

## The British Library's Approach to AACR2\*

Lynne BRINDLEY: British Library, Bibliographic Services Division, London, England.

*The formal commitment of the British Library to AACR2 and Dewey 19 entailed substantial changes to the U.K. MARC format, the BLAISE Filing Rules, and a variety of products produced for the British Library itself and for other libraries, including the British National Bibliography. The British Library file conversion involved not only headings but also algorithmic conversion of the descriptive cataloguing.*

Along with the U.S. Library of Congress and the national libraries of Australia and Canada, the British Library was formally committed to the adoption of the *Anglo-American Cataloguing Rules, Second Edition (AACR2)* and *Decimal Classification, 19th Edition (DC19)* in 1981. This entailed fairly substantial changes to the MARC format as published in the *U.K. MARC Manual, 2nd Edition* as well as the implementation of the new and more sophisticated *BLAISE (British Library Automated Information Service) Filing Rules*.<sup>1</sup>

There is, of course, never an ideal time for making major changes—politically, economically, or technically; and the Bibliographic Services Division (BSD) found itself having a large number of preexisting separate systems, particularly for our batch processing work, which had grown up over a long period of time and had in most cases been tailor-made to the individual products. Whilst relatively small, BSD is nonetheless responsible for a multiplicity of products and services, almost all of which were to be affected to some extent by the change to *AACR2/DC19*. Briefly, then, a comment on the different services and the degree to which they were affected, thus setting the scene for our decisions on machine conversion.

\*Based on a talk given at the Library Association seminar "Library Automation and AACR2," held in London on January 28, 1981. The views expressed in this paper do not necessarily represent those of the British Library or the Bibliographic Services Division.

## SERVICES AND IMPACTS

### *Printed Publications*

The major printed publication of the division is the *British National Bibliography*. It is arguable that for the printed publications (especially the weeklies) there would have been little justification for retrospective conversion. The files could have been cut off at the end of 1980 and started afresh for 1981—it might, however, have precluded, or certainly have made more messy, the possibility of any multiannual cumulations across this period.

### *Microform Products*

These are mostly individual COM catalogues, both within the BL, especially the Reference Division, and externally, provided through LOCAS (BSD's Local Catalogue Service) to some sixty libraries in the U.K. In many ways those libraries that plunged into automation early, building up files of records derived from central U.K. and LC MARC, were likely to be worst affected. Individual machine-readable files had grown very large and exploited not only relatively *current* cataloguing data, but also full *retrospective* U.K. holdings back to 1950. Also we foresaw *no lessening of use* by libraries taking our catalogue service of the U.K. retrospective 1950–80 file after AACR2 implementation. Therefore the grounds for attempting automatic retrospective conversion of records were indisputable.

### *Tape Services*

U.K. exchange tapes, either as a weekly service or through the Selective Record Service, are supplied to nearly one hundred organisations. The same arguments that there will be continuing selection from the retrospective files apply—therefore, for compatibility and ease of use we needed to consider conversion. The weekly exchange tape service makes a clean AACR1/AACR2 break, but obviously libraries have back files of AACR1 records. Mindful of our responsibility to other organisations and agencies utilising our records, we decided to make our own converted tapes of LC and U.K. MARC records available to tape-service customers to aid their own conversions.

### *Online Services*

Regarding the BLAISE Online Information Retrieval System for U.K. and LC MARC, our concern was to ensure continued easy searching and printing across the total span of files. Without automatic conversion it would have been difficult, if not impossible, to ensure consistency in search elements and index entries (e.g.: In U.K. MARC, series fields 400, 410, and 411 no longer exist, so without conversion a searcher would have to remember specific search qualifiers for pre-1981 records, and different ones thereafter). Without conversion the searcher would need a lot more knowl-

edge of MARC and the history of cataloguing practices to formulate effective strategies.

### *Outside Users of MARC*

Last and very much *not* least was a consideration of what we could do to help the now large community of U.K. MARC users in coping with the changeover. This is now a very large and diverse group relying on BSD for the provision of bibliographic records for whatever purpose. Our own conversion enabled us to provide a multiplicity of aids to libraries. Of particular note are (1) U.K. and LC exchange tapes of converted records, and (2) machine-readable and microfiche versions of our own Name Conversion File, which is being used as the basis for the new Name Authority Fiche.

So, in the context of the variety of our services the case for conversion was strong.

### RETROSPECTIVE CONVERSION

The extent of the retrospective conversion exercise is discussed below. In conjunction with this work we were faced with the necessity of rationalising our COM and print product software (Library Software Package), both to enable it to drive each of the previously separate print applications and to ensure that it had sufficiently sophisticated output facilities to cope with the complexity of AACR2/U.K. MARC 2 records, with their increase in numbers of subfields, their repeatability, all or some, and varying sequences, to produce the specified layout and punctuation across our services.

#### *Extent of Conversion*

We are now in a position to discuss the retrospective conversion exercise. Having decided in principle to become involved with conversion, the extent of our involvement had to be established. British libraries have never had the tradition of building and utilising name authority files, and certainly the concepts fit more easily in the North American primarily online system context rather than in the predominantly batch cataloguing systems established in the U.K. The BL therefore found itself without a machine-readable authority file and began to create one from scratch to enable the important heading changes required by AACR2 to be handled automatically.

Again because of the overriding importance of COM catalogues in the U.K., considerable attention was paid not only to automatic heading changes but also to automatic MARC coding and text conversions bringing the descriptive cataloguing elements also into line with AACR2/U.K. MARC 2, so that catalogue records could be consistent on output whether derived from the conversion or newly created.

The third consideration for conversion was our Library of Congress file

(*Books All 1968-* ), used in the U.K. as part of our cataloguing services and as a file in the BLAISE online system. We had always performed certain conversions on LC records to bring them more into line structurally with the U.K. MARC format. However, U.K. libraries using these records for cataloguing purposes still had to undertake substantial editing. It was therefore decided to use the opportunity to enhance this conversion and bring LC records into line with U.K. MARC 2 to make them of maximum use to British librarians.

To summarise, then, the retrospective conversion comprised three main parts:

1. That part which utilised information stored in the *Name Conversion File*, which records the AACR2 and AACR1 forms of names. This enabled the automatic conversion of major, commonly occurring personal and corporate headings.
2. *Automatic MARC coding and text conversions*—this consisted of specifications at MARC tag and subfield level of algorithms for automatic MARC coding and some bulk text conversions. It resulted in records being converted to a pseudo-AACR2/U.K. MARC 2 format, so that all output specifications, whether by profile or by online inversion, had only to cater for the new format.

These two parts of the conversion are inexorably linked, both conceptually and in programming terms, with frequent references to alternative courses of action dependent on whether a match has been found on NCF. The details of conversion are in "Specification for Retrospective Conversion of the UK MARC Files 1950-1980,"<sup>2</sup> prepared in the Computer Services Department.

3. The third facet of conversion was to our *Library of Congress files* (*Books All 1968-* ), to bring records in line with U.K. MARC 2 as far as possible. Only conversions of tags, indicators, subfield marks, punctuation, and order of data elements have been included; no attempt has been made to bring textual data into conformity with BSD practice. The converted records are therefore in AACR2 form to the extent that LC applies AACR2 to a particular record.

The next section highlights major points of each part of the conversion, commenting particularly on aspects of programming and testing.

### *Name Conversion*

The Name Conversion File was built up by BSD's Descriptive Cataloguing Section over nine months of 1980 and comprises authenticated AACR2 headings with the AACR1 form where different. It will form the basis of an authority file of headings and references for future BSD cataloguing and will be the first publicly available U.K. authority file. The file was maintained using existing LOCAS facilities. Pseudo-MARC records were created recording the AACR1 and AACR2 forms of headings in the format shown in example 1.

## FIELD

- 001 (control number)  
 049 (source code)  
 110.1 \$a Great Britain \$c Accidents Investigation Branch  
 (Name Heading in AACRII Form)  
 710.1 \$a Great Britain \$c Department of Trade \$c Accidents  
 Investigation Branch  
 (Name Heading in AACRI Form)  
 910.1 \$a Great Britain \$c Department of Trade \$c Accidents  
 Investigation Branch \$x See \$a Great Britain \$x  
 Accidents Investigation Branch  
 (Reference for AACRII Name Heading)

Name Conversion File Record  
 Example 1

The file being used for conversion comprised some 12,000 records, of which 4,000 had AACR2 heading changes. The remaining records were authenticated by BSD as correct AACR2 headings without alteration. Of the changed headings most were prolific personal and corporate (particularly U.K. government) headings.

The first stage of the conversion process for U.K. MARC records (1950–80) involved all records being processed against the Name Conversion File to replace AACRI with AACR2 headings and associated references.

In programming terms, the name conversion was relatively easy—relatively, that is, in the context of bibliographic programming. The matching program used was not particularly sophisticated. It took each NCF record, identified the 7xx (AACRI) field, created a key of fifty characters stripping out all blanks, embedded punctuation and diacriticals, and then tried to match the key against each 1xx heading in whatever file was being converted. If there was a match on the key, then the program proceeded to match character by character through the data looking for an exact match. If this was not found, then the NCF record was not processed. Example 2 shows this procedure more clearly.

Of course, this file has not converted *all* AACRI headings, but it has ensured that the majority of headings *likely to recur* (i.e., of any significance in catalogue collocation of headings) have been automatically changed.

### *Automatic MARC Coding and Text Conversions*

This is commonly known as the format conversion program and forms the bulk of the “Specification for Retrospective Conversion.” The original specification was extremely complex, particularly bearing in mind the tight time scales that we were working to. The major difficulty throughout all parts of this facet of conversion was having to specify procedures to accommodate the variety of usage of MARC across thirty years, including previously automatically converted 1950–68 U.K. MARC records; it has

## NCF RECORD

- 710 (AACRI) \$a Great Britain \$c Civil Service Department  
\$c Central Computer Agency#
- 110 (AACRII) \$a Central Computer Agency#
- 910 (AACRII) \$a Great Britain \$c Civil Service Department  
\$c Central Computer Agency \$x See \$a Central  
Computer Agency#

KEY: 10\$AGREATBRITAIN\$CCIVILSERVICEDEPARTMENT\$CCENTRALC

Matching on data—would match Central Computer Agency  
would not match Central Cataloguing Agency

N.B. KEY EQUALS 50 CHARACTERS (Upper Case)

## NCF RECORD

- 700 (AACRI) \$a Walker \$h David Esdaile#
- 100 (AACRII) \$a Walker \$h David E. \$q David Esdaile \$r  
1907 -#
- 900 (AACRII) \$a Walker \$h David \$c 1907 - \$x See \$a  
Walker, David E.#

KEY: 10\$AWALKER\$HDAVIDESDAILE

## BOOK RECORD

Before:

100 Walker \$h David Esdaile#

900 -

After:

100 \$a Walker \$h David E. \$q David Esdaile  
\$r 1907 -#

900 \$a Walker \$h David \$c 1907 - \$x See \$a  
Walker, David E. \$z 100#

N.B. Addition of new reference

Name Conversion Matching  
Example 2

been almost impossible to verify absolutely that any of the automatic changes would cover all cases.

Not surprisingly, this was an extremely complex program. It had to allow for manipulating in fairly precise ways nonstandard and variable data, and had to be designed to cope with occurrences in many different combinations. The programmer had to code for these combinations, some of which may possibly never have been used. It is probably the case that certain combinations do not exist, but this could not be guaranteed over such a large number of records until the total file had been converted.

A good example of the complex logic of this kind of processing is found in the 245 field, where seven complex conditions were allowed for:

*FIELD 245*

- (1) If \$e \_\_\_\_\_ then \_\_\_\_\_ else \_\_\_\_\_
- (2) If \$f \_\_\_\_\_ then \_\_\_\_\_ else \_\_\_\_\_
- (3) If \$d or \$e \_\_\_\_\_ or \_\_\_\_\_ or \_\_\_\_\_  
or \_\_\_\_\_ or \_\_\_\_\_ or \_\_\_\_\_ or \_\_\_\_\_

- then \_\_\_\_\_  
 else if \$d or \$e \_\_\_\_\_ or \_\_\_\_\_ or \_\_\_\_\_  
 or \_\_\_\_\_ or \_\_\_\_\_ or \_\_\_\_\_ or \_\_\_\_\_  
 or \_\_\_\_\_  
 then \_\_\_\_\_
- (4) If tags \_\_\_\_\_ then \_\_\_\_\_  
 (5) If 008 \_\_\_\_\_ and \_\_\_\_\_ or \_\_\_\_\_ or \_\_\_\_\_  
 then \_\_\_\_\_  
 (6) If \$h \_\_\_\_\_ then \_\_\_\_\_ and \_\_\_\_\_  
 (7) If \$e \_\_\_\_\_ then \_\_\_\_\_  
 else if first \$e \_\_\_\_\_ then \_\_\_\_\_  
 else \_\_\_\_\_  
 else \_\_\_\_\_

Repeat for all levels of 245.

Another variation on this theme is that the specification catered for what it expected to find. Again, because of the volume and span of data the expected was not always found. For example, a lot of processing of references is dependent on the presence of a \$x. What do you do when you find a record accidentally without one?

A third problem was that of *interdependency of fields* and subsequent actions. A good example of this is found in 110s and related 910s. If a 110 is changed, you may have to *create* a 910, replace a 910 with another one, or reorganise existing subfields. Then you may have to reorder the field and also flag the action to come back to later in the program. Hence you are switching back and forth across fields throughout the program. You cannot simply start at field one, process sequentially, and then stop. Clearly this makes program testing that much more complicated.

However, those were the problems—really a very small percentage of the whole. From all that has been seen of the converted files so far it has been a highly successful exercise. All of the major MARC changes and many of less significance have been converted automatically by this program—Treaties, Laws, Statutes, Series, Conferences, Multipart works—the resulting records being consistent in MARC tagging structure and in significant headings and areas of text.

### *Library of Congress File Conversion*

It has already been stressed that the automatic MARC coding and text conversions for U.K. MARC were very complex programs. Perhaps even more complicated was the conversion program written to transform LC into U.K. MARC format. The main reason for this is that the U.K. and NCF conversions are one-off programs and a great number of the manipulations could be hard-coded. However, it is intended that the LC conversion program will be used on an ongoing basis against each weekly LC tape. Thus each conversion has been treated as a separate parameter to the

program so that it is general purpose and easily alterable in the light of changes of practice by LC. To give you some idea of the complexity, there are well over 600 separate parameters to the program. I say separate, but in fact they are interrelated parameters, so that if a minor change is made to one it can potentially affect many others.

Many of the problems relating to this program could again only be really apparent in volume testing, not in writing. Each parameter written and tested in isolation was satisfactory, but when they began to be put together in modular form, then the problem of unusual combinations began to show.

Although the conversion parameters for LC records are extensive, they cannot touch the cataloguing data, certainly not nearly as much as in the U.K. MARC conversion. There are added problems in the fact that the records coming to us from LC do not show the clean AACR1/AACR2 break that BSD is adopting. We are having to allow for mixed records from LC at least in the foreseeable future. Details of the LC-to-U.K. MARC conversion are published in a detailed specification.<sup>3</sup>

## COMMON ISSUES IN CONVERSION

### *Testing*

It is possible to draw out common problems applicable across all the conversion work, particularly in testing. They are as follows:

1. Variability of records;
2. Complexity of records;
3. Volume of data;
4. Nonstandard data;
5. Repercussions throughout system.

### *Variability*

This is an obvious problem in the handling of MARC records, but particularly pertinent when trying to do such complex manipulations. The record format itself is of course variable—there are very few essential fields or data elements; most need not be present at all; if they are present, they can be there once or ten times. Standards of cataloguing, and therefore MARC coding, have changed considerably over the period in question, adding to the variability. In some exceptional cases BSD practices are different from those prescribed in the MARC manual, e.g., nonstandard use of title references. All of this results in additional difficulties from specification, through programming and testing. On average we found that one conversion process took two to three times the amount of coding required for more normal computer processing.

### *Complexity*

This is linked with variability and was manifest particularly in the fact that it was extremely difficult to ensure that the programs catered for all



conditions. We found that testing threw up oddities not allowed for in the original specification. In an ideal situation with no time constraints a totally tailored and comprehensive test file should have been drawn up for each facet of conversion. This exercise alone would have taken a good year and would still not have catered for the unexpected data problems. In practice, whilst BSD's Descriptive Cataloguing staff were able to provide several hundred records that tested the majority and most important of the conversions, we always faced the possibility of coming across exceptions. This soon became apparent when volume testing commenced and each new file threw up another combination and a different program route not previously tested.

### *Volume*

The third major factor adding to the complexity of the whole operation was the sheer volume of data to be processed. Approximate figures are as follows:

U.K. MARC	0.7 million records
LC MARC	1.4 million records
LOCAS	2.5 million records

The combination of these three factors—variability, complexity, and volume of data—made testing extremely difficult and expensive in machine terms, in that large test batches of material had to be processed.

### *Nonstandard Data*

Like any large file, U.K. MARC has its share of incorrect data, most of it of no particular significance. However, some problems arose in conversion testing resulting occasionally in corrupted records. One example that springs to mind was the incorrect spelling of months in Treaties, giving problems in the 110 \$b conversion to 240.

### *Repercussions throughout System*

A cautionary note, really: we made a decision that postconversion records should not be put back and overwrite existing master files until they had been through validation programs (i.e., those used for validating new input for BNB and LOCAS); it was felt that this was a necessary safeguard against reintroducing any structurally incorrect records postconversion. It was here again that testing threw up timely reminders of just how much the validation programs had been upgraded and changed since many of the original records had been input through the system.

### *Scheduling*

The scheduling of such a large, complex exercise was extremely difficult, with interdependency of processing related to the success or otherwise of overnight runs. A lot of time was spent before the conversion period in

discussion with our computer bureau to ensure maximum cooperation throughout the difficult time. They were extremely helpful in ensuring operator coverage throughout weekends and priority for our work. One of the problems we encountered was having to forecast the approximate number of machine hours that would be required throughout January 1981 when the bulk of conversion work was carried out. At the time the figures were needed we were still in early stages of programming so no volume tests could be run. Equally, although we were experienced in large-volume processing it was difficult to draw any direct comparisons with production work. Additionally, we had to allow for a heavier than normal production work load towards the end of the year, which always sees annual volumes, cumulations, online file reorganisation, and so on. Scheduling therefore was a fine art to ensure correct priorities for production, the bureau's own work, and conversion, and to minimise contentions for files and peripherals.

### *Staffing*

Of interest is a picture of the human resources involved in this project. What is striking is the magnitude of the task achieved by very few people. The overall management of the project was taken on by existing line management within BSD's Computer Services Department. Two project leaders were appointed, one a librarian and one a systems analyst.

The librarian had a team of four temporarily seconded staff who were totally responsible for all output profile specifications (printed products and COM), testing, and implementation. They also did a considerable amount of checking of test file conversion runs.

The systems analyst was a project leader for three analyst-programmers and one JCL writer. Between them they were responsible for LC and U.K. conversion programming and the new filing rules.

Existing operations staff and others as appropriate within the division were called upon for other tasks.

### *Disruption to Services*

Whilst disruption to our normal production services was kept to an absolute minimum, it was decided that it would be necessary to temporarily suspend certain services through the month of January 1981 while the bulk of the file conversion took place. Throughout the period, the BLAISE online information retrieval system continued to be operational: associated online facilities that would normally allow the despatch of MARC records to catalogue files were suspended to avoid any non-AACR2 or nonconverted records inadvertently updating converted LOCAS files. The production of COM catalogues through LOCAS was suspended for a single month, and the first issue of *BNB* for 1981 was not scheduled until early in February. The schedule for the conversion exercise was adhered to with no major slippage except in the case of our LC file conversion; this exercise

stretched on into the spring for a variety of technical reasons largely concerned with the characteristics of the LC data.

## CONCLUSIONS

Having been so closely involved in this project it is difficult to draw out general conclusions as yet. However, there are some already obvious benefits both for BSD and the wider library community: the rationalisation of our software for COM/printed products will lead to easier maintenance and future upgrading; the introduction of the *BLAISE Filing Rules* across all our products is an improvement; the new LC conversion will make our LC files much more easily usable by the British library community; we have the basis of a U.K. Name Authority File for the first time.

This was a vast and sophisticated conversion exercise and will result in U.K. MARC files probably more uniform in structure than they have ever been. It forms an excellent basis for the continuation of BSD services, especially those based on utilising records across the whole time span, e.g., BLAISE information retrieval, Selective Record and cataloguing services. Equally, because our conversion has been so extensive we have been able to share it: the specification, the Name Conversion File, and the converted U.K. and LC files were all available at minimal cost to libraries in the U.K.

Of course, it is not the 100 percent solution—it was never intended to be—so of course if you look hard enough you will find inconsistencies. However, it has proved that very extensive automatic conversion is possible even with today's state of the art of computing and that BSD had led the way, indeed eased the path of transition to AACR2 for British libraries.

## REFERENCES

1. British Library, Filing Rules Committee, *BLAISE Filing Rules* (BL, 1980).
2. British Library, Bibliographic Services Division, Computer Services Department, "Specification for Retrospective Conversion of the UK MARC Files 1950-1980" (unpublished with limited distribution).
3. British Library, Bibliographic Services Division, "Specification for Conversion of LC MARC Records to UK MARC" (unpublished with limited distribution).

---

Lynne Brindley is head of customer services for the British Library Automated Information Service (BLAISE).