FULL MARCs for MANUSCRIPTS

by David C Sutton*

British and American archivists and manuscripts cataloguers have long held the view that both Chapter 4 of *AACR2* and the standard MARC format are so inadequate for archival purposes as to be not worth taking seriously and not even worth teaching on archives courses. This may be regretted, since MARC is undoubtedly the key to national and international data exchange, but archivists have always refused to consider reducing their standards and, specifically, curtailing their descriptions, simply in order to fit into Procrustean systems devised by librarians for librarians. (Such systems significantly place archives and manuscripts

* Paper presented at the Annual Residential Seminar of the Cataloguing and Indexing Group, University of Kent at Canterbury, July 1989. David C Sutton is Senior Research Officer of the Location Register Project, University of Reading

under the highly offensive heading of "non-book materials".

This paper describes three initiatives which have to some extent broken this impasse and which, by introducing new versions of the MARC format into the world of archives and manuscripts, make data exchange and international networking possible. These are the Location Register Project, based at the University of Reading; the new cataloguing system of the British Library Department of Manuscripts; and the RLIN-AMC format, which has transformed manuscripts cataloguing in the USA over the past six or seven years.

Unfortunately, the assumption that MARC is wholly unsuitable for the cataloguing of manuscripts is still, in Britain, taken as read and taken as a starting-point. In recent years, computerised cataloguing has been widely adopted by British archivists, but it has taken the form of a

CIG SILVER

CIG what? In 1991 CIG will celebrate its 25th birthday. Events are moving into place for the year. Obviously, we will be "under one Umbrella" in Leeds at the Groups' Conference. As an indication of our all pervasive nature we are holding three joint sessions with other groups - Audio Visual, Information Technology and International and Comparative. But will you bring a birthday cake with you to Leeds for Friday night?

Negotiations have reached an advanced stage for a sponsored T-shirt (100% cotton, of course - and unbleached if it can be sorted out with the printers). This will be available in the new year; perhaps an ideal late Christmas present. The T-shirt will be printed in two colours with a 4" diameter CIG logo and "SILVER" written below it (the sponsors' logos will appear on the verso). The price will be in the region of £5. Please contact Keith Trickey, 25 Sandy Lane, West Kirby, Wirral, L48 3HY ('phone 051 625 6316) specifying size required to reserve your copy.

CONTENTS

proliferation of individual stand-alone catalogues in a variety of formats. Some of the catalogues are very sophisticated in design, but none of them was created with data-exchange in view.

Certain archivists have moved beyond their technical objections to MARC-based systems to develop positions in favour of stand-alone systems as a matter of principle - arguing that for archivists there is no need at all for physical transfer of data from one catalogue to another or from one network to another. Such a position was strongly argued by Chris Woolgar in an article published in the *Journal of the Society of Archivists*:

> The main principle behind the development of the Machine Readable Cataloguing (MARC) format for the automated cataloguing of books was the need for libraries to exchange data, to develop shared cataloguing systems. It is much less likely that record offices will wish to share data, although they will probably wish to search the data in each other's files. The principal question of strategy, therefore, in the development of an automated system for cataloguing manuscripts is not the standardisation of the detailed structure of information interchange, but the means of access to it.[1]

Many British Archivists, and a few North Americans, still hold the view that the spreading use of systems with incompatible software is not a problem and that archivists need only to be able to access each other's data, not to share it. In other words, the aim is a series of linked nodes, without any development of host databases.

In mitigation of this individualist position, let me remind you of an obvious fundamental difference between archives and books: namely, the fact that each archive is unique. Until the age of word-processors, there has been no such thing as genuinely duplicate manuscripts. And when two manuscripts have been very very similar, they have *especially* needed individual descriptions (including references to their siblings). You cannot catalogue a manuscript by waiting until the University of Bristol has processed its copy and then using their record!

Nevertheless, what Woolgar has not taken into account is the fact that the searcher for manuscript information, as opposed to the cataloguer, can derive benefits from data-sharing networks which are comparable to the benefits derived by the searcher for book information. The searcher who is obliged to scan through the series of linked nodes, using a system such as JANET, for example, is being given a very poor searching service. Thus, someone who is looking for all papers relating to the fourteenth century poll tax, say, or to little magazines published in Paris, is soon going to tire of logging in and out, in and out, in order to search the catalogues of Canterbury Cathedral and then perhaps the Huntington Library in California and then Gloucestershire Record Office and so on.

This is not to denigrate JANET or similar networks. JANET is a very useful tool in many ways, but it is not designed for cross-collection searching. It is, moreover, becoming ever more obvious that the greater the variety of catalogues available on JANET, the more complicated JANET becomes for its users.

I suspect that the arguments now coming forward against networked archival databases are, to some extent, *ex post facto* self-justifications. Having devised a system which makes data-exchange almost impossible, one will naturally be much more likely to evolve a philosophy based on non-exchange.

We at the Location Register have adopted a different philosophy, based on the desirability of sharing data. We have been using a MARC-based system for the listing of manuscripts since 1982. We have since been followed in this by the North American archivists using the RLIN-AMC system and, more recently, by the British Library's Department of Manuscripts and we are quite clear that our choice of a MARC-based system has brought us, and will bring us, a number of important advantages - advantages which will outweigh the restrictions which the use of MARC has placed upon us.

When I took up my position as Senior Research Officer of the Location Register project in October 1982, I was starting from scratch. I had a blank desk, an empty filing-cabinet, no staff, and a computerisation-budget but no computer. The mission was to go out and locate and list all modern literary papers which were available for public consultation anywhere in the British Isles.[2]. The project's founders (my Advisory Group) had taken a decision that this new location list should be "computerised", but had not prescribed a method. The first task of the new Senior Research Officer was to be the choice of a system of computerisation. We can take it, however, that most of the librarians on the Advisory Group were assuming that UKMARC was the norm, while the archivists were generally hostile to it.

I started by dividing the possibilities into three groups:
(1) some sort of stand-alone system based at Reading, using our own programming according to our own precise needs, probably with a fairly large mini-computer;
(2) the use of the facilities of a university computer centre like Oxford, Manchester or, indeed, Reading, in conjunction with a standard archival or database package such as FAMULUS (in 1982 surveys showed that more archivists were using FAMULUS than any other package or format);
(3) the use of an existing database or network to create and store our data (recognising that this would involve our becoming the first manuscripts catalogue to use the MARC format).

There followed two months of discussions, tests, proposals and counter-proposals[3] - at the end of which time a decision was taken that the Location Register should be a MARC file maintained by the British Library's BLAISE-LOCAS service. Among the factors which were taken into account in reaching this decision were the desire to see a major national project as part of the major national database; the impracticality of maintaining our own hardware; costing, and standard of output; and the allocation of a number of local MARC fields to fit the particular needs of a literary location list. I don't want to burden you with lists of sub-fields and indicators, but examples of our local fields are 719 for location (city, town or village); 659 for genre (novels, poems, libretti, account books, etc.); 956 for printed sources of information; 957 for the date information was collected; and 249 for correspondence (distinct from

245 for manuscripts). The distinction between 245 and 249 would enable us to file literary manuscripts alphabetically, but correspondence in a chronological sequence.

So we became a UKMARC file, and we developed a cataloguing style which fitted in with MARC's requirements. And, with our extensive use of local fields, we have found ourselves only rarely inconvenienced by the field-restrictions of the MARC format. It is important to stress, however, that literary manuscripts are a special category, and in fact they do fit rather well into the librarianly prejudice that I described earlier: seeing manuscripts as "raw books". Literary manuscripts, almost alone in the world of archives, *are* raw books. Cataloguing them as items, in the same way that one catalogues a book, as opposed to cataloguing them as collections, can be justified.

This has made our use of MARC much less of a problem. When our choice of cataloguing system was being announced, in a variety of talks and seminars in 1983 and then in a brief article in *Program* in 1984[4], I always stressed that we were not claiming to have devised *the* MARC system for manuscripts: only to have devised one particular MARC-based manuscripts catalogue.

⋆   ⋆   ⋆   ⋆   ⋆

In 1983 I went around the lecture and conference circuit of British librarianship and archives studies, telling people that as far as I knew the Location Register was the only scheme using MARC for the cataloguing of manuscripts. What I didn't know at that time (and I think hardly anyone else in Britain knew either) was that in the United States a group of archivists was developing a MARC scheme on a much bigger scale - a scheme which I hope, in fact, might one day swallow up the Location Register. This was the RLIN-AMC project.

AMC is an amplification of USMARC. Its designers faced the same general skepticism from archivists as we faced when we first decided to use a MARC-based system. They were, moreover, attempting to cover a much wider range than we were - they were designing a system for *all* the archival holdings of Harvard, Yale, Cornell, Stanford, Princeton etc., with the aim of persuading local (state and county) as well as university archivists to come in and join their scheme. They have produced a format with a number of important extensions and additions to USMARC, refinements which it is worth our considering one by one:

First, they removed MARC's restriction on added entries. An RLIN-AMC record for a collection of correspondence might contain several hundred name entries for correspondents in 700 fields. This theoretically simple but technically difficult change has given them great advantages.

Second, in similar vein, they introduced a new field, 520, which is a full-title or full-description field. The 520 field has no keystroke limits - or, if all fields must always have a limit, this one is very large indeed. In several AMC records that I have seen, the 520 goes on for four or five screens. It is in fact a complete replacement for the cherished "archival description".

Third, they introduced a facility (through a 773 field,

referred to as "host item") which enables groups of papers to be linked in a hierarchical tree. Thus the entry in the catalogue of the Fales Library, New York University, for W S Gilbert *Libretti* refers upwards to a more general entry for the W S Gilbert Collection, and the entry for the W S Gilbert Collection in turn refers upwards to the Gilbert and Sullivan Collection which is the Fales Library's pride and joy.

Fourth, AMC provides fields for including processing data and confidential data. These fields, which are essential for archivists, have no equivalent in standard book-catalogues. They include information about conditions of donation or deposit; names and addresses of depositors; level of processing completed and desirable in the future; and so on. The fields, obviously, must not appear in public catalogues, and are accessed through a secret local code.

Above all, RLIN-AMC provides an opportunity to put general collection-level descriptions into a national database. There are a few institutions, especially those (like the Lockwood Library in Buffalo) specialising in "artificial" literary collections comprising miscellanies of individual pieces of creative writing, which are making item-level cataloguing their normal procedure and putting item-entries into RLIN. Most institutions, however, are keeping their item-level descriptions on local catalogues, which may or may not be MARC-based. (I would not expect to find an entry for each W S Gilbert libretto on RLIN). These independent local catalogues are characterised by imaginative names: LITMSS at Princeton, CATNYP at New York Public Library, SOCRATES at Stanford, SIBIS at the Smithsonian Institution. They exist to put flesh onto the bones of the RLIN entries.

At the same time, in North America Chapter 4 of *AACR2* has now been completely replaced for all purposes by a fat document often referred to as APPM, whose full title is *Archives, personal papers and manuscripts: a cataloging manual for archival repositories, historical societies, and manuscript libraries.*[5]

⋆   ⋆   ⋆   ⋆   ⋆

In looking at the exciting recent developments at the British Library's Department of Manuscripts, we see similar problems being tackled in a quite different way.

The Department of Manuscripts is not perhaps one of the places where you would have looked first for innovations and initiatives, but their new cataloguing system is an imaginative idea which combines use of MARC with descriptive freedom *and* with continuity with the age-old Department of Manuscripts traditions - no mean achievement!

The Department's cataloguing traditions centre on a separation of the manuscripts catalogue from the manuscripts index. The catalogue is a paragon of descriptive academic prose, with abundant use of the Latin language and abbreviations like *viz*. and *et sqq*., and with many sections which end with a phrase such as "the names of the writers of letters etc., are given in the Index to the present Catalogue".

In brief, what they have now done is to design a system

---

where the Catalogue is prepared using a version of Microsoft Word, programmed with a BL Style Sheet format, and the Index is prepared using a version of UKMARC. The Location Register's entries provided one of the models for this Index, and it is not surprising therefore that the new BL Index and the Location Register will greatly resemble each other and should be compatible and transferable. Similarly one could say that the BL's Word catalogue will be their equivalent of LITMSS or CATNYP.

⋆   ⋆   ⋆   ⋆

The compatibility and exchangeability of data between the BL Department of Manuscripts, the Location Register and RLIN-AMC is the key to a number of promising developments for the future.

I should emphasise the outstanding success that AMC has had in establishing itself as the national standard in the United States and in Canada. Big projects in North American librarianship, as you probably know, are almost always mainly foundation-funded. The major funding agencies, (including the National Endowment for the Humanities and the Mellon and Pugh Foundations) have now agreed that *no* funds will be provided for non-MARC manuscripts cataloguing projects in the USA. There is a full-time member of staff, an education officer, at the office of the Society of America Archivists in Chicago whose job is simply to promote AMC across the country. In addition, the Council on Library Resources has been funding the "Linked Systems Project", which should eventually permit exchange of MARC records and inter-system searching between RLIN, OCLC, WLN and the Library of Congress. The future of MARC-based manuscripts cataloguing in North America is thus absolutely secure, and not even dependent on the success or failure of RLIN.

At present RLIN-AMC seems to be going from strength to strength as a national manuscripts database. In addition to the large university collections, it is now picking up more and more local input (particularly from state historical societies) and also including a variety of surveys. One major example is the New York State historical survey. The inclusion of surveys in the RLIN database makes it possible for the papers of very small institutions (which could not afford RLIN membership) to be incorporated. Even the *National Union Catalog of Manuscript Collections* (produced in the Library of Congress and always known as NUCMC) is now being loaded into RLIN-AMC.

We have now reached the stage where a good working relationship has been established between the British

Library's Department of Manuscripts and RLIN, and where an in-principle decision has been taken for the Location Register's file to be converted to USMARC (the conversion

will be done by machine) and incorporated into the RLIN-AMC database. This will mean that a searcher at a single terminal searching for information about the whereabouts of the papers of, say, Robert Graves will obtain from the London Register's data information about holdings in the British Isles, including rather unlikely repositories such as the Modern Records Section of the Greater London Record Office or the Regimental Museum of the Royal Welch Fusiliers, as well as data input by RLIN members about holdings in university libraries in Buffalo, New York, Carbondale, San Francisco and so on.

This will provide searchers with a far better service than they could ever get from "JANET-hopping". For the time being, we are backing RLIN as the best international host, but our data could be shifted easily enough to another host or duplicated onto several hosts without technical difficulty. This is, and will continue to be, our justification for choosing MARC in the face of considerable skepticism and even hostility.

[1]   C M Woolgar. The Wellington papers database: an interim report. *Journal of the Society of Archivists*. 9 (1), January 1988. p1.

[2]   For further detail on the scope of the project, see e.g. David C Sutton. A highly original paper chase. *THES*. 5 April 1985. p.13; David C Sutton. Seeking out literary papers. *The author*. Summer 1987. pp.43-44; David C Sutton. Locating literary manuscripts: some good news for researchers. *British book news*. December 1988. pp.894-895.

[3]   I would like to acknowledge the advice given during these discussions particularly by Bob Cooke, David Ferris, Andy Stephens and the late Anthony Long.

[4]   David C Sutton. A computerised union catalogue of literary manuscripts. *Program*. April 1984. pp.170-173.

[5]   Steven L Hensen. *Archives, personal papers and manuscripts: a cataloguing manual for archival repositories, historical societies, and manuscript libraries*. 2nd ed. Chicago: Society of America Archivists, 1989.