1 Introduction

1.1 Background

Since the introduction of automated library systems and of online public access catalogues (OPACS) in particular, one of the greatest challenges for library managers has been the transfer of older records into the online catalogues, particularly of records that were created before the advent of computing. Although some lucky libraries have already achieved this goal, many others are still far away from closing their old card, sheaf or book catalogues. In the mid-1990s it was estimated that in the UK approximately 50 million records (28 million in higher education libraries) were still remaining to be converted (Bryant, Chapman & Naylor, 1995; Bryant, 1997), and 52 million records (of titles published after 1945) in Germany (Beyersdorff, 1993).

It is important to note that retrospective conversion (or "retroconversion") is not the same as retrospective cataloguing (or recataloguing). Whereas the latter refers to the original cataloguing of library material, the former simply means the transformation of already existing manual records into machine readable form (Hills, 1993, p. 47–48; Chapman, 1996, p. 16). Although many librarians would prefer recataloguing, i.e. to create "perfect" records in full conformity with present cataloguing rules, in most conversion projects pragmatic reasons such as time and cost restrict this approach to a minority of "problem" records (e.g. serials). On a large scale, it is normally not feasible to afford retrospective cataloguing (Dugall, 2001, p. 113–114).

However, the mere conversion of existing records into machine readable records is not cheap either; the cost of such a project may well exceed that of the automated library system itself (Library Information Technology Centre, 1994, p. 1). Studies of large conversion projects – where often a combination of various options and techniques is used1 – have shown that on average the cost per record can be between 2.42 and 4.23 Euros,2 which means that in total enormous sums of money are required for larger projects or national programmes as the ones mentioned above (UK: 130–160 million Euros, Germany: approx. 185 million).

---

1 See, for example, Hills, 1993; Library Information Technology Centre, 1994; Bryant, Chapman & Naylor, 1995; Seissl, 1997; Dugall, 2001.

2 For the UK, Bryant (1997, p. 557) reported a mean cost figure of approx. £2.42–3.23 (£1.5–2.0), whereas in the more recent CURL study an average cost of €4.23 (£2.62) was estimated for "straightforward" 19th and 20th century items, and almost twice as much for complex or specialist items (Leeves, Butler & Mealia, 1999, section 8); in Germany, an average cost of approx. €3.5 per item was estimated (Beyersdorff, 1993, p. 304–307). [Currency conversion is based on the rates published on 20/03/2002.]
The *scanning* or *digitization*\(^3\) of the catalogue cards has become common practice in retroconversion – not only as a prerequisite for approaches that involve optical character recognition (OCR), but also when a digital duplicate of the catalogue is needed to support conversion work on a computer screen (e.g. typing, tagging, database searching). For example, Stoklasová (1999; 2000) describes a typical retroconversion project as a three-step process:

- **Step I:** *Scanning* – creation of high-quality images;
- **Step II:** *Transcription* – conversion of the images into (unstructured) ASCII text, either by OCR or manually;
- **Step III:** *Structuring* (tagging) – conversion of the ASCII text into structured records (e.g. UNIMARC), either automatically or manually.

As the scanning step can be done quickly and at reasonable cost, the idea emerged to apply suitable browsing software to the collection of card-images which would make it possible to offer it – both to library staff and to the users – as some sort of auxiliary or provisional online catalogue. The first known example of such a *card-image OPAC* was established at the Princeton University Library in 1994 (Henthorne, 1995).\(^4\) From the mid-1990s on, similar catalogues started to appear in Europe,\(^5\) showing some variation of the browsing component – e.g. simple alphabetical browsing, browsing of (partial) indexes, retrieval of OCR processed text – but always displaying the digital image of a catalogue card as the full view of a retrieved record (Figure 1-1). Some of these catalogues were originally offered on in-house networks, but soon the WWW became the commonly used platform.

Pietzsch (1998b, p. 482) summarizes the advantages of such electronic versions over their card counterparts as follows:

- retrieval speed – only trained librarians can search a paper catalogue faster than its online version;
- saving of users' time – no need to go to the library for searching the catalogue;
- independence from the library's opening hours;
- multi-dimensional search options (e.g. when all headings are offered as searchable text).

---

\(^3\) Following standard practice, these terms are used synonymously here; strictly speaking, *digitization* refers to the conversion of any analog material into digital form, whereas *scanning* means the digitization of specifically image-based analog material (Lee, 2001, p. 35–36).

\(^4\) Bork (1997) mentions briefly that archival applications of a similar kind have been used for some time; however, in the field of librarianship, the Princeton project was presumably the first such application.

\(^5\) Primarily, but not only, in the German-speaking countries.
Fig. 1-1: A typical record display in a card-image catalogue\textsuperscript{6}

Whilst the last-mentioned criterion applies only in certain cases, the following aspects need to be added to the above list:

- printing out / downloading of records – only possible in an online environment;
- online book ordering – appropriate components can be attached to the display of card-images;
- saving of library space – the availability of an online version makes it possible to remove the card cabinets.

However, from a critical point-of-view one could also argue

- that, with the exception of a few sophisticated solutions, no options for retrieval are offered that exceed those of traditional card catalogues;
- that the users – who are used to gaining some added value from computerization – might be frustrated by such solutions;
- that modern information technology is used (or abused) for the resurrection and perpetuation of catalogue cards which basically are relics from a past age and should be buried and forgotten rather than offered on a larger-than-ever scale.\textsuperscript{7}

\textsuperscript{6} National Library of the Czech Republic, General Catalogue I

\textsuperscript{7} This applies primarily to card catalogues in German-speaking countries where antiquated cataloguing rules were used far into the second half of the 20th century.
1.2 Terminology and definition

So far, no standard terminology has been established for OPACs of the kind described above. Often they are referred to as "electronic", "scanned" or "digitized" card catalogues, or – particularly in the German-speaking countries – as "image catalogues". In this dissertation, not only the (synonymous) terms card-image catalogues and card-image OPACs will be used, but also – as an analogy to the widely-used term OPACs – the newly proposed acronym CIPACS (card-image public access catalogues).

CIPACS can be defined as online library catalogues that are based on databases of digitized catalogue cards and more or less sophisticated mechanisms for browsing or searching. For the purposes of this dissertation, online is defined as "available over the WWW", which means that mere in-house solutions will not be considered.

1.3 Aims and objectives

This dissertation aims to investigate CIPACS and to assess their impact on libraries and users. Are CIPACS useful OPAC alternatives or only low-cost interim solutions that are barely acceptable to the users? The general research goal is to establish an informed and up-to-date view of this area, which will be obtained by answering the following questions:

(a) What CIPACS exist so far and where can they be identified? Which libraries in which countries have such catalogues, to what extent and how have they been implemented?

(b) What kinds of CIPAC have been developed and how do these approaches differ in terms of retrieval capabilities and potential?

(c) What are the main problems and issues interconnected with the creation and implementation of CIPACS?

(d) How do CIPAC users cope with these systems? What is their current position in terms of awareness, behaviour and opinions?10

---

8 See also section 3.5.1.
9 This acronym was first proposed by Denis Reardon, University of Central England. Even if some of the image catalogues are electronic versions of sheaf and/or book catalogues, the majority are computerized card catalogues so that in this dissertation, for convenience, "CIPACS" will be used.
10 In the original research proposal, it had also been envisaged to deal with the identification and ranking of criteria for the evaluation of CIPACS; however, after careful consideration it was felt this would require
In order to answer these questions this dissertation provides the following deliverables:

- A world-wide survey of CIPACs that exist so far in various countries and contexts, their main characteristics, features and retrieval capabilities;
- An analysis and discussion of the major issues in CIPAC implementation and use;
- An analysis of the reactions of a selected sample of CIPAC users.

1.4 Scope and limitations

Although this study intends to provide a comprehensive overview of the investigated area, some limitations need to be mentioned. First, the following account will focus on the application of CIPACs rather than on details of the technology (hardware, software, networks) that operates "behind" such catalogues. Second, although there are no geographical restrictions concerning the CIPACs surveyed, the analysis is predominantly based upon such applications for which materials (literature, library questionnaire, web-pages) in English and/or German were available. Third, the dissertation concentrates on libraries; similar applications that may exist in other contexts (e.g. archives, museums, industry) have not been covered. Finally, the user survey is exploratory by its conception and therefore leads to preliminary rather than definite results.

1.5 Methodology

Parts of the dissertation are based on both published and unpublished literature obtained by following up the references found in papers already known, by searching databases/bibliographies on CD-ROM\textsuperscript{11} or available freely on the www,\textsuperscript{12} by using web-based search engines,\textsuperscript{13} by systematically looking at the web-sites of all CIPAC libraries, and by asking these libraries for relevant project literature.\textsuperscript{14}

\footnotesize
\textsuperscript{11} LISA, Information Science Abstracts Plus.
\textsuperscript{12} ERIC, DOBI, Current Cites.
\textsuperscript{13} Mostly Google, but also AltaVista and Dogpile.
\textsuperscript{14} As part of the CLQ.
An unstructured questionnaire – the *CLQ* (*CIPAC* Library Questionnaire) – was sent to all libraries known to operate a *CIPAC* by mid-2001, in order to obtain information on aspects such as the reasons for and details of *CIPAC* implementation, perceived user reactions and details of any monitoring of use, project documentation (see above), policies and future plans. In addition, three extensive interview sessions were held with the creator of the Austrian National Library's *CIPAC*.16

In order to collect information on the search and retrieval capabilities of the various *CIPACs* all catalogues were repeatedly searched on the WWW; for the systems described in section 2.2 a standard type of query was used.

For the study of user reactions a web-based survey of *CIPAC* users was undertaken from February to April 2001. The respondents were recruited by placing buttons with links to the online questionnaire on the web-pages of eleven *CIPACs* in four countries (Austria, Switzerland, Czech Republic, Germany). A semi-structured online questionnaire was used for which versions were available in German, Czech and English. The final out-turn of this survey was 320 completed questionnaires.17

### 1.6 Structure of the dissertation

The dissertation begins with a world-wide survey of existing *CIPACs* (Chapter 2); after a presentation of the author's *CIPAC* web-page the characteristics of four main types of *CIPAC* are analyzed and the features of all fifty *CIPACs* known so far are compared (a comprehensive inventory of these fifty *CIPACs* is included as Appendix A1). In Chapter 3, the main problems and issues involved in the creation and implementation of *CIPACs* are identified and discussed; a case-study of one library's project of further conversion of a *CIPAC* to an *OPAC* is added as Appendix B4. Chapter 4 provides an account of the web-based *CIPAC* user survey and its main results. The study concludes with a review of findings and suggestions for further research (Chapter 5).

---

15 First in December 2000, and subsequently as other *CIPACs* were identified and/or started operation (for details, see Appendix B2; the questionnaire is reproduced as Appendix B1).


17 For details of the methodology applied see section 4.2.