

POSSIBILITIES FOR THE INTERNATIONAL SHARING OF RETROCONVERTED MAP FILES.

By Tony Campbell

One of the main themes of the eighth biennial conference of the Ligue des Bibliothèques Européennes de Recherche (LIBER), Groupe des Cartothécaires (held in Barcelona, 28 September to 2 October 1992) was Retroconversion of map catalogues. Although most of the papers were descriptive accounts of projects proposed or in progress, discussion centred on the practical and theoretical possibilities for the international sharing of retroconverted map files in the future.

It became clear that there were two, apparently irreconcilable, strategies for achieving machine-readable records for historical cartographic materials. Not surprisingly, the approach favoured by each participant reflected both the nature of the materials held in their own library, and the automation decisions that had been already taken. For example, this author outlined the project, now well under way, to convert the British Library's catalogues of printed and manuscript maps.¹ The resulting file, due to be published on CD-ROM in 1995 by Research Publications International, will comprise nearly 250,000 records. The present paper will attempt to analyse the two different approaches and suggest possible ways in which they might be reconciled for the purposes of international record sharing. Reference will be made, at the end, to the strategy proposed by the Consortium of European Research Libraries for dealing with books from the hand press era.

The two approaches are, on the one hand, retrospective cataloguing or recataloguing, and, on the other, retrospective conversion or retroconversion. The first of these, the 'retrospective cataloguing' strategy, stresses the need to maintain the quality of the bibliographical records. The second - perhaps we should call it the 'let's make the most of what we have' strategy - stresses the practical difficulties of maintaining quality, given the large numbers of substandard records. The emphasis of the second is therefore on quantity. This distinction is mirrored in two different approaches to record-sharing:

1. The single record approach is favoured by those advocating retrospective cataloguing. In this scenario, there would be a single, authoritative record for each bibliographically distinct item, created according to agreed cataloguing standards, with each participating library contributing its holding statement to

the common location field. [The British Library's Eighteenth Century Short Title Catalogue is an example of this approach].

2. Common access to records of differing standards and styles - sometimes several for a single bibliographical entity - is favoured by those engaged in retro-conversion. [There is no obvious example of this approach as yet, although it forms part of the proposal of the Consortium of European Research Libraries' 'Primary Database'. These will be described later on].

For modern cartographic materials - essentially maps in series - the single record approach might be fully achievable as far as series level coverage is concerned. However, as the policy in some libraries (for example, the British Library) is to catalogue series only, sheet level coverage would be very incomplete. For historical materials, however, the situation is very different. Older map catalogue entries are generally of a poorer standard than their book equivalents. The elements that mark out map records from those for books (for example, scale) are, simultaneously, the most likely to have been omitted and the most time-consuming to provide. Unlike modern maps, earlier cartographic publications tended to omit dates, leading to considerable variation in the approximations assigned by cataloguers. For these reasons, common access to records describing each copy of a work seems the best way to deal with the large quantities of older converted records.

Few dispute that the single record approach represents the ideal. But has it any chance of being achieved? How will the necessary resources be made available for the retrospective (re)cataloguing of some hundreds of thousands of early maps? For example, only a tiny minority of the British Library's extensive file of map records (those created post-World-War II) are likely to be thought fully 'adequate', when measured against any kind of minimum record standard. This is hardly surprising, considering that the catalogue descriptions have been made over almost two centuries, following a succession of different rules. This situation will be repeated in other large and long-established collections.

While there are existing programmes to (re)catalogue categories of books (defined by period or national production), none such is even contemplated for maps. The single, authoritative record scenario, therefore, would inevitably exclude the great majority of converted records. This would leave most early maps undescribed, and present an insurmountable cataloguing challenge. There is also the particular bibliographic problem raised by maps (usually engravings) in determining, without side-by-side comparison, whether two items are identical or in some way distinct (different states of the plate). Thus the aim of creating a single authoritative record for a specific item (in the manner of a cartobibliography) is

likely to be unrealisable for early material until, perhaps, images of equivalent examples held elsewhere can be readily transmitted to allow the cataloguer to make a direct comparison with the map in hand.

Such problems occur less frequently with modern cartographic materials. These are also more likely to have been catalogued according to more rigorous recent rules and held in automated files. They would therefore lend themselves more readily to the single record approach. However, there would surely be little support for a strategy which separated 'modern' and 'historical' materials (however defined) into separate files constructed along quite different lines.

Where retrospective cataloguing concentrates on the bibliographical purity of the records, retroconversion emphasises the importance of gaining access to the widest possible range of records. It underlines the 'finding' function of a catalogue entry over the 'bibliographical' function. Since the British Library Map Library alone will eventually have available up to 300,000 records, and other large datasets are being prepared elsewhere, it is not unreasonable to imagine a potential combined file of up to a million records, accessible online or via CD-ROM.

Unless a very significant amount of previously unidentified resource is made available, and assuming it is agreed that there is no justification for splitting the continuum of mapping between 'historical' and 'modern', it is unrealistic to pursue the single record approach as a primary aim. Sub-standard records, most would agree, are better than none at all. But what do 'substandard' and 'quality' mean in practice? In attempting to define these concepts, a possible compromise will emerge - based on the thinking of the Consortium of European Research Libraries.

Nature of a future shared file

Though there might not be agreement on all the details, it is clear that many older records for historical material are rightly considered to be 'sub-standard'. There are perhaps three ways of dealing with this problem:

1. We could insist on certain minimum standards and flag those records that fail to meet these. 'Sub-standard' here would probably have to be defined in terms of the absence of data from particular fields or sub-fields.

Distinguishing the quality of data in a given field, for example the style of title transcription, does not readily lend itself to automated identification. Quality would often relate to the date of a record's creation but, for the British Library at least, there is no means by which this could be indicated. These records

could then be omitted altogether from any shared database. However, such a decision would effectively exclude the vast majority of historical records and would thus defeat the object of the exercise. Such a rejection of the possible data (perhaps 80%) would be a massive wasted opportunity.

2. We could insist on certain minimum standards, but present the 'sub-standard' records to the user with some sort of a general 'health warning'. When the deficiencies refer to absent, rather than poor, data, this would be immediately apparent to the user when the record is displayed on the screen (for instance, if the scale had been omitted).
3. We could accept and display all the available records in the same way. This has the merit of comprehensiveness, which many would consider to be the main aim of such a project. If this were the case, all records would presumably output according to a geographical thesaurus, to be mentioned in a moment. As long as all records included searchable dates, the entire file could supply answers to date-range queries as well. Those records which lacked relevant information (scale, author or thematic heading, for example) would fail to output under searches defined in those ways. In other words, as far as those elements were concerned, minimum standards would, in effect, have been achieved and the records suppressed. The fact that some searches would produce incomplete, and hence misleading, answers could be pointed out to the user.

Suggested aims of a shared file

Aims would normally be stated at the outset. It seemed sensible, however, to explore first the range of likely possibilities. If it was agreed to adopt the preferred strategy, the last of the three - in which all available records were welcomed - the combined file might have the following purposes:

1. For modern materials - as a cataloguing tool, in other words, providing fully acceptable records that could be downloaded.
2. For historical material - as a cataloguing guide. When valuable information is held in sub-standard entries, this could be used in the creation of a new record. The 'quality' approach is valid here, since the better the record the more useful it will be for this purpose. Where there are several entries for (apparently) the same publication, the best record would presumably be chosen.
3. For retrieval purposes - to provide as comprehensive as possible an answer to the bibliographer's or researcher's question: what material is there for region or place x, in period, and [perhaps] concerning theme z, in any of the contrib-

uting libraries? Since some enquirers would be interested only in the holdings of a single library, while others would wish to include or exclude specific institutions or search simultaneously as many libraries as possible, a range of output alternatives should be offered.

4. As a cartobibliographical aid - to enable continents, regions or individual countries to identify and download all material relating to their own history, as found in the participating libraries' files.
5. For the future -as a first stage in the creation of a file with single, authoritative records, if and when resources allowed. The combined database could itself bring together the separate records (apparently) relating to each item. In a re-cataloguing operation, these could then be merged into a single record with multiple locations. The argument that it would be difficult to find support for the second stage if it were separated from the first can be answered in two ways. It is better to have the first stage than nothing at all, and the proposal could be framed so that the two stages formed part of a single overall programme.

The Cartographic Context

Whatever the overall approach adopted for a co-operative map file, edited thesauri will probably be the essential pre-condition for international co-operation. Only in that way would it be possible to overcome the variations of language or the erratic spelling on earlier materials, which would seriously reduce the effectiveness of title-word searching. Three types of thesaurus can be identified:

1. *Geographical thesaurus*. The single most important difference between a book file and a map file is that a book catalogue is organized under author headings whereas a map catalogue requires some form of primary geographical control, usually in the form of a hierarchy. Author headings have a degree of predictability, but the ways in which particular libraries have subdivided the regions and areas of the world vary widely and unpredictably. Without edited geographical headings, access would depend on the accuracy with which title words could be anticipated.

'EUROCARD', a multilingual thesaurus of both current and superseded geographical names, was intended to perform just such a function. However, the consortium of European national libraries decided last year not to seek support from the EEC. The project's future is now less certain, but it is hoped that other arrangements will be found to achieve its aims. It is intended that

EUROPACRT would offer consistent access to users to users, who would not need to know which was the 'authoritative' form of their place-name. All recorded name forms would lead to the same group of entries, comprising all relevant material, however titled.

2. *Name thesaurus*. It would probably not be difficult to reach agreement that all author headings were held in their vernacular forms. However, there would certainly be a need for access via other-language forms. The creation and editing of such a combined name thesaurus would require a collaborative effort, perhaps with expert representatives from different countries dealing with their own names. Given that mapmakers, map engravers and map publishers feature infrequently in existing name authority files for books, and given the lack of a fully comprehensive and authoritative biographical dictionary for the mapmakers of the world, much basic work would have to be done.
3. *Thematic thesaurus*. Numerous local thesauri exist to control the thematic element of map collections (administrative maps, canal maps, military maps etc). Since there is no adequate thesaurus available 'off the shelf', many institutions (the British Library Map Library among them) have devised their own. It is hoped that the majority of the concepts used are sufficiently close to allow cross-reference and translation, via automated concordance. The best approach might be for a single, suitable qualified individual to examine existing systems and see how they might be merged. It would be impractical for the new authoritative heading forms to be manually assigned to the large numbers of records involved. However, a multilingual system that allowed continued use of a particular library's own headings, via concordance transparent to the user, would presumably be acceptable.

Consortium of European Research Libraries

Maps are not the same as books, and major difficulties are caused for map librarian when these differences are ignored. On the other hand, map collections are fewer in number and generally far smaller than their book equivalents. Since research grants and project resources are, understandably, more likely to be devoted to books than maps, it makes sense for us map librarians to watch closely those developments in the world of books that may have relevance for us.

The best example is the Consortium of European Research Libraries (formerly the European Working Group on Retrospective Cataloguing), set up to create a database of pre-1830 printed books held in the collaborating libraries. After a number of meetings and a consultant's report, the Consortium agreed on a compromise

between the contradictory 'quality' and 'quantity' approaches described earlier.² Their proposal combines the principles of retrospective cataloguing and retroconversion.

The plan is for the creation of two files, which would reflect the two approaches. The 'Primary Database' would represent a combined file of the existing records supplied by the contributing libraries. Each library would be responsible for its own records. Since these would not be edited centrally, they would be of varying standards. However, they would all be held in UNIMARC format, whose maintenance committee has agreed to accept the necessary amendment. The European Consortium's first target will be to issue a retrieval file, as soon as possible - on CD-ROM or online - containing the records that are readily available.

These records can be considered the raw material for the higher level 'Union File', whose creation would form the project's main, if longer-term aim. Here there would be a single record for each bibliographical entity, catalogued, according to ISBD(A), to an agreed standard. To this would later be added the holding statements of the contributing libraries. It is proposed that name authority would be the responsibility of the country concerned. These authoritative records would be distributed via an online host.

Although the starting-point was somewhat different, it is interesting how closely the Consortium's conclusions mirror those set out in an earlier draft of this paper (distributed in November 1992 via the LIBER mailing list).³ These are broadly repeated here. A number of other potentially contentious issues have been flagged by the Consortium for future resolution, such as funding, charging, the legal complexities surrounding copyright and ownership of the data, update facilities and, most vital of all, the practicality of merging highly varied files.

Suggested future action

By 1996, the British Library Map Library should have moved into the second phase of its long-term retroconversion project - for example incorporating the atlas-map descriptions currently being created by an independent researcher, Rodney Shirley. The Bibliothèque Nationale's map retroconversion and the collaborative German map project should also be complete. Other small and medium-sized files already exist (as documented in the recent LIBER survey), and further projects will no doubt be under way by then, if not finished, both in Europe and elsewhere.⁴ Within the next few years, therefore, sufficient historical and contemporary map descriptions will have become available to make the issue of an international collaborative map catalogue a realistic possibility. Since the planning

could take several years - although the European Consortium hopes to achieve its Primary Database within one year - now is the time to find out if there is serious interest in such a project. The British Library Map Library and our partner in the maps retroconversion venture, Research Publications, believe that such a combined cartographic file is the logical and necessary development to follow the creation of large retroconverted files.

If the foregoing analysis is seen as a useful starting-point, an international group will need to be formed. Its first task would be to define an overall strategy - whether, for example, to follow the European Consortium's example or to strike out in another direction. Thereafter, the detailed requirements would need to be specified, and the following operations, at least, planned:

1. The technical analysis of existing records (and of their mergability) and a detailed assessment of the online and CD-ROM options.
2. The planning and creation of thesauri for cartographic materials: author name, geographical name (EUROCARD?) and thematic.
3. Financial analysis - whether external funding needed to be sought - and clarification of the legal aspects already mentioned in the European Consortium context.

I look forward to the reactions of my colleagues to these suggestions.

Notes

1. 'Retroconversion of the British Library's map catalogues: the art of the possible', due for publication in European Research Libraries Cooperation: the LIBER Quarterly. A more technical description of the processes involved is to be given in the paper 'Conversion of the British Library's map catalogues: the keys to success', to be delivered to the Satellite Meeting (19-20 August 1993).
2. See the (unpublished) Consortium of European Research Libraries, 'Final Report on establishing a database for records of European printing of the hand press period (c. 1450-1830)', May 1993.
3. 'Maps retroconversion: a possible blueprint for future international cooperation', distributed as Appendix 2 to Groupe des Cartothécaires de LIBER, newsletter 26 November 1992 'To all colleagues concerned with retroconversion/cataloguing'.

4. ibid., p.1, listing European library initiatives.

Tony Campbell
The British Library
Map Library
Great Russell Street
London WC1B 3DG
UK