Harmonisation of MARC and Descriptive Cataloguing Standards

by STUART EDE

INTRODUCTION

This paper is intended to be an overview of the current activity towards harmonisation of catalogue record exchange across the Atlantic. LIBER has had a special interest in harmonisation since it passed a resolution to explore harmonisation with the North American library community at its Annual Conference in Prague in July 1999.

The paper will concentrate on current developments MARC harmonisation, because this is where most effort is currently focused, and because format harmonisation is also the logical place to start in the process. However, format change cannot be carried out in isolation, and the data carried within the formats must be harmonised as well to ensure full interoperability of record exchange.

MARC DEVELOPMENT HISTORY

MARC was invented by Henriette Avram and her team at the Library of Congress in the late 1960s. Since then it has been adopted country by country until it is the predominant format for exchanging cataloguing data between libraries. Unfortunately many later adopters felt they could improve upon the original design, or they had special cultural factors that had to be reflected in the format. The result is that national variants were developed. Within the English speaking community alone there were the British UKMARC and Canadian CAN/MARC formats, plus for a time the Australian and South African variants, in addition to USMARC. These variants impeded easy record exchange, and conversion programmes were required before records from one country could be used in another. For a long time libraries had to rely on their national libraries or bibliographic utilities to carry out bulk conversions of monthly exchange tapes which would then be distributed in bulk or selectively.

UNIMARC was developed under the auspices of the International Federation of Library Associations (IFLA) as a solution to this conversion problem, and for use as a national format by countries who were yet to adopt MARC. The idea was that UNIMARC would be a common format into which and from which records could be converted. This would minimise the number of permutations of conversion routines needed. UNIMARC has fulfilled a very useful role, but it has found itself in competition with USMARC (now MARC21), because so many countries have found it easier to adopt MARC21 in order to reap the benefits of joining cataloguing utilities like OCLC and to be able to use library automation packages developed for the North American market.

In practice UNIMARC has been adopted only by a number of countries which came later to MARC cataloguing and which tend to be smaller in terms of publishing output. Even then they have sometimes made local amendments to reflect cultural requirements. For a long time UNIMARC was not greatly used for conversion. This changed with the directive that all European Union funded library projects should use UNIMARC for the exchange of records, and the most prominent uses of the format are in the Consortium of European Research Libraries and in the European Register of Microform Masters (EROMM).

BARRIERS TO EXCHANGE

While one can point to considerable successes in promoting bibliographic record exchange across national boundaries, the differences in national formats have created a number of barriers to the exchange, which have meant the exchange has not been the free and easy peer-to-peer traffic that might have been hoped for. Instead in most cases national libraries or bibliographic utilities have had to act as intermediaries.

The need for intermediation has been a barrier in itself, especially where national libraries are unable to provide record distribution services.

The conversion process can be very complex. Where the structures of the formats are substantially different, e.g. where there is a one-to-many equivalence, sophisticated programming rules have to be devised – often with multiple condition algorithms to be applied. Even then the process is often imperfect, resulting in quality issues or loss of content and structure. For many years - and it is still the case in many countries today - the conversions had to be carried out in bulk by central agencies. In the past decade there have been efforts to develop software that could be used by individual libraries, preferably for on-

the-fly conversion while online to other libraries. The technically successful European Union funded UseMARCon project[1] is a prominent example. However, its success in promoting peer-to-peer conversion has not been as widespread as had been hoped, particularly because the exploitation of the results by software houses and systems vendors has been disappointing.

FACTORS FOR HARMONISATION

An obvious driver for harmonisation is the easier exchange of data, overcoming the barriers identified above. But what are the other drivers?

Reducing systems costs is a significant factor. Because of the size of their internal market, the North American library systems vendors are strong and at the forefront of development. By comparison the choice of European systems is more limited. Widening competition helps to reduce costs, but even then European libraries buying North American systems have found they pay a significant premium for the systems to have the capability to handle non-MARC21 formats.

The cost of maintaining separate national formats cannot be ignored either. While the direct costs of MARC development offices fall mainly on national libraries, the costs to the communities and professional organisations of participating in format revision is not inconsiderable. The accelerating pace of format change to accommodate digital publishing and other trends has increased these pressures. There are attractions, therefore, to be part of a wider community where the costs are shared more widely.

Perhaps the largest single driver is the globalisation of publishing. Because of the economic strength of the United States and its predominance in research output, the multinational publishing houses – even where they are of European origin – have a strong North American influence. This is reflected in the dominance of English as the language for research communication and in the preponderance of books and serials published in or for the North American marketplace. So European research institutions find themselves buying a high proportion of English language publications, for which MARC21 records are readily available. Making the flow of information to researchers as easy as possible is a powerful incentive to libraries to adopt or harmonise with MARC21.

HARMONISATION OF UKMARC, CAN/MARC AND USMARC

These drivers have been at work in the English-speaking world for some time. In the 1980s Australia stopped using AUSMARC and adopted USMARC. More recently South Africa has dropped SAMARC in favour of MARC21.

As the then Director of the National Bibliographic Service of the British Library I opened a dialogue with the Library of Congress in 1993. My main aim was the simplification of cataloguing. By initiating parallel simplification processes in the UK and North America I hoped the formats could be made to converge towards a single, harmonised format. This struck a chord with the then Head of Cataloguing at LC, Sarah Thomas, who had been thinking along similar lines. However, when we consulted the Library of Congress MARC Development & Standards Office the clear message came back that, having just gone through the process of amalgamating the multiple formats for different material types, the US community would not have the stomach for another radical upheaval so soon. An alternative tactic was therefore devised: to harmonise formats first and thereby form a common platform from which to pursue simplification. At this point the National Library of Canada was brought in to the discussions, because it made sense to move forward together.

While the USMARC and CAN/MARC formats were very closely related with just a few differences stemming from Canada's bilingual tradition, the differences between USMARC and UKMARC were - and are - considerable. Chief amongst these is that the subfield structure of UKMARC was designed to mirror the International Standard Book Description (ISBD). The opportunity was also taken to generate ISBD punctuation by software algorithm rather than requiring cataloguers to input the punctuation manually, as in USMARC. A further difference is the way the two formats treat multilevel works. UK practice is closer to continental European tradition. The other major difference is the use of the 006 and 007 encoded fields in USMARC/MARC21. However, this was not seen as a barrier to harmonisation, because their adoption in the UK would enhance the facility to catalogue publications in digital and other non-print media.

Consultation with the UK library and information community in 1995 showed that the majority view was in favour of harmonisation. However, respondents urged the British Library to persuade the North Americans to adopt ISBD subfield encoding and computer generated punctuation at output. ISBD compliance was seen as important to maintain compatibility with UNIMARC. Many libraries were also afraid of the impact on productivity of having to input punctuation. With regard to multilevel works the general opinion was that

a compromise might be worked out incorporating the best from both UK and North American practices.

Unfortunately, an impact assessment in the US revealed a reluctance by bibliographic utilities and systems vendors to undertake the large scale database conversions required, which they did not feel was warranted by the benefits of a harmonised format.

So in 1997 an alternative approach of progressive convergence was worked out. A joint MARC Harmonisation Coordinating Committee was formed, and each of the national libraries agreed to keep future developments in step so as to prevent any further divergence. The partners also undertook to explore the application of technological change together, e.g. the emergence of XML (eXtensible Markup Language) metadata formats as possible successors to MARC. The British Library began a process of progressively adopting USMARC fields, starting with unique fields that would have little impact on users and working towards the more complex changes that would require systems or database changes by users. The intention was to consult the community at each stage and stop when users felt the process had gone as far as they were prepared to go.

Meanwhile Canada and the US harmonised their formats. The harmonised CAN/USMARC format was called MARC21, and fairly soon it was promoted as an international format in virtual competition to UNIMARC. It is unfortunate that should be the case, but it is probably a reflection of the economic realities in the marketplace.

In 1999 the British Library consulted the Book Industry Communication Bibliographic Standards Technical Working Group, which acts as the advisory body for UKMARC development, on the next phase of harmonisation. They advised that the community might prefer a „big bang" approach over the staged method of change, because in that way all the systems changes could be made at one time. At the request of the British Library, the Working Group drew up a detailed proposal for what should be contained in the converged format, and this would form the basis of a consultation document. It was recognised that the situation had changed in other ways since the last major consultation in 1995; for instance, there was a growing trend by academic libraries to procure MARC21 based integrated library systems, and collaboration with North American bibliographic utilities had increased. Therefore the opportunity was taken to ask again the fundamental question whether the community wished to keep UKMARC largely unchanged, go for a converged format to the extent proposed or adopt MARC21 *in toto*.

349

The result of the major consultation exercise in Autumn 2000[2,3] was a clear mandate (57%) for adopting MARC21. There was still a large minority (30%) who wished to retain the unique features of UKMARC and who therefore voted in favour of the partially converged format, but only 7% wanted to keep UKMARC largely unchanged.

MARC OR XML?

An issue that the British Library raised as part of the consultation exercise was whether the community wished to wait until it became clear whether a metadata standard being developed for e-commerce or for describing digital documents, such as ONIX (ONline Information eXchange) or Dublin Core, might offer a viable alternative to MARC. However, the feedback from the consultation open meetings was that this was unlikely to happen very quickly, and it was in any case another reason for switching to MARC21. MARC21 would provide a common platform from which to move together towards the new format, when the wider availability of support services and conversion tools would make the transition easier.

LIBER

While the debate had been going on in the UK, LIBER had, of course, set its own strategy. By resolution at the Annual Conference in Prague in July 1999 the LIBER Secretariat was given the authority to explore harmonisation of MARC formats and descriptive cataloguing codes with the North American community. From experience gained on the REUSE project collaboration between German libraries, OCLC and the Library of Congress, and from preparatory work carried out by Monnika Münnich for LIBER, the most contentious issues were identified as the treatment of multilevel works in MARC21 and divergent authority control practices in Europe and the AACR community. Interestingly LIBER members viewed ISBD subfield encoding, which was the predominant issue in the UK, as of secondary importance.

After some informal soundings through the British Library the President of LIBER made a formal approach to the Library of Congress, who hold the chair of the Committee, early in 2001. This was particularly timely, because it coincided with the reactivation of discussions in the MARC Harmonisation Coordinating Committee, and the Director of the Deutsche Bibliothek was invited to attend the MHCC meeting in May as LIBER representative. This

meeting was exploratory, and it was agreed that follow-up meetings would be held over the succeeding months.

GOVERNANCE

Besides the technical issues under discussion, governance of the format was on the Committee's agenda. In giving the British Library the mandate to move towards MARC21 the UK library community made the stipulation that the UK should be closely involved in the governance of MARC21. LIBER is also interested in ensuring that governance reflects the interests of international users.

From discussions at the meeting in May 2001 it emerged that the Library of Congress wishes to remain the steward of MARC21. This is unlikely to be contentious in itself, provided other countries are given an appropriate voice. The role of other institutions in decision making and the nature of the advisory structure were identified as areas needing further discussion. At present the MARBI committee, convened under the auspices of the American Library Association, is the sole advisory committee. The National Library of Canada has a representative on MARBI and is consulted in advance on proposals going to the committee. However, there are no representatives from other MARC21 countries on MARBI. The future governance options include enhancing international representation on MARBI or setting up other regional forums, eg a European advisory committee.

The meeting agreed to adopt a stepwise approach, and the first step is likely to be the inclusion of the British Library in the existing agreement between the Library of Congress and the National Library of Canada. Proposals for more far reaching changes to the governance are to be drawn up for discussion at the Autumn 2001 meeting. It will be interesting to see how those discussions turn out.

DESCRIPTIVE CATALOGUING

The harmonisation issues raised by LIBER extend beyond format issues in to the area of cataloguing codes, e.g. name authority practices. These go beyond the remit of the MARC Harmonisation Coordinating Committee and fall within the scope of the Committee of Principals for the Anglo-American Cataloguing Rules. This will, therefore, require a separate dialogue, though the overlap in membership between the two committees, i.e. the national libraries

of the UK, Canada and US, should make it easier to coordinate the discussions.

The International Conference on the Principles and Future Development of AACR held in Toronto in 1997[4] recommended that the internationalisation of AACR should be explored. There are two aspects to internationalisation: governance and internationalisation of the code itself. The Committee of Principals decided that internationalisation of the code should take a lower priority for the time being, because the restructuring of the code to comply with IFLA's Functional Requirements of Bibliographic Records[5] would go a long way towards the internationalisation goal, and was in any case a necessary precursor. At the same time the harmonisation of serials cataloguing between AACR, ISBD(S) and UNESCO's International Serials Data System will also help better alignment of cataloguing practice[6]. Another reason for postponing work on internationalisation of the code was the heavy workload borne by the Joint Steering Committee for the Revision of AACR, who are charged with formulating the restructuring programme.

However, internationalisation of the governance of the code can be treated separately, and the Committee of Principals commissioned the British Library to draw up a discussion document for consideration at its May 2001 meeting. This is just the beginning of the process, but the issue is now firmly on the AACR agenda.

NEXT STEPS

This is an exciting time for international collaboration, but because of the far reaching nature of the changes required it will take some time to realise the objective of an environment where libraries across Europe, North America and further afield can more readily exchange catalogue records one with another across language and cultural barriers. The process has to be taken a step at a time, but the key players are now engaged in dialogue, and the only way is forward.

In summary the next steps are:

- Review of the governance structure for MARC21

- Exploration of the format differences and development of proposals, especially working out a compromise on the treatment of multilevel works

- Review of the governance structure for AACR

- Restructuring of AACR

- Investigation of the technical implications of internationalisation of the cataloguing rules.

I hope that by the next LIBER Annual Conference there will be some solid progress to report.

REFERENCES

1 <http://www.bl.uk/services/bsds/nbs/usemarcon/usemarcon.htm>.

2 All change for UK library cataloguing format – 19 March 2001, <http://www.bl.uk/pr2001/14.htm>.

3 <http://www.bl.uk/services/bsds/nbs/marc21survey.htm>.

4 The principles and future of AACR : proceedings of the International Conference on the Principles and Future Development of AACR : Toronto, Ontario, Canada, October 23/25, 1997 / edited by Jean Wiehs. — Ottawa : Canadian Library Association ; London : Library Association Publishing ; Chicago : American Library Association, 1998. — ISBN 0-88802-287-5 (ALA); 1-85604-303-7 (LA)

5 Functional requirements for bibliographic records: final report. UBCIM Publications - New Series Vol 19. IFLA Study Group on the Functional Requirements for Bibliographic Records (September 1997). K G Saur, München 1998. <http://www.ifla.org/VII/s13/frbr/frbr.htm>.

6 <http://www.nlc-bnc.ca/jsc/current>.