

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 |app. 34. Pl. A-E. Edward Stanford: London, 1858.

300 |afol.

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 *infects*

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.