeing that the Bibliographical Services Section of the City of London Libraries had set up a blog this year, and am following their approach with interest.

Another aspect of being seen virtually is through publishing. If you feel that you have something to say about your job or a project you have worked on, then why not do so in a published format? You will be raising the profile of yourself, your project and your department; adding to your CV; and an article could possibly lead to potential future opportunities or collaborations. If writing for a peer-reviewed journal is too daunting, then perhaps offering something to CIG's *Catalogue & Index* periodical might feel more achievable. If you don't think you have anything to write about, then perhaps offering to do some book reviews might be a way forward.

If you do want to publish in a journal, find out what is available, look at the kind of articles and topics different journals publish on, and check out their impact factors. Of course, the ones with the highest impact factors might also be the hardest to get published in, so you may have to lower your sights slightly. If a journal is fairly young and hasn't had time to build up a reputation they may be more willing to publish new writers. If your first attempt is rejected, pay careful heed to the feedback – it may be that you just picked the wrong journal to submit to. Look out for the chance to collaborate with someone more experienced or even an established academic. A few years ago I was the second author on a paper that was published in the journal *Clinical Genetics*, and was about the Human Genetics Historical Library for which I am the sole cataloguer. (2) The main author, Prof Peter Harper, was the originator of the project, and a renowned authority on various aspects of genetics. He was keen for my involvement in the paper, and has subsequently been encouraging me to write up the project for a library journal. Contacts like these are extremely valuable. Which brings me on to…

Unexpected opportunities

Due to their very nature you cannot plan for unexpected opportunities, apart from being open to new possibilities. Sometimes you need to volunteer for those obscure projects, because it may lead to you gaining new skills, getting published, becoming a project manager or even ultimately a promotion. Of course there are also times when we need to learn to say 'no' because we are already overworked, but occasionally we need to make the space for something because it could prove more valuable in the long run; finding that distinction can be tricky.

Engaging with academics can fall within the 'unexpected opportunity' category. Depending on your exact role you may not have the chance to interact with academics within your institution, but if you can, then you may have unique opportunities to promote the role of the cataloguer, and enhance their use of the library. As an example, I have recently co-curated an exhibition on Arthurian material within our special collections, with one of our academics. The opportunity arose, partly because of my role with the rare books, but also because I took several evening classes with this tutor, and was able to point out material she was not aware of that could enhance her research and teaching. This was certainly an unexpected opportunity, and it did arise partly through activities undertaken outside of work; but you never know when something might happen.

Conclusion

Although most of the ideas I have raised may appear fairly basic, sometimes it is easy to overlook the obvious. Almost by definition, cataloguing takes place behind scenes, in basements and back offices, or remote admin buildings; and if we are doing a good job, creating great records that ensure items are easily found on the library catalogue, then nobody notices we are there. The role of the cataloguer is undeniably changing, and we should be ready to embrace new challenges and possibilities; but if no-one knows about you, then they might not realise that the cataloguer is the best person to take on those changes.

If you are going to engage with high visibility cataloguing you need to ensure visibility for yourself, and for your department; this means being visible within your institution, as well as to the wider world; it means being visible in person, but also virtually/online. It may also mean promoting your individual persona as an information professional in general, as well as a being a cataloguer. Be seen and be willing to take unexpected opportunities!

References

1. For more information on this scheme see Earney, S. (2011) 'Do something different days' at Cardiff University library service. *SCONUL Focus* 51: 70-72. http://www.sconul.ac.uk/publications/newsletter/51/20.pdf

2. Harper P.S. & Pierce K. (2010) The Human Genetics Historical Library: an international resource for geneticists and historians. *Clinical Genetics* 77: 214–220.

Karen F. Pierce, Cataloguing Librarian, Cardiff University

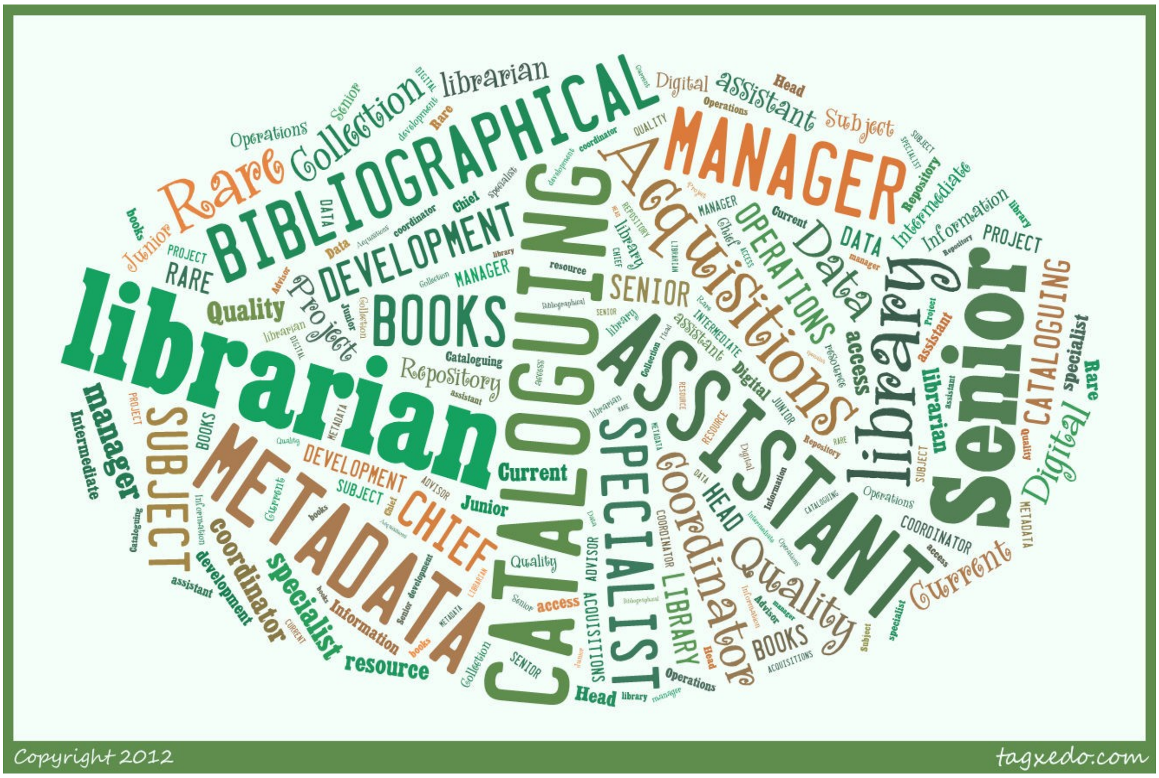
@Darklecat

http://darksideofthecatalogue.wordpress.com/

Trigger's cataloguer: professional development by default

Rachel Playforth

Trigger's cataloguer is my version of a philosophical problem known as Trigger's Broom (or more classically, the Ship of Theseus). Trigger, as fans of eighties situation comedy will surely know, is a character in *Only Fools and Horses* who works as a street sweeper. He once claimed to have had the same broom for 20 years, although it had gone through 17 new heads and 14 new handles. The philosophical question is: can something that's had all its original parts replaced still be thought of as the same object? So in terms of professional identity, can I still be considered a cataloguer if I'm not doing any cataloguing? And what if I AM cataloguing, but I'm called something else altogether? When I used the High Visibility Cataloguing blog to look into our evolving job titles, as well as what we actually do, it was clear that many of us are having the heads and handles of our cataloguing brooms gradually replaced.



In my own case, I started at the British Library for Development Studies nearly four years ago as a cataloguer, but have moved into the role of repository coordinator. This involves managing a digitisation project with international partners and launching our own institutional repository. So does a role change always involve wholesale reinvention, or could I bring my cataloguing skills with me? Cataloguing for me has always been about increasing access, and full-text repositories are part of that same goal, so although I'm doing a lot less traditional cataloguing there is an obvious overlap in the skills required. I'm still using my cataloguing expertise to describe items and make them more accessible.

However, putting my skills into a new context where I was managing not just metadata, but people as well, did involve what you might call professional development by default. Here are a few of the lessons I've learned (the hard way) about managing change and juggling roles:

1. First, catch your expert

When faced with a new area of work, I ask myself four questions:

1) who else is doing this?

2) are they doing it better?

3) can they teach me?

4) can they do it instead?

Although you might be doing a wider range of tasks, you can't possibly become an expert in every area. A good use of time is probably to find out who IS an expert in these areas, either in your workplace or elsewhere, and try to make use of their expertise. We often don't get formal training when our job changes, especially if it's a gradual change, but there are lots of formal and informal networks out there which you can call on for help. It's even worth asking the question, should I be doing this at all? Duplication of effort is common in large organisations, so make sure with every project that you're not reinventing the wheel.

2. Stay in control

Be strategic about taking on new responsibilities. Even when it feels like things are just being thrown at you, you have a choice about how to react. Work out how (and if) a role change would benefit you in the short and long term and let you follow your interests. But do be open-minded and prepared to find a new passion.

3. Do less with less

Much as we're always asked to, you can't do more with less. It's more a question of making the less better - quality rather than quantity. So my first tip is an obvious one: say no. But not just to the things you don't want to do; sometimes you should also say no to things you DO want to do, which is much harder. Make sure you have a clear job description so you feel comfortable about turning people down, too. Managing information overload and social media as well as formal obligations can be tricky – the active library blogosphere and Twitter community give us glimpses into a multitude of new ideas, developments and events, and offer the illusion of keeping on top of it all with lists, digests and RSS feeds. But it is an illusion. A narrower focus isn't going to make you a bad librarian, and sometimes it's better not to know what you're missing. Even a small amount of judicious unsubscribing and unfollowing can make a difference – ignorance can be bliss!

4. The myth of time management

You can't manage time. You especially can't make more of it, no matter how organised you are. What you can do is manage yourself, and other people's expectations. A really useful approach when you're trying to work out when and if you can deliver something is to remember the difference between effort and duration. A task may involve seven hours' worth of effort but if you have a full day of meetings and then a day's leave before you can start it, you need to budget three days to do it, and tell other people that that's how long it will take. (This is especially important for part-time workers.)

5. People vs cataloguing

I'm being reductive here but cataloguing is essentially a linear process - something is not catalogued, we follow some steps, then it is. Collaborative work is subject to a lot more iterations, delays, conflicts and shades of grey; there's no single point in my repository project where I can say 'that relationship has been successfully managed'. Other people always have an agenda, and their goal for the project may not be quite the same as mine. It's also undeniably scary taking on a new role in a new area of work. A useful mindset to get into is one where you believe that nobody wants you to fail (because after all, why would they – it very rarely benefits people when their co-workers and project partners are failing!) This is also known as 'reverse paranoia'. The more reverse paranoid we can be, the more we'll be able to develop as cataloguers, and beyond...

Cataloguing in Wales: Are there professional benefits to exploring regionality?

Karen F. Pierce — **Cataloguing Librarian, Cardiff University**

Introduction

This paper will take a brief look at what I am calling 'regional issues' (such as physical, geographical, language); the embryonic development of an all-Wales group, and co-operation between geographically (and institutionally) separated libraries on one catalogue.

A common identity

> *"…a social category…into which one falls, and to which one feels one belongs, provides a definition of who one is in terms of the defining characteristics of the category – a self-definition that is a part of the self-concept."* Hogg, Terry & White (1995, 259) (1)

During the Cataloguing and Indexing Group conference in Sheffield in September, the attendees came together, from many different parts of the country, primarily because we all had a shared 'interest' in cataloguing and metadata. This 'interest' loosely defined us as a group, albeit a group made up of overlapping smaller groups: cataloguers who were CILIP members, cataloguers who were not CILIP members, metadata specialists who were not cataloguers, academics who teach cataloguing, etc. etc. At the very least we were all 'attendees of CIG's annual conference 2012'. Because of our common interest in the topic of this conference, we had a defined 'social identity' and as individuals, defined 'role identities'; and to a lesser or greater extent we may have felt a sense of belonging to this artificially created temporary community.

I'd like to think that as members of this group we could see the potential benefits of belonging, such as being able to help one another, offering support when needed, exchanging information and ideas, networking, and improving the profession. While the community of CIG attendees may have only been temporary, we are all still members of the wider community of cataloguing and metadata specialists, with the same benefits. While initially thinking about this community I asked on Twitter if anyone could come up with a collective noun of cataloguers, and these two were my favourites:

"*A manifestation of cataloguers*" @Heather_Jardine
"*An authority of cataloguers*" @HelBader

This community potentially encompasses not only Britain, but the rest of the world; and I am sure many people have overseas contacts with whom they have networked and exchanged ideas.

Sometimes, however, it is more practical for a group to have boundaries, whether these are geographical or numerical. A world-wide community can be very positive, but the functionality of a world-wide group, less so. Thus as individuals with a shared 'social identity' we tend to band together in smaller groups, sometimes these are temporary, such as at a conference, sometime more permanent, such as being members of CILIP or members of a particular library service. We are usually members of several overlapping groups, and most of the time are not consciously categorising ourselves as such; but may unconsciously be drawing benefits from the collective groupings we are involved in. The main focus of this paper will be on cataloguers in Wales.

Cataloguing in Wales

As I believe there are benefits in having an identity as a group, I have been wondering about the creation of a group consisting specifically of cataloguers in Wales. We have a defined geographic boundary and can tap into a strong regional identity, but so far, because no group has been properly formulated, we only have an amorphous grouping without a specific social identity.

Unlike Scotland, there is currently no regional branch for the Cataloguing & Indexing Group in Wales. Some CILIP specialist groups do have regional branches in Wales, e.g. the Health Libraries Group, and the Career Development Group, so there is a precedent. Although I have no idea how many cataloguers there actually are in Wales, there are plenty of potential library services where they might be. But getting in touch with everyone and encouraging them to meet up is another problem.

I'm not going to dwell on the geographical problems we have, as I'm sure Scotland has similar issues, and probably parts of England too. Never mind that it's difficult to get to events in London, or other parts of Britain, just moving around Wales can be a challenge at times with mountain ranges dividing the country, and the fact that to get from North to South Wales by train you have to go via England. When groups, or committees, that have a cross-Wales membership want to get together they often choose a mid-Wales location to try and be fair to everyone – what this usually means in practice is, everyone goes to the Victorian Spa town of Llandrindod Wells, although relatively central it is, however, equally awkward for everyone to get to.

A conference

In March 2012, the Conversations with Cataloguers in Wales Conference was the first attempt in a long while to bring together cataloguing and metadata specialists who work in Wales. There were sessions on reclassification projects, and some specialist libraries and collections highlighted the cataloguing challenges they dealt with on a daily basis. With over 40 attendees, mostly from the South and the West, it demonstrated that there was at least an audience, or market, for a regional conference and possibly a regional group. At the end of the day there was a discussion about whether we wanted to set up an all-Wales cataloguing group of some description, and if so, whether we should be affiliated to CILIP and CIG, or remain independent. A room full of cataloguers sat in a circle, sadly resorted to type and the discussion was fairly faltering; however there was a general agreement, that yes, we did want an all-Wales cataloguing group of some sort, and maybe at some point it could be affiliated to CILIP, although it turned out that the vast majority of people in the room were not CILIP members.

An all Wales group

If we are to form an All Wales cataloguing group, these are some of the ideas that might result:

Training sessions held in Wales
Regular meetings
Bi-annual conference
Social events
Email discussion list
Blog
Exchange of ideas
Visits
Support network

This is not really much different to what CIG currently offer, so why should having a separate group be any different?

Firstly, there is the question of CILIP itself. It was apparent at our meet up that most people weren't members and had no desire to be members of CILIP. They were therefore looking for a group that didn't necessarily have any affiliations, but that stood alone. A group that wouldn't have to answer to any other organisation. The idea of some kind of affiliation to CIG was raised, which would entail committee members at least being members of CILIP. This was an option, but would garner the risk of a two-tier system, where only CILIP members could be committee members.

But if CILIP and CIG can't attract many members from the cataloguing community in Wales, would an independent group do any better? Is it the organisation itself that puts people off, or is it the financial costs?

In Cardiff I help to run a staff development group for library staff called CLIC. We are cross-sectoral, and mainly operate on a zero budget. Our events are free to attend and usually over-subscribed. Maybe we could run an All-Wales cataloguing group on a similar model.

One of the main reasons people wanted an All-Wales group seemed to be in order to organise training sessions in Wales. Too often relevant sessions and briefings and events are held in London, which often proves impractical or expensive to reach; although I would say, to be fair, that CIG are now organising training sessions at various locations around the country – and at the start of November we had the FRBR session come to Cardiff. Even within Wales there are complaints when too many things get organised in Cardiff; as in many ways Cardiff is to Wales, what London is to England; it is the capital situated in the south east, and often assumes that everything should happen there. I've already mentioned the Llandrindod Wells option, when people try to pick a middle ground; but the most sensible option is to move events around the country (as CIG are now doing). Many events do come at a cost, of course; but if it is going to be free to join the group, then maybe people won't mind paying for some of the training.

Social events and reciprocal visits are also more likely to occur when held in your own region, as once again it all comes down to travel; in addition it often seems more useful to connect with people in your own back-yard, someone who can probably relate to your day to day problems because they are based in a similar situation; and share the same 'social identity'. A regional group can offer a similar package to the national group, but just on a more localised basis.

AWHILES partnership

One positive example of an all Wales group that I would like to highlight is the All Wales Health Information and Library Extension Service, commonly known as AWHILES. (2) This is a working partnership between NHS Wales and Cardiff University health libraries which dates back to 1972. Originally the university health libraries were part of the Welsh National School of Medicine; this became the University of Wales College of Medicine, which then merged with Cardiff University in 2004. The service provides health information to support patient care, education, training and research for NHS Wales staff, and Cardiff University staff and students on placement. Originally, its prime purpose was as an interlibrary loan service between the hospital and university libraries involved, with assistance available for reference queries and literature searches. The service has expanded over the years to encompass many more aspects of professional librarianship; has regular meetings between all the librarians involved, and an annual conference. Most pertinent to this paper is the shared LMS and cataloguing support. At the 40th anniversary conference in June this year, one of the key points to be highlighted was the success of having a shared library catalogue. (3)

Currently 27 NHS libraries and three university libraries form the AWHILES partnership. Over a period of several years the AWHILES libraries were added to the University catalogue (which of course includes all the other CU libraries). For many of the NHS libraries, this entailed adding their stock item by item and having to negotiate the idea of 'cataloguing'. The majority of these libraries are very small, and have only two or three members of staff. The librarian's job encompasses all aspects of librarianship, with the focus on providing the best service for their users, many of whom are working in stressful situations. Cataloguing therefore isn't their main priority, and they are not trained cataloguers. Only one of the hospital libraries actually has staff whose main focus is on acquisitions and cataloguing; this is the largest library located at Glan Clwyd in North Wales, and their records were of a noticeably higher standard from the beginning.

Since 2004 there has always been someone within the University cataloguing department who is the 'AWHILES cataloguer'. All the AWHILES librarians were given initial Voyager cataloguing training, and any subsequent training is facilitated by this person. As this is currently part of my role, I also encourage the AWHILES staff to contact me either via phone or email whenever they have a problem. It's much easier and quicker to sort a problem when it occurs, rather than to find it by accident months later and spend ages trying to unravel what has taken place. Over the years the standard of cataloguing has improved, and even though they are under no obligation to add more than just basic details, most of the librarians take pride in creating fairly decent records. If they end up with something they deem as complicated, such as a DVD, or an unpublished local report, they know they are welcome to send the details to me and I will create the record for them. Every month we also receive two reports from our systems team listing all the new records the AWHILES libraries have added, and all their 'on order' records. As most of the libraries don't use the Voyager acquisitions module, records for their items are only added at point of receipt, rather than at point of order. The cataloguing team at the University then check and upgrade all records where necessary. As the AWHILES Cataloguer I get to attend the annual conference, and therefore know most of the librarians in person, which also helps the communication process.

To summarise, we have 27 libraries plus the University site libraries sharing one catalogue. These libraries are spread right across Wales, thus are not necessarily geographically near one another. Cataloguing of material is done at individual sites, but overall quality of the catalogue is maintained at one location. In many respects this may not be that different to a large university library service, or to a public library service, it is just that the physical distance between cataloguers is a lot greater. As a working model of a shared catalogue it is successful. Part of this success is down to the close working relationships the librarians involved have, despite their geographic separateness. Success is also due to there being one institution in overall control of the catalogue – discrepancies and basic records can be tidied up, and there is overall quality control.

Conclusion

The title of this paper is 'Are there professional benefits to exploring regionality?' however, I think perhaps what I was really asking was, how much benefit would an All-Wales cataloguing group be, versus, or in addition to, a national group like CIG? By creating a smaller, regional, group identity, I believe that more people will be willing to join in, as they will believe that the local group can do more for them than a national group. Within Wales we have plenty of examples of professional regional groups that work well. We understand the difficulties of our locale, and the best ways round them.

I have no criticism of CIG, I believe we need them too. But, I also think it's time we stood up for ourselves, and instead of complaining that we can't get to training sessions and conferences because they are on the other side of the country, we should ensure that we organise them ourselves, or encourage the trainers to come to Wales. By banding together we can create a bigger audience for such sessions; and make it less expensive for ourselves. We can perhaps get together for more informal sessions and talk about the problems we may be facing; we can also feel less alone.

This all sounds obvious, so why hasn't it been done before, why isn't there a CIG Wales? (Is that just a CILIP membership problem?) Or is it because traditionally cataloguers aren't known for being visible or active away from their desks? When I was organising the conference earlier in the year I found it quite difficult to contact cataloguers I didn't know around Wales. I used the contacts I had, I used social media, I used CIG, I used various mailing lists, and I still didn't manage to contact everyone. Cataloguers aren't listed on library web-pages, and if they are locked away in the basement it can be quite hard to get hold of them.

However, as you will have gathered I do believe that there are benefits to regional groups, and I am hopeful that we can create one in Wales. I will be getting together with interested parties in the new year to see about the best way forward. We have already managed to host the CIG FRBR event and will hopefully be able to run, or host, further cataloguing and metadata training sessions in the future. If there are any cataloguers in Wales reading this, please do get in touch and help me make the group a reality (PierceKF@Cardiff.ac.uk; @Darklecat; http://darksideofthecatalogue.wordpress.com/).

References

1. Hogg, Michael, Terry, Deborah J. & White, Katherine M. (1995) A tale of two theories: a critical comparison of identity theory with social identity theory. *Social Psychology Quarterly* 58.4: 255-269
2. http://www.walesdeanery.org/index.php/en/awhiles/891-about-awhiles.html
3. The concept of a shared LMS for *all* the Welsh higher education institutions is currently being explored by a JISC funded project. The Welsh Higher Education Libraries Forum (WHELF) are undertaking a 6 month study to develop a business case for a shared LMS across Welsh Higher Education Institutions and partners. This would include the 27 AWHILES libraries as they are already connected to Cardiff University. The project aims to summarise the key requirements and issues to consider for the implementation of a shared LMS, and the key benefits and impacts of doing so.

From UDC to DDC: reclassification at the University of Bath

Elly Cope Information Librarian (Chartered) - Acquisitions

The University of Bath took the decision to commence the reclassification of Library stock from Universal Decimal Classification to Dewey Decimal in 2009. The project has been very Academic Services driven with the Cataloguing team's involvement growing as the project has expanded.

In case any University of Bath specific terminology creeps in here's a brief glossary:

T.S. or Technical Services – Acquisitions, Document Delivery, Periodicals and Cataloguing.

The 'workroom' is the office that houses the T.S. departments. It's open plan and cataloguing and acquisitions sit in the same area.

The following are terms we use for Academic Services.

'Upstairs' and 'levels' refer to the way the library is laid out. We don't have a single sequence for the stock but rather it's grouped by subject so the subject staff sit on the same level as their stock.

Infos: although the title slide of my presentation identifies me as an information librarian T.S. Infos are very rarely referred to as 'Infos' as we all do very distinct jobs where the A.S. Infos do very similar jobs but for different subjects.

Deweying and Deweyfication are probably not unique to Bath but they are ways in which we refer to the process of reclassification.

Library hierarchy

Within the library each floor has a faculty librarian with responsibility for two subjects and for the faculty represented (e.g. Level 5 is Humanities and Social Sciences) as well as overseeing the other subject librarians on the floor.

Technical Services is headed up by the Deputy Librarian and then it splits with a bibliographic services librarian and a document delivery librarian. My line manager is the bibliographic services librarian for Acquisitions and Cataloguing. They take an overview of the project but most procedural decisions are down to me.

In cataloguing there are two-part time library assistants – one works Monday to Wednesday and the other works Thursday and Friday so there's no overlap. This did cause some problems when the procedural changes were coming thick and fast as there was no direct communication between them about how they were dealing with things.

A brief history of the project

Traditionally in the library it has always been the Academic Services staff that have done the classification work. I'm trying to get it into Technical Services as I think it sits better with us but we'll see what comes out of the reclassification project.

Rumour has it that UDC was adopted by the first University of Bath Librarian because after the Second World War it was too expensive to buy things in from America and they couldn't afford the Dewey schedules. As UDC was considered similar and was cheaper they went with that. Whether this is true or not is questionable but I like the story.

New schedules were purchased haphazardly, if at all, and the ones in use started to be amended, annotated and changed. Before the reclassification project started it was worked out that 35 different classification variants were in use in the library.

UDC schedules 1943, 1961, 1972, 1985 and some UDC 1993

History was being classified to Dewey 20

As well as these there were 28 individual in-house schedules for different areas of stock designed by members of staff to cover areas where UDC was lacking.

Unsurprisingly, this led to growing student dissatisfaction. In one month 44 enquiries at the Level 5 Information Point were related solely to UDC filing order when colons and brackets were used. Negative feedback was also starting to creep into the more official performance measures, including the annual LibQual survey. So this gave some impetus to the library to address the problem.

The other impetus was the possibility of making use of the shelf-ready services offered by Coutts and Dawson.

Task group: formation and methodology

Following the negative feedback in the LibQual survey the library Communications Task Group recommended that classification be reviewed with the aim of implementing a simpler, more intuitive system.

A task group was formed to look at the problems and work out a strategy for moving forward with the aim of having a better classification system.

They looked first at a review that the Graduate Trainee had done that year of the state of classification in the library. It was a very comprehensive report that included background, options and a questionnaire.

They also conducted a literature review looking for reports on how other libraries had handled reclassification and/or retro-conversion projects and also studies on how students actually search for and find materials. While there wasn't much around on these areas there were several ongoing projects that were identified. LIS-link was used to send questions and ask for experiences. Responses were mixed and 'success' was harder to quantify for some more than others.

They also looked at how much it would cost per book to either out-source classification or to keep it in-house. It would have been £1.75 for a fully shelf-ready book with classmark and record from Coutts. The cost/time analysis for internal cataloguing, processing and classification averaged out at 33 minutes per item and the cost would be £4.63.

Of course, the other financial consideration was how much a new scheme would cost; initial outlay and then any updates that might be published.

Task Group: identifying options

The Task Group had six interlinking areas to consider before making any decisions. Using these areas and the options each provided they would work out how best to proceed with classification and any retro-conversion:

Option A was to do nothing, maintain the status quo and try and find more staff time to keep the schedules up to date and all infos trained in all the different classification schemes in use in the library. That wasn't really a viable option – as student expectations and staff workload had been steadily increasing it was recognised that something had to change.

Option B was which classification scheme to use in place of UDC. They did consider an in-house scheme but, again, in terms of time and maintenance it wasn't really a valid option. Ease of use for both students and staff was the core driver behind the decision and as it was felt that more people are familiar with Dewey then, if the change was made, it would be to Dewey.

C was the question of whether, if they did change to another classification scheme, it would be applied to the whole library or whether they would limit it to a test area or one floor or faculty.

D was whether or not there would actually be any benefit to outsourcing the classification and cataloguing. The report suggests that they were keen to use the Coutts shelf-ready service for book processing and for their quality of records and service.

The next question was whether or not there should be any retro-conversion of existing stock. The feeling was that, in order to avoid things getting even messier over time, all stock should be retro-converted. However, practicalities suggested that small, targeted projects on identified problem areas might be the best option.

Finally, they looked at the management of retro-conversion. Would it be best to bring in an outside company to do all the stock in one concerted effort? It was recognised that with the staffing levels in place it would be very difficult to do a single hit to retro-convert everything.

Task Group: conclusions

Having considered the options the Task Group presented their preferred options. They suggested that the best thing to do would be:

Adopt Dewey across the whole library (updating to newer editions when available)

Use the Coutts shelf- ready service for new purchases

Retro-convert all existing stock in one short concerted effort.

As we don't often get exactly what we want it was the alternative options that were largely adopted through the pilot scheme that was run in 2009.

2009 pilot project

The pilot project was run in 2009 and targeted the 720s (architecture) as this was an area that had received a lot of complaints.

The aims of the projects were:

"To retro-convert architecture books within the 720 range in our UDC sequence to DDC"

"To acquire new architecture books likely to fall within the DDC 720 range in shelf-ready format."

As the majority of the Dewey numbers were going to be found through downloaded records, and not all records have numbers from the latest edition of Dewey, the schedules were checked for any major changes in the numbers. It was determined that anything from versions 19-22 would be acceptable as the numbers would likely be the same from any of those schedules.

It was calculated that there were 7,108 items in the 720 range but by using ISBNs Coutts were able to find acceptable Dewey numbers for 3,500 of those. BDZ records were also searched and numbers for a further 3,118 were found. Any books where numbers hadn't been found would have to be classified in-house. The retro-conversion would be done by the Academic Services information librarians, no out-sourcing, but there would be new shelf-ready items coming in from Coutts.

Pilot project procedures

At this point there wasn't much for Cataloguing to do. The team checked the records coming in as a quality control exercise but had very little else to do with the project.

Academic Services had the bulk of the work to do. They had a spreadsheet of titles and Dewey numbers that had been found with the version from which they'd been created. If the number was from Dewey version 19 or above then it was accepted without being checked. If it was from 18 or below the numbers were checked and then either accepted or changed to a version 22 number.

Those 1,122 items that didn't have a number were removed from the open shelves and placed in a storage area, issued on the LMS to the user ID 'RECLASS' so if anyone came looking it would be obvious where they were. If there were any items on loan, particularly where there was more than one copy of a title then a hold for 'RECLASS' was placed.

Once the numbers were on the spreadsheet Joblink staff (aka students on casual contracts) would update the bibliographic record with the new number and put a spine label on. They would also place a green dot sticker on the spine. This was done so that it would be easy to identify any scatter of stock that resulted from the project and give an indication of whether this might be a potential issue in the future.

Pilot project: results

In total the pilot took 22 weeks, which is nearly half a year. In total, 6,768 items were reclassified and some 340 were identified as missing and were therefore deleted, and 1,068 moved outside the 720 range with some moving to Level 5 to join the more sociology related subjects and some moving to the engineering subject areas. It became clear that these were mostly in anticipated areas and where the library had previously employed a high level of customisation in the classification.

Sixty-seven were ordered as shelf-ready and it was a pleasant surprise that the library had only been charged the price of 'extra items' rather than the standard price for a new item.

After completion a short questionnaire was used to collect feedback from architecture students making enquiries on Level 3. The comments were all positive and even the Head of Architecture weighed in saying that the new system was much easier to use despite having been used to the placement of books in the old classification scheme.

Expansion of the project

This is really where cataloguing became properly involved in the project. It was decided not to use any shelf-ready services, which meant that the cataloguing team were now responsible for downloading records and finding acceptable Dewey numbers.

The project was initially extended to other Level 3 subjects – the engineerings (but not literature) and all new books coming in bought by those fund codes would be classified to Dewey. At this stage, because Level 3 doesn't buy many books the increase in workload for cataloguing wasn't too bad.

The decision had been taken to continue with the retro-conversion as well. The Academic Services infos would start at the beginning of the Level 3 sequence and work through until Level 3 was complete before moving on to the next Level.

It also meant that when a new book came in any other editions or extra copies would also be captured, given the new number and re-labelled. Initially this was done by the cataloguing team and it was certainly manageable even though we were now dealing with retro-converting any extra copies and other editions – literally going to get them from the shelves, updating the records with the new Dewey number and creating new spine labels.

However, in early 2010 the faculty librarian for Level 5 requested that purchases for the Management fund code be included in the process. This would mean that new books for Management would be given a Dewey number and any extra copies or other editions would have to be captured and updated. At this point the workload on cataloguing increased massively as Management textbooks seem to have a lot of new editions. It quickly became unsustainable as there were days when the library assistants were out looking for books more than they were at their desks.

They were also getting demoralised by the sheer numbers of books still going through as UDC and knowing that at some point they'd have to see them again and Dewey them. At this point the Bibliographic Services

Librarian and I started to push for all fund codes to go to Dewey and that the other editions would either have to be left or AS would have to start doing them. To prevent other editions being at different shelf locations the faculty librarians agreed to incorporate the other editions work into the Academic Services reclassification/retro-conversion workflow.

Cataloguing kept the extra copies work but enlisted the help of the book processing assistant who would now be the one to get the books from the shelves and re-label them. This meant that we could work faster and the books would be rounded up much more efficiently so they wouldn't be on the shelves with the wrong number showing for very long.

Current cataloguing procedures

We now have the policy that any books that pass through the Workroom, including new purchases, repairs or loan category changes, will get a Dewey number and have their records updated. We did have a large number of UKMARC records still on the catalogue and although they have now officially been made into MARC21 records due to the implementation of Primo they are still very poor quality and so seeing the older books gives us a good opportunity to update these records.

We attempt to find a Dewey number in the acceptable range identified for that fund code (for example Management accept 20-22 while the sciences only accept 22 and 23). We can download records from BDZ and Library of Congress but also check the British Library, British National Bibliography and new service OCLC classify for possible numbers. If one is found either from an 'unacceptable' Dewey version or without a version number then the Cataloguing Info will check it in Dewey 23 to see if it's usable. If we can't find a number then we send it to the relevant Academic Services Info to be classified. Once they've done that it's returned to us and we finish the cataloguing process.

If the book we're dealing with is an extra copy then we will make a print out of the record before updating the call number, place a hold on any items on loan and then pass the printout to the book processing assistant who will fetch any copies from the shelves and re-label them; when the holds come in she re-labels them and files the completed sheet.

If there are any other editions in stock then we make sure the Dewey number we've used is added to all the other records in an 082 field and we make printouts of all these records. We then send them to the relevant Academic Services info. This is because the call number on the OPAC hasn't been changed so there isn't such a pressing need to re-label them and the stock is more easily fetchable by subject staff on the levels.

Our Systems Librarian has created a programme which he called 'ReclassificationThing'. When a book's barcode is scanned it pulls the number from the 082 field through into the call number field and it creates a suffix using the hierarchy of 100s, 700s, 245 and then creates a label using this information. This has sped the process up considerably and now Levels 3 and 5 have their own labelling machines. Each floor also has its own Reclass user ID in order to place holds for any items out on loan.

Retro-conversion

The retro-conversion of Level 3 continues all year round with the Academic Services infos timetabled to do half a day a week on it. This involves fetching a trolley of books, identifying Dewey numbers for them and re-labelling. The Task Group are hoping to be able to identify funding for extra help with the re-labelling and the other editions work as this has started to reduce the number of books being reclassified.

Each summer since the start of reclassification there has been a targeted summer project on an identified problem area. It started with the 720s and since then Mechanical Engineering and Computer Science have been completed. Due to the implementation of Primo there wasn't one this year but likely contenders for 2013 are Biology & Biochemistry and Management. I would be happy with either of these as they are very textbook intensive and if we could get them all done then the reclassification workload at busy purchasing times would be significantly reduced.

The future

We're unlikely to adopt shelf-ready services in the near future – the systems we have in place are working well and the trial with Coutts didn't impress sufficiently to encourage the outlay on it. The summer retro-conversion projects will continue, as will the work on converting Level 3.

My hope is that as the library gets more Deweyfied there will be more of an impetus to change the stock layout to a single sequence. While there is some merit in having the books grouped by faculty and subject it's probably more confusing than helpful and if the purpose of making the change to Dewey was to make it easier for users to locate stock then we should probably make it as logical as possible. This may be a long term goal, however, as there are some who are very wedded to the faculty system.

There will of course be changes to Dewey. We have adopted Dewey 23 and we were lucky this time in that there were only enhancements to numbers in our main subject areas while some other areas had considerable changes. We may not be so lucky with future editions and it may be a problem that we have to deal with when Dewey 24 is published.

Over the three years I've been involved with the project I've managed to establish my role and the work of the Cataloguing team as central to the project. It has evolved so that the main spearheads of the day-to-day workings of reclassification are the Cataloguing Info and the Information Librarian on Level 3 (where it all started).

However, the original Level 3 Info left last year and so there has been a new member of staff to bring up to speed. The benefit for me was that an internal candidate got the job and I already had a strong working relationship established with her.

We also have a very recently appointed new Head Librarian at the University of Bath and she is very keen on the Dewey project. We have high hopes for a new collection management policy that will enable easier withdrawal of outdated stock. This would mean that we could clear some of the UDC stock from the statistics and never have to Dewey it.

The main thing is to keep going. It does seem like a Herculean task and one that probably won't be finished during my time at the University. Current calculations predict that, at the current rate, we have another 17 years to go before the library is totally Deweyed. However, I believe that the project is improving things, not least the standard of the catalogue records. The procedures are now embedded and are no longer subject to change on a frequent basis so things may speed up again. The reasons for embarking on the project remain valid and while ever we can keep sight of that we'll keep soldiering on and try to keep in mind the improvements it has already made and how good it'll be when it is finished.

From Club Reading Room to Research Library: developing cataloguing and classification practices in a specialist sports library

Neil Robinson **Research Officer, Marylebone Cricket Club**

Visitors to the MCC Library at Lord's will probably find that it still looks very much like the reading room of a private club: the two leather chesterfields and coffee table for relaxed, sociable reading, the walls lined with oak shelving and the two tables with limited study space do not suggest a library where much serious research takes place. But while the Library still serves as a convivial, and dry, location for Members of MCC to congregate on match days, on non-match days research use is growing rapidly.

Although MCC began formally collecting cricket memorabilia as early as 1864, there was no real attempt to create a proper library until the opening of the new Pavilion in 1890. In 1893 the Club accepted a donation of a full set of *The Badminton Library of Sports and Games* from Captain HB Sutherland. This is usually taken as being the starting point for the Library collection, but the previous year the Club had turned down the donation of an entire library collection "for the time being", while in 1884 the minutes record the purchase of bookcases for the Members' Room. Some of the mid-19th century books in the Collection were clearly presented to MCC at the time of publication. It is hardly a picture of consistent decision making, and this rather random method of assembling a library continued well into the 20th century.

The first home of the Library was in a small, square room to the left of the Pavilion entrance. From 1893 to 1973 the Club's growing collection was housed in this cramped space which offered limited access to external researchers (i.e. non-Members). Over the course of those 80 years, more and more of the Collection had to be removed to a slightly damp basement and by the end of this period there was room for no more than two seats for visitors to sit and read or research. The new Library was no more a purpose built facility than the old; it was developed out of a series of disused bathrooms on the top floor above a converted rackets court which was now the Imperial Memorial Gallery (later the MCC Museum). The location offered fewer restrictions to access for external researchers, but less obvious access – and, crucially, reduced visibility – for Members. But the space was an improvement and with the extension of 1985 doubling the available shelf-space, there was room for the collection to expand for the next thirty years.

So much for facilities, but until 2006, no professional, qualified Librarian had ever been appointed to manage the Collection. So what were working practices like during the formative years of the Library?

For so much of its history MCC was the realm of the gentleman amateur. Founded in 1787, not until 1864 did it appoint its first professional Secretary. The first full-time groundsman appeared around the same time. The appointment of the first Curator of Collections did not take place until 1946 when, in the aftermath of the Second World War, the Club decided to open its Collections to the public for the first time by creating the Imperial Memorial Gallery in a disused rackets court behind the Pavilion. It was fortunate for the Library that the move to create a public museum came at this time. The acquisition of two major donations – the library of AL Ford in 1930 and that of Sir Julien Cahn in 1945 – had radically altered the character of the Library Collection, making it for the first time a collection of historical importance.

Planning for the new Museum and all associated collections issues was placed in the hands of the newly formed Arts & Library sub-committee. Responsibility for day-to-day management was given to a remarkable woman: Diana Rait-Kerr. Given that this was the realm of the gentleman amateur, it would be easy to scoff that Diana's principal qualification for the job was that she was the Club Secretary's daughter. True enough, she had no formal training or qualifications in Museum or Library work but she was a woman of remarkable gifts who quickly became a respected authority on cricket history and literature. Keen to do her job as thoroughly as possible, Diana read a book on librarianship and invented her own classification system. Rait Kerr 1, as we call it, shows the clear influence of Dewey:

Rait Kerr 1: the first classification scheme

- **1. History of Cricket**
- 10 General History
- 11 Early History – 1725
- 12 History 1725-1850
- 12.2 Scores and Press Extracts
- 13 History from 1846
- 14 History of Lord's & MCC
- 14.23 MCC Club Periodicals (not scores)
- 151 Reminiscences of Cricketers and Others born before 1870
- 152 Reminiscences of Cricketers Born since 1870
- 153 Reminiscences of Club

- **2. Records, Scores, Biography**
- 20 General statistical and reference works
- 201 Curiosities of Cricket
- 202 Test Cricket
- 2021 Test Cricket: England v Australia
- 2022 One Day Internationals
- 203 Statistics of Cricket in England
- 2031 Gentlemen v Players
- 204 Statistics of Overseas Cricket
- 21 Biographies of Individuals and families
- 22 Collected Biographies

The scheme comprises nine classes, with subdivisions, and uses decimals or additional digits to create further divisions or represent subject content from elsewhere in the schedules. It was simple and easy to understand, but sixty years after its introduction there had been no major update and, apart from the introduction of a handful of additional numbers, no attempt had been made to account for the radical changes in the game since the late 1940s, the birth of new nations like Pakistan and Bangladesh or the growth in publications which address the game from a social-historical point of view. There were also some inconsistencies in the scheme. It was remarkably Anglo-centric: the only subdivision under 202 – Test Cricket was 2021 – Test Cricket: England v Australia, leaving all the other Test playing countries to be lumped in together. There was also a curious separation of autobiography and biography into different classes, obliging any library user searching for all the biographical information on a particular cricketer to look in at least two separate parts of the Library.

Perhaps the clearest sign that any classification scheme is overdue for an update is when one of its sections becomes a dumping ground for items difficult to classify. With Rait Kerr 1 that section was 13 – Modern History from 1846. But unlike the *Number 13* in the famous story by MR James, our number 13 did not disappear in daylight; it stayed there, two bays of shelving, throbbing with malevolence. Any member of staff faced with the task of finding a book shelved in 13 would approach the task with understandable reluctance. It was practically impossible to find anything there in under five minutes, while any time a new title had to be classified in 13 it felt like a little defeat. It was clear that number 13 had to go and Rait Kerr as a whole had to be updated.

For anyone approaching a re-classification there are probably two essential questions to address:

- Was the original scheme well conceived?
- Does the scheme need to make philosophical sense or simply make finding items on the shelves easy?

The answers to these questions will help to inform the decision whether to simply update the existing scheme or to devise a new scheme entirely from scratch. There is also the question of managing expectations. No classification scheme can ever be perfect, and the detailed representation of subject content is particularly difficult to achieve in a cricket library. So much of cricket literature is made up of books which are part autobiography, part history, part instruction manual, part guide to the Laws and part essay on cricket's current state and future development. From the outset of the reclassification it was crucial to remember that such books would remain difficult to classify. They should, however, be considerably easier to find on the shelves.

From a personal point of view there was the additional factor of inexperience. I was asked to take on this project soon after starting at Lord's, with no more than two years' professional experience under my belt. Consequently I felt that I did not yet have either the professional capacity or the knowledge of the collection to tackle with confidence the design of a new scheme from scratch. So, the choice was made to adapt the existing scheme which, in its basic conception, was sound enough. This made it easy enough to tackle directly many of the inconsistencies in Rait Kerr 1.

The process of designing the update took around one year to complete. A vital part of the process was talking to other specialist libraries about how their classification schemes handle certain issues, and our colleagues at the Melbourne Cricket Club Library were especially helpful. It was important not to be afraid of borrowing ideas from elsewhere. For example, the Melbourne Library's approach to books on cricket tours was shamelessly appropriated, while the Test cricket section was expanded by borrowing the treatment of this subject in *Wisden's Cricketers' Almanack*, i.e. dealing with each Test playing country in order of its accession to Test status and subdividing by contest between those countries. An easier issue was the reconciliation of biographical material into 150. The result, Rait Kerr 2, was a scheme with a vastly increased number of classmarks available for use, making the process of classifying new titles easier for Library staff, while for Library users, the process of browsing was made far more simple and efficient. The dreaded number 13, now numbers 130-135, was reduced from two bays of shelving, to two and a half shelves. Browsing was also aided by the addition, for the first time, of spine labels to all open-shelf stock.

Some examples

- **100** **History (General)**
- **101** **Social & Cultural History**
- 102 Amateurism and Professionalism
- **110** **Early History to 1725**
- **120** **History from 1726 to 1863**
- 121 Hambledon
- 122 All England XIs
- **130** **Modern History from 1864**
- 130.1 1864-1914
- 130.2 1915-1962
- 130.3 1963-
- 130.4 The Origin of the Ashes
- **131** **Scandals & Controversies**
- **132** **Contemporary Opinion**
- **133** **Broadcasting**
- 133.1 Radio
- 133.2 Television & Newsreel
- 133.3 Internet

- **200** **Statistics & Records (General)**
- **201** **First Class Cricket (General)**
- **202** **Statistics of Cricket in Great Britain**
- **203** **Statistics of Early Cricket to 1863**
- **204** **Statistics of Cricket Overseas**
- **210** **Test Cricket (General)**
- **211** **Test Cricket (England)**
- 211.2 England v Australia
- 211.3 England v South Africa
- 211.4 England v West Indies
- 211.5 England v New Zealand
- 211.6 England v India
- 211.7 England v Pakistan
- 211.8 England v Sri Lanka
- 211.9 England v Zimbabwe
- 211.91 England v Bangladesh
- **212** **Test Cricket (Australia)**
- 212.3 Australia v South Africa

Implementation presented its own challenges. It was decided to commence the operation to coincide with a redecoration of the Library in early 2008. All the books were crated up and stored in an empty gallery on the top floor of the Museum. Once redecoration was complete, the crates were returned one by one and the titles manually reclassified before reshelving. Some bulk changes, where the numbers of entire sections were changed without additional subdivision, were carried out remotely before reshelving began, but it was important still to handle each title in turn and assess whether a more relevant number had become available.

The whole process was carried out in one month of seven-day working for myself and one keen volunteer. A side-effect of this, apart from severe exhaustion, was a much improved knowledge of the collection acquired from handling each title individually. This led, inevitably, to the realization that some of the changes in the update might have been managed differently. But this is an inevitable part of any reclassification process and shows the necessity for continuous updating to keep any scheme fresh and workable. Perhaps more importantly, for a solo librarian working in a specialist collection, the vastly improved knowledge of that collection which resulted was a significant benefit in a job which requires a mixture of librarianship and historical awareness.

Having completed the reclassification project, the focus of the department turned to cataloguing. Once again, as in the late 1940s, the Library was able to benefit from an initiative aimed principally at the Museum. Up to 2008, no catalogue of MCC's Museum and Archive Collection had ever existed, apart from some printed catalogues of the Club's fine art Collection. With the approaching Bicentenary of the current Lord's ground coming in 2014, the year which would also mark the 150th anniversary of the Club's collections, the decision was taken to acquire a new Collections Management System as a shared platform for all three elements of the Collections with a view to creating an online catalogue in time for 2014.

The Library was able to lead the way in this project, since it was the only element of the Collections already to have been catalogued in any meaningful sense. The first computerised MCC Library catalogue was introduced in 1989, replacing the printed card catalogue which remained the only way users could access information about the Collection. Computerisation was the brainchild of Douglas Foskett, then a leading figure both in the Library Association and on MCC's Arts & Library sub-committee. The catalogue itself was created on an Access database; this had limitations of its own, but it also imported one major problem from the card catalogue. There was no link between the printed catalogue card and any individual copy.

For any historic collection, knowledge of provenance is vital. In our case, we had very little knowledge and then only thanks to any physical inscription which might link a copy to information on the card. There had also been no attempt to conduct an inventory of the number of copies we had of any individual title. This was not addressed when moving the catalogue on to Access, since the cataloguing method used was simply to migrate data directly from card to database. Consequently, in many cases we knew we had several copies of a given work, but not precisely how many, where they all were or where most of them had come from.

The Access catalogue was limited in several ways: there was no facility for distinguishing between edition level and copy level information, so when multiple copies of a title were held this was either represented by creating duplicate edition records, or by adding the note "another copy" to the record. Sometimes both of these methods were used; sometimes neither. Access also lacked the facility for the creation of authority files, so when an author's name was entered it might be recorded exactly as printed, as in a statement of responsibility field, or in the form of the author's full name. There was nothing to prevent the creation of a variety of different versions of a single author's name, all of which would later be imported as authority files to our new system in 2008. An identical problem existed with subject terms; the random, uncontrolled entry of such terms had led to a total lack of consistency and a remarkable excess of terms stored in the system.

This of course made searching for information on the catalogue problematic, but there were technical issues to do with Access which also played their part. The most significant of these was our database's inability to search for anything containing an apostrophe. Given the presence of an apostrophe in "Lord's" this was a major inconvenience to say the least.

The first stage of planning for the cataloguing project involved setting realistic goals. We had a deadline of five years before the Bicentenary, but at this stage we had no cataloguing staff and no new system to catalogue onto. The Library was by far the most well-catalogued element of the Collections, so it made sense that this would form the backbone of the online catalogue, with "highlights" of the Museum and Archive joining it. A three -man working party was put together to select a suitable collections management system – and it was a great help that one of those involved was Steve Jones, head of IT at MCC, who was actively interested in the project having formerly worked at the National Gallery and co-written one of the early versions of the MODES system for museums.

The choice of Adlib as our new software was largely due to its modular design: separate library, museum and archive databases, each allowing documentation according to the relevant professional standards, could be bolted together to form a single platform for a diverse collection. But purchasing new software is an expensive business, and having a budget for this meant compromising on the budget for staff.

At the outset of a major cataloguing project, there was no budget to recruit extra professional staff, leaving volunteers as the only viable option. Working with volunteers can be a rewarding experience, but they come with varying degrees of training and skills and it can be difficult to acquire the necessary aptitudes without paying for them. The amount of time each volunteer spends on a project is also worth bearing in mind: if a volunteer spends just a short time on a project then the project itself fails to benefit from the knowledge the volunteer has gained and continuity can be adversely affected; if a volunteer spends too long on a project then the position they occupy ought really to be a salaried one and the relationship is veering uncomfortably close to exploitation.

The MCC cataloguing project was fortunate in that one of the volunteers recruited happened to be a professional librarian from India, with extensive cataloguing experience, who happened to be temporarily resident in the UK while her husband worked here. She stayed with us for around ten months and got through an enormous amount of work with great accuracy and consistency.

Before cataloguing could begin it was essential to reorganize the Collection so that all items would have a fixed location, avoiding the time-consuming necessity of having to re-enter location information later in the process. An external storage facility was closed down and hundreds of duplicate items, rarely consulted annuals, match programmes and periodicals brought back on site. They were shelved in steel cabinets along a corridor adjacent to the Library, dubbed “cupboard world” or “the corridor of uncertainty” by staff, with each shelf numbered. Items were grouped together according to the priority order in which they would be dealt with. The following categories were used:

- Individual editions – i.e. books which are not annual publications.
- Rare books – these were stored separately in our archive facility.
- Annuals – i.e. books such as *Wisden* or club and county yearbooks.
- Match and tour programmes.
- Periodicals.

Individual editions were tackled first since they made up the bulk of the stock on open shelves.

Work began with the first cabinet of duplicate individual editions. A book would be selected from the duplicates and its record on the system identified. The master copy in the main library would also be identified and an attempt made to link either copy with the information already on the catalogue. If this could be done, a unified edition record would be retained, with all copy information held in the two copy records within it. Each copy record was given a number which corresponded with a number written in the copy itself; this would either be the original accession number showing when the book was acquired (e.g. L.68.17) or in the absence of this a pencilled temporary number allocated in sequential order (e.g. TL.2009.510). Any duplicate edition-level record would be deleted, assuming that the information it contained could be firmly tied to one of the copies identified.

Once this was completed, both copies would be returned to their shelves, their spines marked with a small red sticker to show that the copy had a record on the system. This allowed staff to physically mark their progress through the cabinets and, once the cabinets were completed, to spot gaps in the main library where unduplicated items remained to be dealt with. This marking process, dubbed “measling” by staff, was simple but remarkably effective. The other method of tracking progress was statistical, as new copy records were created with L. or TL. prefixes, old ones with imported MCC prefixes were over-written or deleted. Using Adlib’s copies database it was possible to assess how many copies with each prefix existed and thus calculate the percentage of records which had been dealt with.

As a result of this work, the new MCC Library catalogue contains far more usable information than the old Access database. In cases where multiple copies of a single title are still held, there is a single edition-level record containing several copy records, each bearing a number and a description which links it definitely to a copy in a fixed location. At the time of going to press, only the periodicals remain to be catalogued, and these will be catalogued at title level rather than recording individual copies. When the data was migrated from Access to Adlib, 11,595 records – a mixture of edition and copy level information – were imported. As the project nears completion, more than 19,000 copies have been identified.

All of this work has been done manually, with cataloguers handling the books, reading enough of them to be able to enter accurate subject terms and entering data at their own workstations. This remains very much the method at the MCC Library; there is no use of MARC tags, no importing of bibliographic data from external sources, no shelf-ready stock. The method is old-fashioned, hands-on, dust in the fingernails, artisan cataloguing. Leaving the cataloguing project aside for a moment, the MCC Library acquires around 250-300 new titles per year, excluding periodicals. At this level of stock acquisition the cataloguing workload is light enough to be manageable for a solo librarian, and there is therefore little economic sense in paying to import records from elsewhere. Furthermore, manual cataloguing by staff with subject knowledge will generate far more comprehensive and accurate catalogue records than can be acquired from any external source; such detail is vital for any library wishing to be taken seriously as a centre for specialist research.

Periodicals are expected to be fully catalogued by January 2013. In the spring of next year, the MCC catalogue will go online, first to a limited audience of MCC Members and staff, then to the wider public. Recent experience with the online project *Taking the Field* (www.takingthefield.com), plus the impact of the Library's own greater online presence and the bi-annual publication of *MCC Magazine* has demonstrated that outreach projects can generate much greater interest in special collections and drive access to them. The awareness these initiatives have generated has had a remarkable effect on the number of research visits to the MCC Library. From a figure below 100 during 2010, by the end of November 2012 the number of visits for the year had reached 286, a vast improvement rendered all the more notable by the fact that as well as being closed to researchers on match days, the Library was also unavailable to researchers for the whole of July while Lord's was involved as host of the London 2012 Archery competition.

As well as attracting more visitors to the Library, it is anticipated that more enquiries will also be generated. Currently the Library receives about 850 per year. While receiving additional enquiries might place a greater burden on staff, it is hoped that the presence of the online catalogue and a selection of frequently asked questions on the Library website will help to weed out the more basic enquiries, particularly those relating to the presence of certain items in the Collection.

But the work needn't end there. Work is already underway aimed at unifying cataloguing standards and creating an agreed list of subject headings with our colleagues in Melbourne. It is hoped that this work will lead eventually to the creation of a comprehensive online bibliography of printed material on cricket, which would be the key resource for researchers requiring information on cricket literature and would increase further access to the Library Collection. Whether this happens or not, the creation of such standards would be a useful end in itself and help to cement what is already a productive partnership between the only two professionally run libraries in a very specialized sector.

The journey from Club reading room to research library has been an existential one. As the Collection grew too large to be housed in the Pavilion and was moved to another location, access for MCC Members was made more difficult. The improved working practices highlighted here have helped greatly to attract new audiences to the Library and new interest in its Collection. If the MCC Library is not to be a widely-used facility for Members and does not make money for the Club, attracting new audiences from outside the Club and acting as a showcase for MCC's commitment to the history and heritage of cricket becomes its primary justification for existence and helps to guarantee that its Collection will be available to anyone with an interest in cricket for years to come.

Harvesting repository (EPrints) records for use in an LMS (Aleph)

Katrina Clifford **Senior Information Advisor, Kingston University**

As with many things, making a small decision can be the thin end of the wedge (in a good way!) in terms of both creating work for yourself and making improvements. That was the way with this project but I think the overall benefits have been worth it.

In early 2011 the decision was made to include details of all the higher level research theses (PhD, MPhil and Masters by Research) in the Kingston University Research Repository (http://eprints.kingston.ac.uk). It was just the metadata as we've not yet gone down the electronic theses route. The repository records were to include elements such as supervisor's name and funder, elements not listed in the catalogue records. This meant each thesis (about 600) was checked by hand and the details entered into a spreadsheet. This spreadsheet was then uploaded to the repository by EPrints as part of a paid upgrade.

The process of going through the theses led us to thinking about the whole process of adding theses to stock. It became apparent that there was a certain degree of double-handling going on which meant theses were taking some time to be available. From participating in the British Library's EThOS scheme, we knew that getting the theses available quickly was important as they were often requested very soon after being added to EThOS. We already knew that the catalogue records were not as rich as those in the Repository as these were often catalogued quickly by a non-cataloguing member of staff. We began to wonder if, instead of detailed records being added to the repository then being added separately to the catalogue, we could take the information from the repository and import it directly into the library management system.

By this point we had moved to a new library management system (Aleph) in summer 2011. Looking in the manual it showed where text-based records could be imported but very little information was available on what format the text had to be in in order to display properly. I did some hunting around on the internet – at times it felt like I was looking for a needle in a haystack! Finally I found a document from the University of Liege in Belgium (http://orbi.ulg.ac.be/handle/2268/494) which explained how to enter the bibliographic information in a spreadsheet in order to output it in a text file that was suitable for importing.

Meanwhile, our technical analyst was looking for a programming project to work on in conjunction with EPrints so we asked if there was a way to automatically fill the spreadsheet from the Repository. This meant I had to explain which parts of the record I wanted and in what format I wanted the text to appear in the spreadsheet. You can choose which elements to export from the Repository; for example, we don't take the 'Research Area' as it's not a Library of Congress subject heading. In the records on EThOS, however, 'Research Area' is incorporated as 'Keywords'.

A colleague of mine explained how to interpret the XML and how to write what I wanted it to look like so that a computer could read it in terms of spacing etc. In the end it worked out better than we'd hoped as a Perl script was written to output from the Repository straight to a Word file ready to convert to text for importing. (1) So now, once a search has been done in the repository, the results can be exported into Word and imported into Aleph via a text file.We've chosen not to go back and 'upgrade' the old theses, although we may consider scanning in abstracts of them all as and when time allows.

Of course, there are often different ways of doing the same thing. Since setting up the export I have found out I could have used MARCEdit to carry out this work. At the time I had considered this may be a possibility but I couldn't find the information and we had the support for creating the export through EPrints.

As with many projects, at times it felt like I'd opened a bit of a can of worms with this! When we chose to add a link to the EThOS record in the Repository we found that some of our theses were not listed on EThOS. Many of our early theses were awarded by the Council for National Academic Awards (we're a former polytechnic) and some of these were incorrectly assigned to other institutions. Sara Gould and Heather Rosie at EThOS have been brilliant at working with us to correct these errors. Furthermore the process of adding new theses to

EThOS has been greatly streamlined. Before this, title pages and abstracts were photocopied and sent by post but now the records can be exported directly to EThOS from the repository (they use the date stamp to ‘pick up’ records added/updated).

Overall though, despite the work involved in getting the project to this stage, there have been a lot of plus points. I’ve got a better understanding of the workings of the repository. Also, I thought I was pretty good at being able to explain cataloguing but this really did test it, not only in explaining it to other people but how to write things in a way that a computer will understand and output what you want and I’ve learnt a lot. The main bonus, however, has been one that will benefit many people in that the catalogue (and EThOS) now more comprehensively covers what we have, with greater findability.

The slides of my talk, along with further information on creating a spreadsheet and text file can be found at http://eprints.kingston.ac.uk/23511/ (on our institutional repository). I’m happy for people to contact me to find out more out what we did if they are interested. (k.clifford@kingston.ac.uk)

Reference

1. Information on writing an API can be found at http://wiki.eprints.org/w/How_to_write_plugins)

Media cataloguing practice at Sheffield Hallam University

Helen Garner

This lightning round talk looked at developments in media cataloguing practices at Sheffield Hallam University.

About us

Sheffield Hallam University is a large institution (over 30K students) on two sites. We have three cataloguers plus two information assistants who work on media cataloguing. We use the Millennium LMS and have the Summon discovery tool. We export our catalogue data into Summon, known locally as "Library search" and no longer promote the standard Opac. Media material is used in all areas of teaching and we have a number of specific media collections such as Feature films and Computer games

Size of the media collection

The media collection is quite large - nearly 44,000 items, 15,000 of which are feature films. We began recording off-air material in streaming video format in 2008 and now have over 8,000 titles. Annual cataloguing statistics indicate that we catalogued a total of 4,658 media titles in 2011/12, with 1,481 being in streaming format.

The volume of material we are required to catalogue can be a challenge. I will discuss our strategies for dealing with this.

Cataloguing strategies

As we purchase a significant amount of feature film material, we subscribe to the BDZAV service which is an additional subscription to the standard book service. We search for material using the EAN (barcode number) and can obtain detailed records for Region 2 films and TV entertainment material. In the last two years we have downloaded 900 records. This might not sound a lot but as the records are extremely detailed with plot summary, cast list, extra features, accessibility notes etc they represent quite a time saving. Only minor edits are required. We currently remove all the 700 fields for the cast. Unfortunately, not all material we buy is covered. For example, we often buy Box sets and then catalogue films individually and also purchase a significant amount of Region 1 material. This needs to be catalogued in-house. I understand that BDS will go live with print RDA records in January 2013, but we will have to wait longer for BDZAV records to be created in RDA format.

We also purchase video content direct from publishers like TV Choice, Classroom Video, Viewtech Educational media and Concord media. We sometimes find records on BDZ AV but often this content doesn't have an EAN and records can't be found. We therefore catalogue this material in-house. We now try to obtain a licence from publishers so that we can make content available as streaming video, e.g. all TV Choice titles are available in streaming format, off-campus and world-wide.

Off-air developments

We began using the Tripleplay IPTV system in 2008 to record TV and radio onto a server. We then took out the ERA plus licence in 2011 so we could make content available off- campus. We have been looking towards streamlining processes and workflows and in 2012 we stopped producing a DVD version of the recording. We have now moved towards streaming only (with a few exceptions). This will represent a saving in staff time, materials and space.

Dave Pattern, keynote speaker, talked to us about Library Impact data. We also value usage data and are now able to get information from Tripleplay about total hits and unique users for the overall portal and the constituent elements. The reports cover the most frequently watched titles and titles which have been available over three years and not viewed. We have known for a while that our DVD circulation figures are down and can now evaluate usage in Tripleplay. This data could also be used to inform programme selection strategies.

Media catalogue records in a Discovery system

When our Library catalogue data went live in Library Search in May 2011 we had a number of data issues specific to media records. Firstly, we had a lot of in-house catalogue records which only had the fixed field language code in Millennium. These were mostly foreign language feature films and were appearing in English language facet in Summon. This was clearly a problem, so we updated about 4000 titles with 008/041 fields. We achieved this clean-up before many staff noticed! When we mapped the content types we used in Millennium to the available options in Summon, we were unhappy that there was no streaming video or audio content type. We had offered a similar facet to users when we had the Encore search. We canvased the Summon user group to gain support for these content types and once we gained backing from other customers it was set up. We also pushed to have the "online content" refinement expanded to include streaming content. This gave the content type a yellow online sunburst, which is quite visual to users.

At Hallam we have a history of using a range of local GMDs for non-standard material, e.g. VoD, Computer game, Blu-ray DVD and many more! It was not possible to display the 245 |h in Summon. We had concerns that users would miss them, especially as search terms. We therefore decided to map 245 |h to a searchable 500 note field. In reality, I think that the content facets and icons have adequately replaced the GMDs and this should help with the transition to RDA.

However, I am aware that we will need to ensure that our RDA records display sensibly in both the standard WebOpac and Summon. I feel that we will need to upgrade older media records to have RDA features at some stage.

Shelf-ready at the University of Sheffield: speeding things up without dumbing things down

Emily Bogie — **Metadata Specialist, The University of Sheffield**

In 2005, book purchasing at the University of Sheffield increased dramatically in preparation for the opening of the Information Commons, a new learning and research centre. This increase in book purchasing coincided with a restructuring of what was then known as Bibliographic Services and the retirement of a number of metadata specialists. In order to avoid a backlog of items waiting to be catalogued and classified, the decision was taken to move towards the purchase of shelf-ready items.

The introduction of shelf-ready items into the collection was a gradual process. Over a period of three years, shelf-ready items were trialled and the procedures associated with these items finely tuned. Initially, we received small samples of items which were shelf-ready in terms of processing. These items did not come with bibliographic records, nor had they been classified. At this stage, the Head of Bibliographic Services checked each item in order to ensure that they had been processed correctly. As this stage was successful, fully shelf-ready items were then trialled; those which had been processed, catalogued, and classified. Initially, spot checks were made by acquisitions staff in order to ensure that the correct bibliographic record was attached to the correct item and that the classification number on the spine label matched the classification number in the item record. Due to the high volume of items being received in preparation for the opening of the Information Commons, however, acquisitions staff were unable to continue to carry out these checks, and the items began to go straight to the shelves after receipting.

By 2008, the volume of items being ordered had decreased and it was decided that it was feasible to begin checking the quality of the bibliographic records attached to shelf-ready items in order to ensure the integrity of the library database. Metadata specialists began to check a list of imported bibliographic records for shelf-ready items on a daily basis. During this process, a number of issues were identified such as records with essential fields missing, e-book records attached to print items and incorrect classification numbers. When metadata specialists identified a problem with a bibliographic record, a message was added to the order record in order to alert acquisitions staff and to intercept the item before it reached the shelf. The physical item was then seen by the metadata specialist and the bibliographic record amended or overwritten. In light of the issues that this process raised, the daily checking of imported records became a permanent procedure within the metadata team. This decision was deemed to be of particular importance in light of the imminent move to a new resource discovery layer with the ability to make use of more of the information contained within a MARC record.

The shelf-ready process begins with order sorting. A number of item types are excluded from the full shelf-ready process - music scores, special collections items, CDs and DVDs. These items are ordered from a non shelf-ready supplier or, from Dawsons or Coutts with ‘no shelf-ready’ instructions. Acquisitions staff are able to request different levels of shelf-ready, from basic processing to full shelf-ready, which includes processing, cataloguing and classification. As part of the shelf-ready import, a file of fulfilment data is sent to the Talis gateway which is then consolidated into a single file and transferred to our server. Every day a script runs to process the fulfilment data which, in turn, updates the item records with barcodes and classification numbers. The MARC record import works by taking all standard numbers, such as an ISBN, from the existing record, which has been imported during the ordering stage, and searching for a matching record in the following databases: British Library Consolidated, Library of Congress Books ALL, BDS Books, RLUK, WorldCat, Dawsons, and Coutts. The first matching record found is then imported. The order of the databases searched was amended a number of times during the shelf-ready trial period in order to increase the probability of a high quality record being selected.

When the items are received, acquisitions staff complete a number of brief checks, such as ensuring that the item matches the order and that the barcode is correct. Items are forwarded to the metadata team if required. If no further action is required, the item status is changed to ‘in stock’ and the item is forwarded to the correct library site to be shelved.

Each morning the metadata specialists receive an email which lists all of the imported bibliographic records for that day. These records are then checked one by one on Talis in order to assess their quality and to make any necessary edits. A number of criteria are addressed during this process. The metadata specialist first checks for any essential fields which may be missing from the record such as authority fields and subject fields. If these are not present, a message is added to the order record so that the physical item can be seen before it reaches the shelf. Unnecessary fields are then deleted in order to avoid data cluttering, for example, the 263, 300 (if the pagination details are missing), 521 and 852 fields. The accuracy of data within fields is assessed, such as ensuring that any dates are consistent and that the names in the authority fields are in the correct form. It is also necessary to remove any data within the record which might be misleading for library users, for example, any mention of electronic access in a print record, such as an e-book ISBN or an 856 field. Occasionally, a record is assessed to be inadequate in terms of quality and so is overwritten with a record from an external database.

The classification number is verified as a follow-up task, as the main priority is to get the items to the shelves quickly and to make them easy to locate on the catalogue. The alteration of classification numbers for shelf-ready items comes under the remit of a separate reclassification project.

As previously mentioned, a metadata specialist will request to see an item if the bibliographic record that has been imported is not of an adequate standard. At the start of 2012, the metadata team began to keep detailed statistics regarding shelf-ready items. From January to August 2012, of 1,936 shelf-ready bibliographic records imported, metadata specialists requested to see 385 items. This is just under 20 per cent of all of the records imported and represents a manageable amount of items for the metadata team to physically handle. It also demonstrates that checking records that are imported as part of the shelf-ready process is a worthwhile procedure.

In conclusion, the shelf-ready procedures that we have in place allow us to speed up the process of making items available to users, whilst continuing to maintain the quality of our bibliographic data. In terms of the future of shelf-ready at the University of Sheffield, it is expected that the processes involved will change following the move to a cloud based library management system in June 2013, although at this stage it is not known to what extent.

Quality Control of Shelfready at the University of Warwick

Christina Claridge **Metadata Librarian**

This trial explored the possibilities of introducing shelfready processes to the University of Warwick library.

The processes for bookstock at Warwick University Library are particularly complex. They include all three divisions of the library, passing through four different teams before arriving on the shelves. Book requests originate with either the Reading List team or the Academic Services teams, move through Acquisitions and Cataloguing [Data Services] before arriving at the Collection Management team.

This paper will focus mainly on the work done in the Data Services team.

The shelfready trial ran from October 2011 to July 2012, and covered 743 books (803 by mid-August). As the library holds over 90,000 volumes, this is a very small percentage of the total collection, and covered limited subjects: Medicine and sciences; part of philosophy; and Law (using Moys classification).

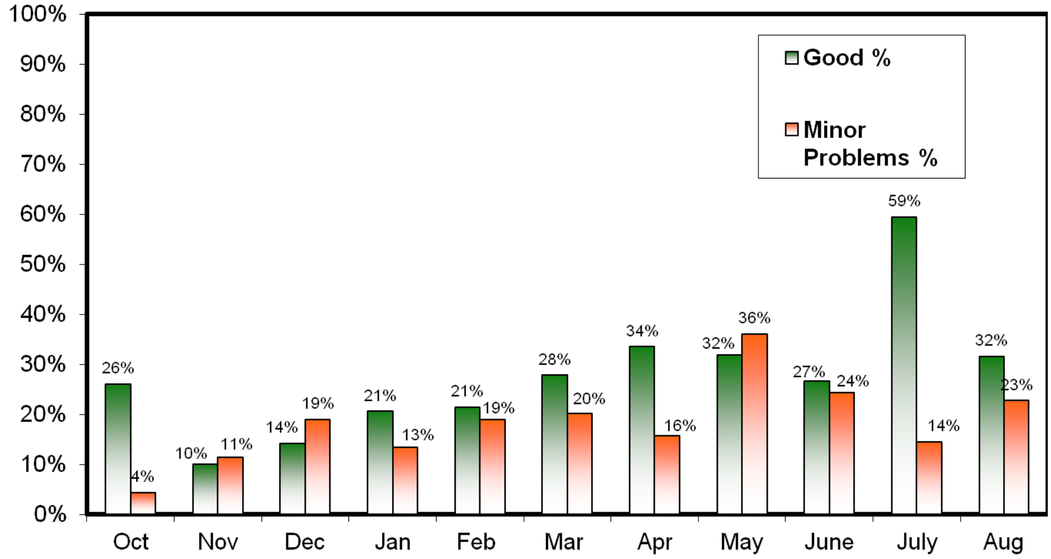
During the trial, every physical book was checked against the bibliographic records, and every bibliographic record was printed out, and changes to the record were noted. Changes were then transcribed into a spreadsheet, so that patterns could be tracked through the trial. Books were then sent to Collections Management to be made available and shelved or relabelled if necessary.

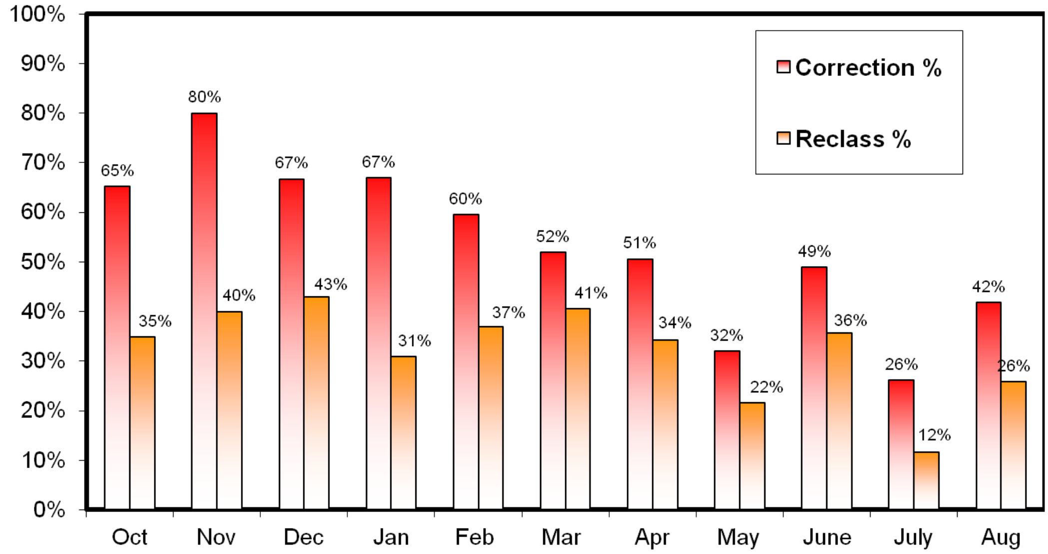
The criteria used for checking consisted of a series of questions:

- Was the classmark correct?
- Does the record need significant corrections?
    - Correct title (including series) and authors.
    - Correct information about the book (numerical data).
    - Correct subject headings and MESH heading if possible.
- Are there any other issues? (856, 504, 505, 440 fields)
- Processing problems: label over edition or volume no. or title.
- Item problems: Warwick requires multiple locations and status within individual bibliographic records.

The problems with introducing shelfready covered both internal and external issues. Internally, the structure of the library teams worked against shelfready books. The feedback system was slow and difficult, and there were delays in seeing results as there was sometimes a backlog of books to check. There were also problems with the addition of item records to bibliographic records, a process which was done within the library. There were also some early teething problems that resulted in entire batches of records having the same load problems. This was eventually solved by introducing an entirely new load table, only for shelfready books, into the library management system.

There were also external issues. Records were of lower quality, especially when compared with earlier editions, and books needed to be reclassified. There were also problems with the quality of the processing; labels were placed over the edition, author or title information on the spine.





As of September 2012, this is no longer a trial. The physical books are no longer being checked against the records; only the records are being checked by a cataloguer. Errors in records are no longer being collected, and the cataloguer checking the records must now fetch the book from the shelf or recall the book from the borrower if there is a problem with the shelfready record. The question of what level of error will we and our partners accept has been answered.

A first-timer's view

Stephen Carlton

Having been awarded one of two sponsored places at this year's CIG conference, I was especially determined to get as much as I could from attending. To this end, I composed a brief list of things that I wanted to achieve while at the conference. The list included vague things like:

- Get a better understanding of RDA
- Identify gaps in my knowledge
- Talk to some interesting people
- Live tweet the conference

I've been to a couple of conferences since qualifying last year, and both were fairly terrifying. Being in a room with loads of infinitely better qualified peers can be intimidating at the best of times. However, in order to be able to tick off all the objectives on my list, this was something I was going to have to deal with. Thankfully, recognising people from Twitter made this a little less scary than it might have been.

The theme of the two day conference was 'the value of cataloguing'. In the current climate of cuts and upheaval within the information profession, particularly in the cataloguing community, the value of the work we do is often overlooked. The conference was an opportunity for us to celebrate our achievements, share best practice and look to the future.

I'm not even going to attempt to write about all of the sessions over the two days, suffice to say that they were all incredibly inspiring in their own different ways. I am going to talk a little about a few sessions that I enjoyed the most, the sessions that I feel I took most away from and the sessions that made me feel proud to be a cataloguer.

While studying for my Librarianship degree, I must have found a way to cite Dave Pattern's work at the University of Huddersfield in almost every essay I wrote, irrespective of whether it was relevant to my argument or not. It was quite exciting for me, then, when it was announced that he was going to be doing the keynote session at this year's conference. The controversially titled 'If you want to get laid...go to college' (It's a Frank Zappa quote, apparently) brought us up to speed on the JISC funded Library Impact Data Project (or LIDP from now on), which the University of Huddersfield is a lead contributor to.

It seems almost obvious to us information professionals that library services are extremely beneficial to library users. There are countless ways in which our libraries help to make people's lives better. We have all probably heard stories about the positive impact that libraries have made. It is, however, nice to be able to see that laid out in pretty charts and graphs. Which is exactly what Dave Pattern gave us. Also, pictures of cats.

The main aim of the LIDP is to analyse library usage data in order to better demonstrate the value of university library services. One of the key ways in which Dave and his team have done this is by attempting to find a link between library usage and the grades that students achieve. There's been some fascinating results so far and I would definitely recommend having a look at the project's blog if you get the chance.

One graph, for instance, seemed to show that the time of day that students access e-resources has an effect on the degree classification they go on to achieve. First class degree students, for example, tend to access e-resources earlier than their peers. Third class students, however, seem to login later in the day and continue using e-resources into the early hours of the morning. This might suggest that these students often find that they have to work late in order to finish assignments at the last minute. These students also seem to have an intriguing dip in usage at about 8pm, which Dave amusingly attributes to the 'Eastenders Effect'!

In return for being able to attend the conference, I had agreed to live tweet the proceedings for the benefit of people who could not be there. This started reasonably well but half way through the first day I had completely lost the ability to simultaneously tweet and listen. There must be a knack to it and I certainly don't have it. In the end I managed to send a couple of tweets per session, but it wasn't the flowing stream of wisdom that I'd hoped it would be. One useful tweet I did send, though, was nicked from Simon Barron's ace presentation 'The fundamental interconnectedness of things'. He said that 'cataloguers are the cartographers of the abstract' and I thought it was so profound that it needed tweeting immediately. So I did.

He said plenty of good stuff, actually, about how we're moving away from strict hierarchies towards the idea of knowledge as a network. Because everything is connected. I guess this is what FRBR is for, in that it helps to describe relationships and links between all kinds of things. Not just books and other books, but also manuscripts and objects and images and audio (and lots of other things) too. It got me thinking about linked data, which I know very little about, and it encouraged me to write the words 'read up on linked data!' in red ink in my notepad.

Aside from the presentations themselves, the Q&A sessions after each talk were particularly illuminating. It was interesting to hear people discussing and sharing ideas with each other, about RDA especially. This was most definitely the hot topic over the two days, and the presentations covering the implementation of RDA and the subsequent discussions gave me a real insight into how other organisations are preparing for it (or not preparing for it in some cases).

Céline Carty's update from the ALA conference contained a really useful piece of advice: that it's good not to be a trailblazer! Let other organisations adopt RDA early and learn from their experiences. I'm pretty sure that we won't be adopting RDA when the British Library and Library of Congress implement it in March 2013, but this statement made me feel a lot better about it. Also, Céline provided some excellent links to RDA training materials that you can access for free on the web and I will definitely be using these at some point soon.

Cataloguers are encountering difficult times at the moment, and there were a few sessions on how we can better demonstrate our value. The amazing High Visibility Cataloguing (Or HVCats) initiative, which is well worth following via their blog and on Twitter, was mentioned several times over the two days. Céline Carty, a member of the HVCats team, popped back to the podium to give us an update on recent developments and talk about the Cat23 CPD project which is in the pipeline. It looks like there's going to be loads of useful resources available through the site in the near future, so keep your eyes peeled for that.

Between sessions, the opportunity to talk to different people was fascinating and terrifying in equal measures. I spoke to people from loads of different sectors, some influential Twitter types and, less impressively, colleagues from within my own organisation. This was one of my favourite parts of the two days and gave me some much needed networking practice. Practice that I badly needed! Apologies to anyone reading this who was forced to listen to me ramble on about braille and giant print.

I think, overall, I managed to put at least a partial tick next to all the criteria on my checklist. There's loads of sessions that I haven't mentioned here that helped me to clarify things in my head, gave me ideas for things I want to get involved with in the future and generally made me proud to be a member of the cataloguing community. I left Sheffield on the Tuesday feeling like a saturated sponge, almost overwhelmed by how much information there was to try and absorb. In a good way. I'm now looking forward to the next CIG conference and maybe, by the time that comes around, I'll have something to present myself?

Book review

Indexing: from thesauri to the Semantic Web / Pierre de Keyser.

Oxford: Chandos Publishing, 2012.

xxi, 249 p. ISBN: 9781843342922

Reviewed by **Neil Nicholson, National Library of Scotland**

This book is the result of a course the author (head librarian of the Katholieke Hogeschool Leuven, an institute for higher education in Louvain, Belgium) taught at a library school.

Having read so much on RDA recently, I was hoping this book would be a welcome diversion - and so it turned out. The author has provided a fascinating insight in to new trends in indexing without what he calls "the pure technical aspects".

The book is well laid out, in 12 chapters, each preceded by a useful abstract and key words. The first chapter gives a useful overview of traditional (library) indexing tools of subject indexing and thesauri. This chapter also deals with some practical aspects of controlled vocabularies – where they can be found and how they can be created or maintained. Subject headings, such as Library of Congress Subject Headings, are also studied.

The second chapter gives an overview of the arguments used in the discussion between the supporters of manual indexing and those of automatic indexing. Arguments against manual indexing are that it is slow, expensive, not detailed enough, that it does not lead to better retrieval and that there is no consistency between indexers. Arguments against automatic indexing are that it does not provide an overview of the index terms, does not solve the problem of synonyms and variants, does not take the context into account, does not allow browsing related terms, and that it is too complex for computers. Significantly, the author notes that the results of automatic indexing are of lesser quality than those of manual indexing.

Chapter 3 deals with automatic indexing of text material, which has been applied for many years in full-text databases, repositories, and search engines but, as the author notes, some of the applications have been disappointing, with simple word-by-word indexing, many errors and no advanced techniques.

Subsequent chapters deal with automatic indexing of images, moving images, and music. Taxonomies and ontologies are also examined and the author concludes: "ontologies and taxonomies can be very good instruments for structuring data, and as a consequence, also for searching and comparing data. Up until now, many applications have been built that use one or more aspects of ontologies and taxonomies, but their full power will be released when the Semantic Web is realized".

The content in the chapter on metadata formats and indexing will be most familiar to cataloguers as it gives an overview of what the author considers the most important metadata formats, such as Dublin Core, MARC21 and ONIX, and the way they treat subject indexing.

The final chapter focuses on the Semantic Web, looking at XML and XML schema, the Resource Description Framework (RDF), and the Simple Knowledge Organization System (SKOS). The author believes that SKOS "guarantees that the vast knowledge hidden in the traditional instruments libraries used for indexing, i.e. subject headings, thesauri and classifications, can contribute to the realization of the Semantic Web", an optimistic note to end on.

In summing up, I thoroughly enjoyed reading this book and found it well researched, with plenty of references to further reading, well written and very informative.

Catalogue & Index is electronically published by the Cataloguing and Indexing Group of the Chartered Institute of Library and Information Professionals (CILIP) (Charity No. 313014)

Subscription rates: free to members of CIG; GBP 15.00 for non-members & institutions

Advertising rates: GBP 70.00 full-page; GBP 40.00 half-page. Prices quoted without VAT.

Submissions: In the first instance, please contact the Editors: Cathy Broad (library@ethicalsoc.org.uk) and Heather Jardine (heatherjardine402@hotmail.com)

For book reviews, please contact the Book Reviews Editor: Neil T. Nicholson, Cataloguing & Metadata Services Team Leader, National Library of Scotland, e: n.nicholson@nls.uk

ISSN 0008-7629

CIG website: http://www.cilip.org.uk/specialinterestgroups/bysubject/cataloguingindexing

CIG blog: http://communities.cilip.org.uk/blogs/catalogueandindex/default.aspx

Tags from the CIG blog on specific areas of interest:

authority control book reviews Catalogue and Index cataloguing CIG activities CIGS classification committees and working groups conferences Dewey digitised material Dublin Core events folksonomies linkblog MARC metadata news RDA Semantic Web social software standards taxonomies UDC