

REPORT ABOUT DEVELOPMENTS IN RETROCONVERSION FROM THE GROUPE DES CARTOTHECAIRES DE LIBER

By Jan Smits

On request of the IFLA, Geography & Map Libraries section part of the 8th conference of the Groupe des Cartothécaires de LIBER, held in Barcelona in 1992, was devoted to the subject of "retroconversion". This was to have a follow-up during the IFLA conference in 1993 in Barcelona. This paper is based on the papers read and on the ensuing discussions. All papers will be published this year in one of the issues of "ERLC The LIBER Quarterly" (See list at the end of this paper. Summaries in English, Spanish and Catalan are distributed with the conference papers).

During the LIBER-conference some 9 papers were devoted to the subject retroconversion by speakers from 7 countries. Some of these may be read to this meeting in an adapted form, as the process of retroconversion is continuing and new insights or developments may have changed the nature of these papers. When I summarize the papers I can say the following. 4 were concerned with pre-conversion research -OCR = optical character recognition (Belgium, Spain), keyboarding (Denmark, Great-Britain), codification of elements (Great-Britain, Spain), enhancements of records (Denmark, Great-Britain); 3 of them were concerned with actual conversions (Denmark, Germany, Great Britain); 2 were concerned with end-products (Denmark, bibliography; Great-Britain, CD-ROM), and 3 were concerned with co-operation (France, The Netherlands). When I wrote this report unfortunately only summaries of the two Spanish papers were available.

Let us have a look at the four aspects I mentioned one by one.

Pre-conversion research

Obviously a lot of this research is done in general by libraries who are concerned with retro-conversion. The most efficient would be Optical Character Recognition (OCR) which is a technique already used within and outside the library-field. However as the data to be captured is of different quality depending on the age of the records to be converted it takes a lot of error and trial before satisfactory results are available. OCR combination with database- or textbase-programmes can

result in data with less than 1% errors. Of course unclear type and handwritten data may make it necessary to use a mixture of techniques to convert the whole of a catalogue. The OCR-technique could well serve better if libraries in Europe would file more proposals for the 'European Communities Telematic programmes, Libraries programme, Action Line I, part 2': "Retrospective conversion of catalogues of important collections at international level". Till now an insufficient number of proposals have been submitted and accepted.

Another way of converting the data is by keyboarding. This will make the amount of errors almost neglectable, presuming that some paleographers will help. One of the items to be carefully checked against the original catalogue card or bibliography description are the shelf-marks or call numbers, to prevent lost items. But the costs of keyboarding will be higher than using an OCR-technique. However the next phase may be made cheaper by this. Because when we have captured the data in one way or another they have to be encoded to be able to serve in a MARC-environment. This encoding can be done simultaneous with the keyboarding - which means that it has to be done by qualified personnel- or afterwards.

The last phase is the enhancement of the converted record. This can mean that further searchable elements are added - by depending on the description or by the document in hand- and the addition of new headings (personal and corporate names, area, subject, etc.).

Actual conversions and end-products

One retro-conversion project is completed. The Royal Library of Denmark converted by keyboarding and enhancement all its 1986-1989 records to their database REX (using the danMARC(CM)-system, an adaptation of UKMARC) and in this way could produce *The Danish National Bibliography, Cartographic Materials 1986-90*, which was presented during the conference. However as the policy of the Royal Library is to have all catalogues retro-converted by the year 2000, the harder part will come yet. As far as is known now, this will mainly be done by OCR-techniques. However Mrs. Vejlsgaard will inform you how the retro-conversion in her department has and will run.

Another project was well on its way already. In Germany three state libraries (German State Library, Bavarian State Library, State and University Library Göttingen) decided to create a database through a project called *Cataloguing of historically relevant pre-1850 map holdings*. In its first phase this IBAS-database (using the marc-format MOSES) will comprise approximately 140,000 records of sheet maps and maps from atlases. However this is not a pure retro-conversion project. Some 40,000 map-records of a previous project indeed will be made ma-

chine-readable. But all records will be re-catalogued by autopsy, as the Deutsche Forschungsgemeinschaft (DFG, the subsidizing body) saw consistency in the quality of records as more important than quantity. The other 100,000 records to be created are not retro-conversion records, but retro-cataloguing records, i.e. they have not been catalogued before. This project is only the beginning of a far greater project: an union catalogue of maps and atlases in Germany.

A retro-conversion which usually is outside the scope of the ordinary mapcurator is that of the Ministry of Defence, Military Survey, Acquisition and Library Group in Great Britain. Till november 1992 they retroconverted some 485.000 map records (approximately 40% of the whole collection) at a rate of ca. 520 records a day into their system MODMAP. This system is based on a ORACLE SQL forms application, with functions enabling graphic searches and plots designed around GIMMS software. Some of the functions are a higher version of Cartonnet¹ which enables searchers to retrieve maps by using geographical co-ordinates. Unfortunately for us they do not use a MARC-format, which makes possible exchange of records rather difficult. Few of us can say that initial database population is carried out by 40 data entry staff, which makes for the high figures. The Ministry of Defence is very satisfied with the new retrieval possibilities MODMAP offers.

The conversion in the British Library, Map Library of the 15 volume *"British Museum catalogue of printed maps, charts and plans"* and its supplements will be well on its way by now. The project is contracted to a private company, which eventually will produce a CD-ROM containing almost 250,000 records. The conversion is done by keyboarding and concordances will be run against the file to impose consistency. As with the German project the converted records will be added to a single file, which contains converted and current records, to maintain a continuum. Though the British Library would like to opt for quality out of circumstances they opted for records which are not enhanced by autopsy, but which are flagged to show their origin. This has been done out of practical necessity, because the project can only take place in a limited span of time and because the money is available only now.

Cooperation

In France there were two parallel developments which helped to convert existing records. The area-thesaurus of the BN-OPALINE system was loaded with all the historical names of the French cities. Together with some activities in the regions of France, and the efforts of the Commission de Documentation du Comité Français de Cartographie this provided opportunities to combine work on retro-

conversion and on an union catalogue. I hope that Mr. Duchemin can assure us that his euphoria about this sudden surge for co-operation is still in vogue.

As said before the same kind of co-operation is taking place in Germany. This of course has also the potential of developing into a union catalogue which may consist of retro-catalogued and current records.

There was one paper by Mr. Dirk de Vries of Leiden University with a distinct call for quality through co-operation. As this paper will not be published I shall quote it at large.²

His first remark was that European map librarianship was at risk to fall behind with developments in the world of books, where retro-conversion appeared much more pressing by the sheer fact of the quantities involved. In the past years two conferences have taken place in München which derived "from the fact that projects involving retrospective cataloguing were under way or about to be initiated almost everywhere in Europe, without a common European concept, often even without precise knowledge of achievements elsewhere which might be taken into account".

On the problems of quantity and quality Michael Smethurst, director-general of the British Library and president of LIBER, explained that, when most European research libraries started huge retroconversion projects of their card catalogues, they had no other choice than the *conversion* of the existing records. The number of records are too large to permit re-cataloguing. As long as it is for the rather simple purposes of retrieval and circulation of current material and the OPAC is primarily used as a finding list, only conversion of the existing card-title will be enough. But when you need exact and consistent documentary information on behalf of scholarly and historical research and you have not the collection at hand, than a mere finding list of converted titles of different quality and by consequence inconsistencies, will turn out to be inadequate and re-cataloguing based on the examination of the document will be inevitable.³

This resulted in a wish to form an pan-European database for books of the hand-press period (1453-1830), based on the following reasons:

- It will give better service to scholars and researchers;
- Sharing cataloguing means sharing costs;
- Recognition of a common European printed heritage asks for exchange of bibliographic information at an international level;
- Libraries in Eastern and Central Europe need external assistance to provide access to their research collections.

Based on a study by the Center for Bibliographical Studies and Research of the University of California⁴ the Working Group gave the following recommendations:

1. the creation of a single computer utility in Europe to house a database of ca. 1450-1830 records on the European printed heritage;
2. the adoption of UNIMARC as format
3. the structure of (ISBD(A) (= antiquarian) becomes the suggested cataloguing code for new titles;
4. the records for the period 1830-1950 should be assembled, as a separate project, as quickly as possible by means of retro-conversion projects.

The main drift of the rest of his paper was focused on how we can and should profit developments like this in a technical and an operational way.

EUROCARD

Several papers referred to the *EUROCARD*-project. One of the problems one encounters when retroconverting is headings. Most catalogues, lists, bibliographies use a heading-system which usually is out of date or not detailed enough for present purposes. To alleviate this problem a collection can of course create a new system of headings which can be used for retro-conversion and current cataloguing.

However when we would like to exchange bibliographic data it would be easier if headings can be exchanged as well, to alleviate the work-load of the receiving party. Of course there are universal systems like Dewey, UDC, etc., but especially area-headings are not very detailed and are not easily usable for historical entries.

In 1991 some 7 national libraries came together to discuss this matter, which resulted in a proposal under the "European Communities Telematic Programmes", called *EUROCARD*. This was a project proposal for the development of a multilingual geographical names thesaurus for the benefit of (European) end-users and libraries with large and/or important cartographic collections of modern and historical [= old] maps. The key aims of the project are:

- improve direct access for end-users, via geographic access points, to large and/or important cartographic collections ...;
- improve subject access for end-users to geo-referenced materials in non-cartographic collections ...;
- develop a cataloguing tool for (large) cartographic collections;
- reduce the costs of cataloguing cartographic materials;
- develop a common standard for access and encourage the general adoption of this standard;

- support future exchange of cartographic records via UNIMARC.

Multi-linguality means that the entries will be multi-lingual. but also that the system itself will be multi-lingual.

However I have used the word *was* on purpose. Not that *EUROCARD* is dead, but the European Commission did not accept the proposal in first instance. At the moment the proposal is redrafted, emphasizing especially that it concerns a thesaurus and not a standardized classification, and it meant for end-users in and outside the library field. Hopefully will be developed one way or the other, with or without acceptance by the European Commission.

By the way. The Munich Consortium wants to use the *EUROCARD* thesaurus for its area subject-headings in its 1453-1830 database!

Discussions

These focused mainly on the controversy *quantity* versus *quality*, though closer examination showed that it is not really a controversy, but more an economic principle. Tony Campbell referred to the search for quality as a problem rather than an ideal. Not that he is against quality, but forced by pragmatism he has to be satisfied with less, at least for the present. In the Munich consortium they have opted for quality by the method of autopsy and consequently have to put up large amounts of money to achieve their goal. Some like Lothar Zögner have opted for quality, but their project is concerned with retro-cataloguing and not with retro-conversion. However when one opts for quantity this does not necessarily mean that quality never will be achieved. Retro-converted records without data-enhancement can be flagged or coded. In this way their quality can be recognized in a single database for retroconverted and current material. If the flagging or encoding is done in levels of quality than later groups of the same level may be selected and retrieved for enhancement of data.

However if one is willing to co-operate and one has not a too tight time-schedule one can guard quality by limiting oneself in the records to be converted. As I have understood from Pierre-Yves Duchemin they limit themselves first to "French" records and later hope to acquire quality records of foreign holdings from other libraries. Indeed when a profile of a database can be made and matched with ones wishes a profitable exchange can ensue. The profile should give answers to questions as: How many records are there of a certain country or area with a certain subject and in a certain era, and what is their bibliographic quality. With the possibilities for data-manipulation computerized systems provide it must be possible to create such profiles.

Because of restraints in personnel and means everybody agreed that retroconversion programmes may not interfere with current routines. If possible work should be contracted outside the department. If necessary library-schools may give short specialized courses to train non-library personnel in such a way that they can work for retroconversion. If collections in a certain area or country do not all at the same time retroconvert their catalogues, but one after the other, skills and experience of the thus educated personnel may be used more profitably.

Dirk de Vries questioned if it would not be better to use the personnel and means one could possibly get for retro-conversion for cataloguing the backlog map departments may have. Or would it be more advisable to start cataloguing the maps in books and serials and in this way save them from oblivion!

Finally I would like to remind you that the 'LIBER Library Automation Group', on order of the 'Council of Europe' (Council for Cultural Co-operation), prepared *Guidelines for retroconversion projects*, published in the *LIBER Newssheet* and the *IFLA journal*.⁵

List of papers read and to be published

Agenjo, Xavier (Biblioteca Nacional, Madrid, Spain): The automatic purge of cartographic records originated by retroconversion processes: the integration module ARIADNA.

Campbell, Tony (The British Library, Map Library, London, United Kingdom): Retroconversion of the British Library's map catalogues: the art of the possible.

Duchemin, Pierre-Yves (Bibliothèque Nationale, Paris, France): Retroconversion of french cartographic material card catalogues: an overview of the situation.

Herrero, Ana (Biblioteca Nacional, Madrid, Spain): Retroconversion of cartographic material in the national library.

Parkin, Murray (Ministry of Defence, Military Survey, Surrey, United Kingdom): MODMAP.

Milis-Proost, Greta (Universiteitsbibliotheek, Gent, Belgium): Experiencing of the Ghent map catalogue.

Veijlsgaard, Susan (Det Kongelige Bibliotek, København, Denmark): Retroconversion by keyboarding of the Danish National Bibliography, Cartographic Materials 1986-1990.

-
1. Morris, Barbara A., R.G. Healey and T.C. Waugh, From map catalogue to database with graphics in view. In: Bulletin Ligue des Bibliothèques Européennes de Recherche, No. 28. 1986, p.22-29.
 2. Vries, Dirk de. Retrospective cataloguing of maps in Europe: a call for quality and co-operation. (Not published).
 3. Smethurst, J. Michael, Opening address. In: Retrospective cataloguing in Europe: 15th to 19th century printed materials: proceedings of the international conference, Munich 28th-30th November 1990. München, Saur, 1992. 194 p.
The proceedings consists of two parts: I: country reports on retrospective cataloguing, 14 articles; II: From principle to practise, 18 articles.
For more information on this one can read: "European Working Group on Retrospective Cataloguing: Final Report. In: ERLC The LIBER Quarterly, Vol. 2 (1992), No. 1, p. 47-60".
 4. Snyder, Henri and Heidi L. Hutchinson, A comparative and analytical study of cataloguing rules employed in Europe for the cataloguing of the hand press: prepared for the Working Group appointed by the organizers of the Munich 1990 conference on retrospective cataloguing and conversion in Europe 1500-1900: Revised draft. 1991, 195 p.
 5. Bossers, Anton and Derk Law: Guidelines for retroconversion projects. In: IFLA journal, Vol. 16 (1990), No. 1, p. 32-36.

Jan Smits
Koninklijke Bibliotheek
P.O.Box 90407
NL - 2509 LK The Hague
Netherlands