

# *Bibliofile* for Serials Cataloging

Péter Jacsó

**ABSTRACT.** Microcomputers have for years been affordable for librarians. They could choose among much software for cataloging support, but typically the data had to be created by themselves. The Library Corporation pioneered the first CD-ROM product for libraries, the *Bibliofile* database. It offers over four million MARC records from the Library of Congress. Nearly half a million of those records are for serial publications. The records selected can be edited and downloaded for an in-house catalog database in the US-MARC communications format. Catalog cards and spine labels may be printed. This article evaluates and illustrates the database, the procedures of searching, editing, displaying and exporting records. *Bibliofile* gets a high score for price/performance. Recommendations are made to introduce some new features.

---

At the time of writing this review Dr. Jacsó was Visiting Research Scholar at the School of Library and Information Studies of the University of Hawaii. He was working on a project related to CD-ROM databases for serials cataloging and acquisition. He also taught continuing education courses about CD-ROM technology and application at the University of Hawaii, where he will teach again as visiting Associate Professor in the Spring and Summer semesters of 1990. Formerly he was a visiting professor at Rosary College in River Forest, and at Columbia University in New York. He is on leave from his position as head of the Computer Science Library and Information Center in Budapest, Hungary. Péter Jacsó was a consultant for UNIDO and UNESCO in projects designed to computerize library and information services in developing countries. He is member of the editorial advisory boards of *Infomediary*, *Revista Española de Documentación*, and *Information Development*, and is a regular speaker at various international fora for librarians and information specialists.

The author is grateful to Professor Carol Tenopir of the University of Hawaii, who was his scientific advisor, for advice, recommendations and guidance carrying out this project and for reviewing the projects interim and final results and the manuscript of this article. The Library Corporation and Ms. Mary Ann Zimmerman are thanked for making available a review copy of *Bibliofile*.

## INTRODUCTION

*Bibliofile* from The Library Corporation (TLC) represents a landmark application of CD-ROM technology. Back in 1985, it was the first commercially available laserdisk product, and one intended specifically for the library community.

*Bibliofile*, as it first appeared in February, 1985, was a useful product, and since then eleven new, revised versions have been introduced. There are now several CD databases available to support, among other functions, serials cataloging.

In a project sponsored in 1989 by the International Research Exchange (IREX) Program of Princeton, New Jersey and hosted by the School of Library and Information Studies of the University of Hawaii, several such databases were analyzed and tested by this author.

In addition to *Bibliofile*, these included Supercat (Gaylord), Lascercat (WLN), CD-CATSS (Utlas), Ulrich's Plus (Bowker) and The Serials Directory (EBSCO). The producers of these databases made available review copies that were used both on IBM PC/XT-s and PC/AT-s (and compatibles), all of them with 640K of RAM and 30Mbyte of hard disk. The CD players used were the Hitachi 1503 and its equivalent under a different label, the AMDEK 2000. The version of *Bibliofile* reviewed is 4.03.

Several papers, referred to at the end of this article, have been published on *Bibliofile*. These papers typically give a bird's-eye view of the complete product or discuss the management issues of using a CD database for cataloging. This review, on the other hand, will focus on a particular application of *Bibliofile*, and will attempt an in-depth look at that. This area is the cataloging of serials using *Bibliofile*. The article evaluates the scope and coverage of the database, the content and searchability of the records, as well as ways in which the records may be edited, displayed, printed and transferred to a local information storage and retrieval system. It is suggested that much of what is said can be of interest from the standpoint of other serials related functions, such as preorder verification and the creation of check-in records for serials control. Evaluated also are the installation, documentation and support aspects of *Bibliofile*. Though these are not specific to serials applications, they are indis-

pensable for an in-depth evaluation of the system. Indeed, a large part of the evaluation may be of interest to the cataloger of any type of material.

Finally, this review includes several suggestions for bettering a product that is already of outstanding quality and price.

## METHODOLOGY

*Bibliofile* was tested against a real sample, the collection of serials of the Computer Science Library and Information Center in Hungary, a library that is considering converting its card catalog of serials to an online catalog. This library has the necessary hardware (IBM PC/AT compatibles), and software (Micro-ISIS from UNESCO) to create the catalog, but its serials record is not in machine-readable format. (An online monograph catalog has since long been available.)

The collection consists of 423 serial titles. Its profile by country of publication and language is illustrated in Figure 1. (The abbreviations used in the figures are the ISO codes for languages and countries.)

Not surprisingly, many Hungarian titles appear in the sample. The subject specialization of the library indicates that an overwhelming majority of the titles started publication in the past 20 years and remain active. These factors may not make it the most typical of collections, but did not influence the evaluation, except in the case of the hit rate, and they did in effect provide a reasonably realistic setting for the project.

All serials were checked by ISSN, if one was available. If the ISSN were not available or, when available, did not retrieve a record or retrieved a wrong record (because the ISSN of a former title was available in the source file, for example), the serial was searched by title. If this did not retrieve a record, variations of the title, allowing for misspellings or wrong word order or punctuation, were used.

In addition to searching, displaying and printing, the record editing and transfer features of *Bibliofile* were also tested. No attempt was made, however, to create original records, because that would

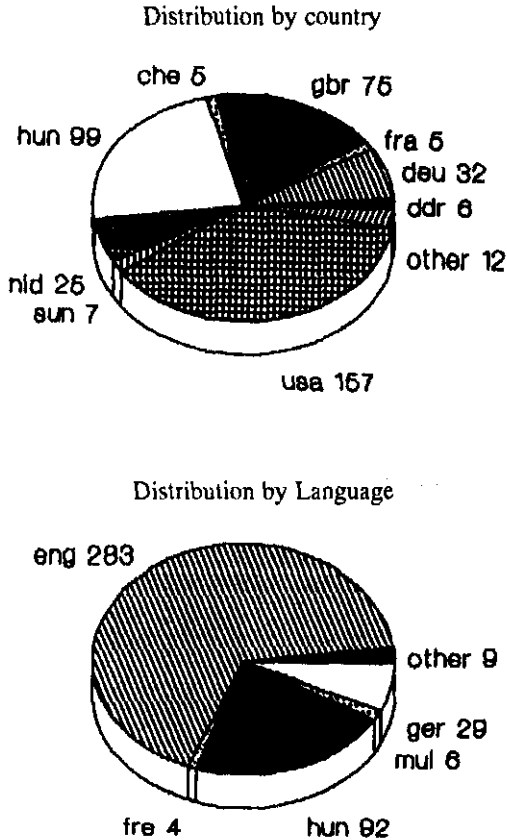


FIGURE 1. Profile of the Collection by Country and Language

be done by the stand-alone information storage and retrieval software, Micro-ISIS.

### **DATABASE SETUP AND CONTENT**

There are over four million records in *Bibliofile*. Out of that, about 500,000 are for serials, although not all are for unique titles, as shall be discussed later. There are various ways of allocating millions of records on CD. Records may be grouped chronologically, by country, by language or by document type, to name a few

options. It is not necessarily the optimum space usage on disk which is the decisive criteria in allocating the records among the disks. Utlas for example, dedicates a separate disk for serials on its CD-CATSS system.

*Bibliofile* basically splits the records into two major groups by language of publication. Three disks contain English language materials, and two others include non-English language materials, regardless of country of publication.

The various indices to the records are stored on the disk with the highest number within the group, Disk #3 for English, and Disk #5 for non-English language materials. (Physically the disks for non-English materials also carry #1 and #2 on their labels, but it is much more unambiguous to refer to these as #4 and #5 disks.) This is a reasonable setup from the viewpoint of the producer, because only the last disk in the group needs to be remastered for updating. For users who wish to perform retrospective conversion of monographs, serials and audiovisual materials, an allocation by type of material may be more appropriate in general, as this allows parallel conversion of different materials at different workstations. The allocation method used by *Bibliofile* favors the serial cataloger who works with a multilingual collection, as the retrospective conversion work may be split into two groups, if there are two workstations, by the language of the serials.

Figure 2 indicates how the hits (matching records) were distributed on the five disks. Clearly, a workstation with two drives can simplify the identification of appropriate serials records.

*Bibliofile* started out with 3 compact disks (2 for English language materials and 1 for others). As new records are added to the database, the number of disks increases. Though the system can be used from a workstation with a single CD-ROM player, a two-drive configuration is strongly recommended. *Bibliofile* in its enhanced version can accept an eight-drive configuration. This kind of CD-ROM setup is called daisy-chaining and should be differentiated from networking, which permits the use of one or more disks from several workstations. Networking actually requires not only another version of *Bibliofile*, but also an extra investment in hardware and systems software. Given the predictable rate of increase of new catalog records added to *Bibliofile*, it is safe to estimate that almost

No hits from Disks #1 and #4

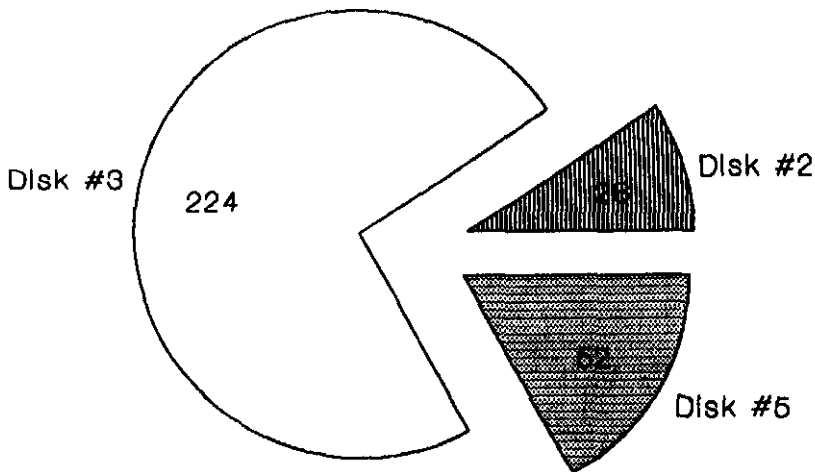


FIGURE 2. Distribution of Hits on the Five Disks

every year the number of disks required to store the complete LC file will be one more, unless an efficient data compression technique is applied. This will be very much needed in the long run. The arrangement of records within the primary grouping of English versus Foreign is not published, though it could help in the optimization of the sequencing and grouping of input. Experience in the project suggests that bibliographic records for monographs and records contributed by the National Library of Canada are stored on Disk #1. One could call TLC for this information, but it would be better if it were available in the manual.

Because of the importance of language as a demarcation line, it might be also worthwhile to mention in the users manual that all records without language codes are to be found on the disks for English language materials. This is particularly relevant for those who catalog sound recordings. The serials cataloger might, for instance, wish to know that many records describing a serial have English as the language code in the control field, even though the title might be thought of as non-English. This is worth remembering

when there is no match for titles beginning with *Zeitschrift, Revista*, etc. on the foreign disks.

Generally speaking, availability of statistical data about the file would be very helpful. Such data would orient the customers as to what they might expect regarding the size of the database and of its distribution by type of material, language and country code, publication year, etc. In the online version of MARC, one can carry out prefixed searches by language, type of document, year of publication, etc. These were meant to limit a search, but can be used on their own to get statistical information. None of the CD-ROM versions of LC MARC offers these data elements as primary search keys. These indices certainly would increase space requirements considerably, but they could be put on the disk when premastering it, if not as access points, at least as static auxiliary information in the form of an introductory help screen.

The records used in *Bibliofile* are MARC-S records, which describe titles cataloged by the Library of Congress since 1965. Popular titles since the turn of this century also have bibliographic records in the database. CD-ROM sources other than *Bibliofile* exist, which feature a higher number of records — and a much higher price tag. These include contributed records from libraries. Wealth of records, however, is not an absolute indicator of value, and it may have disadvantages in the form of unwanted duplicates, triplicates, etc.

*Bibliofile* records are stored in a proprietary internal format, but may be displayed in traditional MARC tagged format and, more importantly, downloaded in the USMARC communications format.

### HIT RATE

One of the most important criteria used in judging the appropriateness of a database for cataloging support is the hit rate, i.e., the percentage of titles for which one or more records can be retrieved. The overall hit rate for the sample was adequate. Out of the 423 titles, bibliographic records could be found for 304, thus representing a 72% hit rate.

This hit rate is much better, an impressive 89%, if Hungarian titles are excluded. The miss rate for Hungarian titles is not of great

concern, as bibliographic data for vernacular serials is fairly easy to obtain locally.

It is a bigger surprise, that for serials from West Germany only a 47% hit rate could be achieved. The hit rate for British periodicals, 80%, was also below expectations. The periodicals involved were well-known, most with a circulation figure of well over 100,000. Figure 3 illustrates the overall hit/miss rate and gives a breakdown of the latter by country of publication.

Only 52 of the 304 queries (17%) retrieved more than one record by ISSN or precise title search. Multiple hits caused by inexact search key, along with non-unique titles without appropriate qualifiers, were not considered. The relatively low rate of multiple hits means that in most cases it is not necessary to wade through several records for the same item, as is typically the case with the online utilities and with the CD-ROM versions of their databases. UTLAS, for instance, produces by ISSN search a high rate (approximately 65%) of multiple hits.

### OVERVIEW OF SYSTEM COMPONENTS AND FUNCTIONS

The major components of *Bibliofile* may be grouped, as illustrated in Figure 4, under three categories; catalog production, system configuration and batch utilities.

The catalog production module of the system deals with three

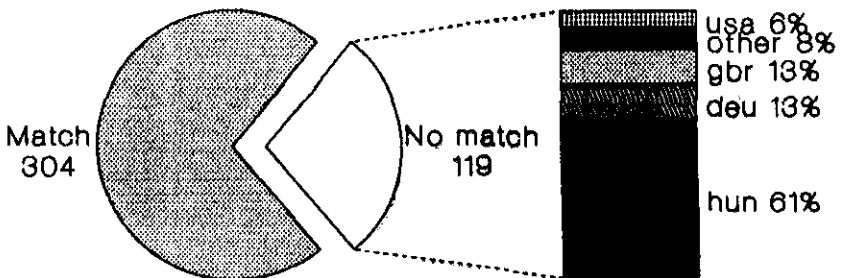
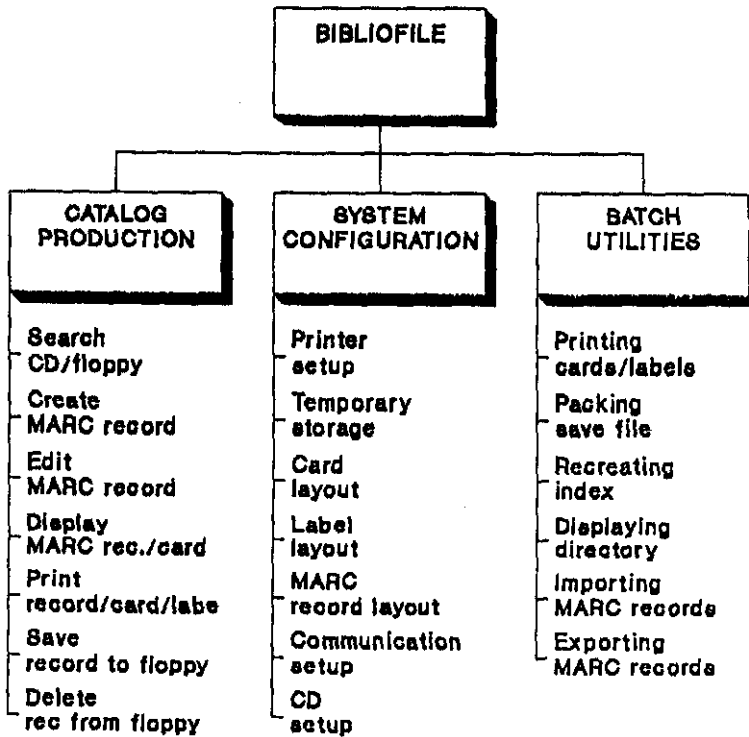


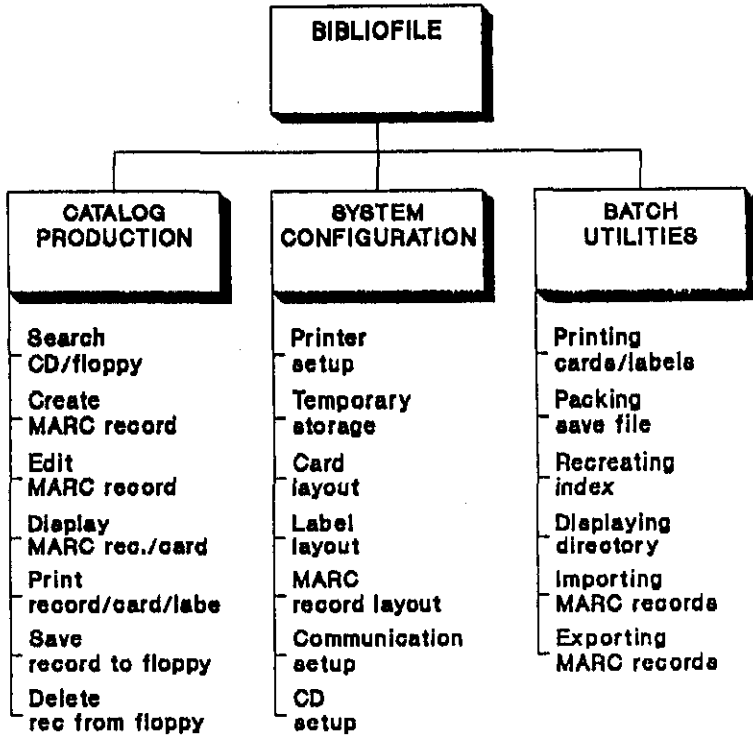
FIGURE 3. Hit/Miss Rate by Country of Publication

FIGURE 4. Components of *Bibliofile*

elements: MARC records, catalog cards and spine and pocket labels. How these elements are used is illustrated in Figure 5.

In addition to the above functions, there are options to change from the English language subset to the foreign language subset of the database and to toggle between the menus of the individual functions. The master menu and the catalog production menu are illustrated in Figures 6 and 7.

This arrangement of the functions in menus is reasonable, well-structured and convenient, but it is felt that the printing of MARC records should be included in the F1-F10 group of functions, and the diskette searching function along with the exit option should be relegated to the ALT-Fx group. This would seem to be a more logical grouping of functions in respect of the typical workflow pattern.

FIGURE 4. Components of *Bibliofile*

elements: MARC records, catalog cards and spine and pocket labels. How these elements are used is illustrated in Figure 5.

In addition to the above functions, there are options to change from the English language subset to the foreign language subset of the database and to toggle between the menus of the individual functions. The master menu and the catalog production menu are illustrated in Figures 6 and 7.

This arrangement of the functions in menus is reasonable, well-structured and convenient, but it is felt that the printing of MARC records should be included in the F1-F10 group of functions, and the diskette searching function along with the exit option should be relegated to the ALT-Fx group. This would seem to be a more logical grouping of functions in respect of the typical workflow pattern.

**BIBLIOFILE MASTER MENU**

Select an application from the menu by pressing the key listed to the left of the function.

**CATALOG PRODUCTION:**

C = Catalog Production  
S = Configure CPS  
U = Batch Utilities

**OTHER APPLICATIONS:**

X = Exit to DOS

**OTHER BIBLIOFILE PROGRAMS:**

E = Switch to English CPS  
F = Switch to Foreign CPS

**CURRENT PATH:**

C:\BIB

---

Select desired application:

FIGURE 6. Options of the Master Menu

**BIBLIOFILE CATALOG PRODUCTION MENU**

Select a function from the menu by pressing the key listed to the left of the function.

**CATALOG PRODUCTION FUNCTIONS**

<b>F1 = Search MARC</b>	<b>F2 = Menu/Exit</b>
<b>F3 = Edit Record</b>	<b>F4 = Display Card</b>
<b>F5 = Input Original</b>	<b>F6 = Print Cards</b>
<b>F7 = Delete Record</b>	<b>F8 = Print Label</b>
<b>F9 = Save Record</b>	<b>F10 = Search Diskette</b>

<b>ALT F2 = Master Menu/Exit</b>	<b>ALT F3 = Print MARC Record</b>
<b>ALT F5 = Insert Control Number</b>	<b>ALT F6 = Print Single Card</b>
<b>ALT F8 = Search Disc Queue</b>	<b>ALT F10 = Display Diskette Directory</b>

For HELP with any of these functions, press <ESC>

---

**SELECT FUNCTION**

FIGURE 7. Menu Options for Catalog Production

2709) format. It would be useful if there were a batch input function to input search keys in a batch from a file as an alternative to the individual entry of search keys from the keyboard. For libraries which have ISSN and/or title information concerning their serials available in machine readable format, this would increase greatly the throughput rate of conversion. There would be irrelevant records in the output file in the case of non-unique search keys, but the clean-up of the resulting hit file of matching records would not take much time. Using batch key entry and retrieval would release the workstation for other tasks.

Utility programs are available with which to compress (pack) the data on the diskette which stores the retrieved MARC records temporarily. It is a well-known problem in data processing that deletion of records does not necessarily make space available for other records. The packing function addresses that problem. Another batch utility function facilitates the recreating of the index file in case its integrity is compromised.

The system configuration function of *Bibliofile* offers an awesome array of options for customizing the format and content of the cards, labels and records to be produced.

The system's workflow is logically organized. Its major phases are outlined in the flowchart given in Appendix 1. Although the flowchart fails to indicate all available options (as in printing a worksheet of the record for editing or or in displaying of the directory of queued items), it provides sufficient information for gaining an overview of the functions and files of *Bibliofile*.

The files which the users refer include: the master file and the index file on CD-ROM; the queue file which holds information derived from the search key to permit automatic retrieval of records not on the same disk as the index file; and the output files which temporarily store the retrieved records along with their indices. The size of the floppy diskette limits the number of records stored in the output file to 500. It is technically possible to store records on the hard disk, but this would require purchase, at extra cost, of *Bibliofile's* Catalog Management System. The limitation, however, is not serious, as one may have several diskettes to store the the retrieved records. Furthermore, the increasing popularity of high density disks will alleviate this problem.

A maximum of 99 items can be stored in the queue for records not located on the index disk itself. This limitation is likely caused by constraints of the random access memory. *Bibliofile* requires a minimum of 512Kbyte RAM, but 640 Kbytes are strongly recommended. One wonders, if with the latter capacity the number of queued items might be increased. Again, this is not a severe limitation until 99 records are retrieved by swapping two disks at maximum, i.e., by replacing the index disk by the #2 and #1 disks to retrieve records from those. If, however, the number of disks keeps increasing, this 99-item limit should be also increased, possibly by some technique which would let a hard disk simulate the RAM buffer for the queue, as a sort of a reverse caching. Such an arrangement would be slow, but less inconvenient than one's acting as a disc jockey.

### SEARCHING

Many CD-ROM databases, let alone online database, permit searching records according to several criteria and by sophisticated combinations of these. In the case of abstracting/indexing and directory CD-ROM databases, this is possible because the number of records in their files is relatively small for the available capacity (over 500 Mbytes). Therefore indexing by all terms in the records is not hedonistic. When several millions of records and their indices must be stored on compact disks, one must be parsimonious in indexing.

Furthermore, a database for catalog production, as opposed to online catalogs and reference databases, has a characteristic search feature commonly referred to as known item searching. This implies that having the item or its appropriate surrogate in hand, one is likely to know the author and/or the title of the item, and may know the unique identifier number of the document, as well as the place and date of publication, the extent of the item and type of material.

*Bibliofile* is meant to prepare a set of catalog cards, labels and/or MARC records for a given item. It is not intended to be used as a tool for subject searching, and its search facilities must be evaluated in this light. On the other hand, one expects in known-item search-

ing to retrieve mostly one and only one record matching the search key.

Primary search keys are:

- International Standard Book Number (ISBN)
- International Standard Serial Number (ISSN)
- Library of Congress Card Number (LCCN)
- Government Printing Office Number (GPO)
- Lead terms of the title
- Author name

The efficiency of retrospective conversion much depends on the quality of the sources used to derive the search keys. If one has a shelflist of serials with LCCN, the retrieving of records is a breeze. Unfortunately, this unique search key is not always available and so other less appropriate access points must be used.

### ***Searching by Standard Numbers***

The standard numbers are the most efficient keys for searching, as they generally provide unique hits, i.e., a single record in response to searches. These keys are, however, not always unique. The same ISSN is used, for example, for the printed and the microfilm edition of a serial. In such cases more than one item may be retrieved by a single ISSN. Several versions of a bibliographic record with the same ISSN are often the unwanted byproduct of cooperative catalogs used to create CD-ROM databases, as *CD-CATSS* of UTLAS or *Laserquest* of GRC.

*Bibliofile* suffers little from this problem because it uses a single source for records, the LC MARC file. This explains why in the present project a very low rate of multiple hits was experienced in ISSN searches. Only 61 ISSNs retrieved more than one record out of the 365 which produced a match at all. Figure 8 illustrates the distribution of retrieval results by number of hits. Most multiple hits by ISSN search were the result of a change in the serial itself, which although the change did not justify the allocation of a new ISSN, may have necessitated creation of a new bibliographic record. In some cases plain errors in a record demanded creation of a new record. An extreme case is illustrated by the *Quarterly Journal of*

---

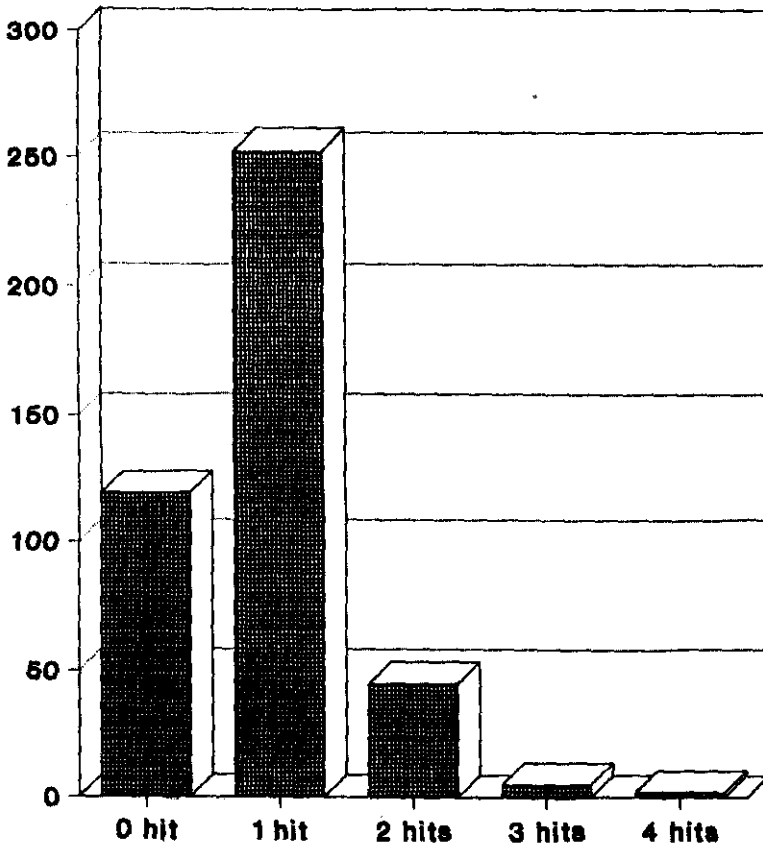


FIGURE 8. Hit Distribution

the Library of Congress, the ISSN of which retrieved 8 different records. One may wonder if only the latest version of each record should be kept, in *Bibliofile*. If all the versions of a record are kept then a simple utility which would facilitate the selection is badly needed.

Many times the hit list of short records provides enough information to select the right record. From the hit list in Figure 13 it was easy to select the record wanted on the basis of the place of publication code. As shown in Figure Figures 10 and 11, however, there

are situations when the hit list does not provide sufficient information for the decision. In such cases, the users typically print out all records which seem to be relevant, then compare them offline. If the records are on separate disks, this a tedious process.

It would be convenient to invoke a utility program which would produce a comparison chart of the tags of records concerned. A sample chart is illustrated in Figure 9. A chart of this sort can accommodate 10 records in this form, but with sideways printing twice as much—a reasonable limit. Many librarians would never download a record which has no LC classification number or subject heading. Looking at the chart would immediately rule out the downloading of many versions, and so spare the time needed to look at the records themselves, let alone the tedium of comparing records field by field to spot minor differences. As this is in large part a quantitative tool, it would not of course reveal such things as a correction in title, for example. That would need the kind of elegant solution found in Gaylord's Supercat CD-ROM database of LC MARC records, which compares two records on the screen and highlights the differences.

A useful feature of *Bibliofile* is that it is not necessary to specify whether one's number is an ISBN, LCCN or an ISSN because the search software recognizes what sort of number is used from its pattern, e.g., two four-character groups separated by a hyphen represent an ISSN. This also means that an ISSN is entered as it is seen. One may take this for granted, until using other CD databases for cataloging, as these require not only an indication that the following number is an ISSN but also the need to prune the hyphen from the ISSN, a requirement making the key more error-prone during entry.

There is a major disadvantage in searching by ISSN. If the search key yields multiple hits, only the ISSN and the identification number of disks which hold the records are displayed (see Figure 10). This is due to the questionable assumption that ISSNs are unique codes and will retrieve only one record. It is important to have other information displayed along with the ISSN in the case of multiple hits, such as the date of record creation, a code to indicate if the record is for the microform version of the serial, etc.

Two minor distractions occur in searching by ISSN. One is that if





1 000417939  
 2 000417939  
 3 000417939  
 4 000417939  
 5 000417939  
 6 000417939  
 7 000417939  
 8 000417939

2 2 2 2 2 2 2 3 3

Disc 3: 0

F1 SEARCH MARC HITCOUNT = 8

0041-7939..... Au Year-Year Place Pages Subject Type

FIGURE 10. Multiple Hit List by ISSN

1	QUARTERLY JOURNAL OF THE LIBRA	1964	US:DC		G 3
2	QUARTERLY JOURNAL OF THE LIBRA	1964	US:DC		G 2
3	QUARTERLY JOURNAL OF THE LIBRA	1964	US:DC		G 2
4	QUARTERLY JOURNAL OF THE LIBRA	1964	US:DC		G 2
5	QUARTERLY JOURNAL OF THE LIBRA	1964	US:DC		G 2
6	QUARTERLY JOURNAL OF THE LIBRA LI	1943	US:DC	Library Sci	G 3
7	QUARTERLY JOURNAL OF THE LIBRA UN	1964	US:DC		G 2
8	QUARTERLY JOURNAL OF THE LIBRA UN	1964	US:DC		G 2

---

Disc 3: 0

F1 SEARCH MARC HITCOUNT = 8  
quarterly journal of the libra Au Year-Year Place Pages Subject Type

FIGURE 11. Multiple Hit List by Title

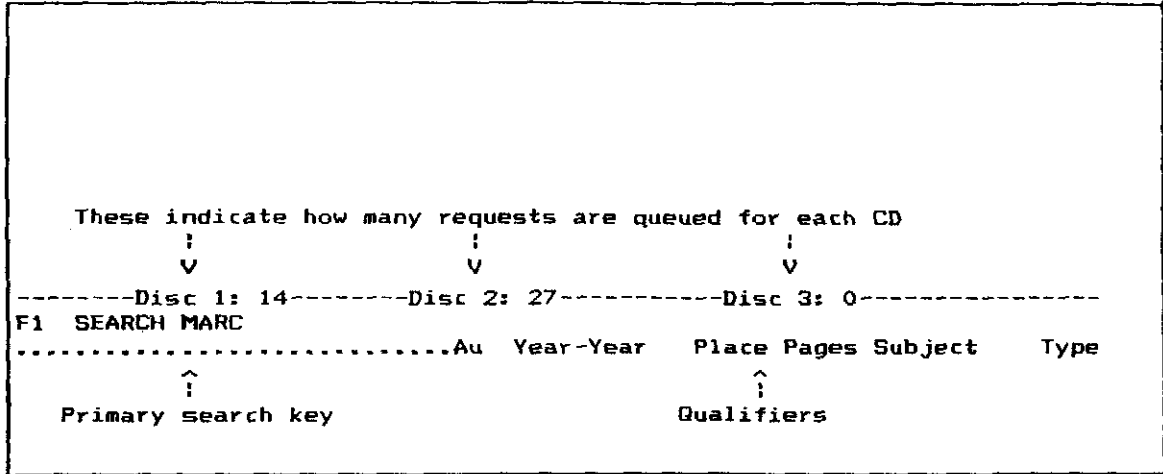


FIGURE 12. Query Template (explanation added)

1	INSIGHT		1985	US:CA	Law	S 2
2	INSIGHT		1985	US:NY	Medicine	S 2
3	INSIGHT		1985	US:NY		S 3
4	INSIGHT		1987	CN:ON		S 3
5	INSIGHT		1987	US:OH		S 3
6	INSIGHT /		1985	US:DC	Periodicals	S 3
7	INSIGHT /		1985	US:MA		S 3
8	INSIGHT /		1986	US:CT		S 3
9	INSIGHT 6		1986	CN:ON		S 3
10	INSIGHT ON COLLECTABLES		1987	CN:ON		S 3
11	INSIGHTS		1987	US:CA		S 3
12	INSIGHTS	LI	1987	US:DC	Library Sci	S 3
13	INSIGHTS 6		1987	US:NJ	Law	S 3
14	INSIGHTS 6		1987	US:NJ		S 3
15	INSIGHTS		1987	CN:MB		S 3
16	INSIGHTS FOR SUCCESS		1987	US:WI		S 3

Disc 3: 0

F1 SEARCH MARC HITCOUNT = 16  
 insight..... Au 1985-1987 Place Pages Subject sype

FIGURE 13. Sample Hit List

the last character of the number, i.e., the control character is an X, it is replaced by a zero in the result. The other is that a leading zero is added to and the hyphen is removed from the ISSN on the hit list. There is no apparent reason for these anomalies.

One improvement might be made in searching by ISSN, apart from doing the search from a batch file. If an incorrect number is entered, the system should warn about this, and so eliminate slavish searching in the index. The control digit as the last character of the ISSN was introduced for computerized checking. It takes no time for the program, and helps the user to realize the reason for the message: Last search returned 0 hit.

Searching by ISSN means searching by the 022, a subfield of the MARC record. ISSNs from former and successor title fields are not retrieved, neither are the invalid and cancelled ones from the 022z and 022y subfields. This is only a problem if using a source which carries one of these ISSNs. It may even be the chief source of information, the title page of the serial itself!

### ***Searching by Title and/or Author***

Searching by standard number is not always possible because many items may not have standard numbers, or if they have, these numbers may not be available to the user. In such cases searching by title and/or author is a useful alternative.

Up to 30 characters of the title may be entered. Definite and indefinite English articles should be omitted. In the case of generic titles, it is recommended that as many characters of the title be entered as are necessary to make the title (almost) unique. This technique works in most cases, but for serials with generic titles, like the *Journal of the American Society of Information Science*, which can only be entered as the "Journal of the American Societ", a large hit list is produced.

Qualifiers, as will be discussed later, may be used, but only the year of publication, if it is known by the user, is appropriate in this case. Title word searching requires word indexing, which is a device not really feasible in such a large database. OCLC-like search key generation from characters of the first, second and later words of the title is not particularly user friendly. A convenient compro-

mise could be the implementation of optional elliptic searching. In case the first 30 characters of the title do not provide a distinctive enough unit, the user may enter a few characters from the beginning and the end of the title with some special separator character in between. The title above may then be entered as

Jour< . . . >Information Science

and would retrieve the appropriate record, and only that record. However, this would certainly increase both storage requirements and response time.

There is an implicit truncation in title searching. Every record with a phrase index (generated from the first thirty characters of the title) that matches the search key entered would be retrieved. This is convenient for titles that although lengthy are distinctive in their first few characters, as

Dr. Dobbs' Journal Of Microcomputer Software

On the other hand, implicit truncation is a lesser bonanza when a title consists of a single word that is also the lead term in many titles, such as Computers. One can zero in on the required title when browsing in the hit list displayed, but it would be preferable to have a special character to indicate that the characters entered represent the whole title, i.e., that truncation should be disabled.

Searching by title means searching by the 245a subfield. It is not possible to retrieve a record by its former title, as for example Junior Libraries, the former title of School Library Journal, as no bibliographic record exists for it. This is not a problem, however, because *Bibliofile* is a cataloging support tool, not a finding tool. The only problem occurs when the source being used indicates a title which is stored as an alternate title in the bibliographic record. It is not possible to retrieve a record by its variant title, not even when that is the one on the cover page of the copy used for cataloging.

As there is no way to recognize automatically whether the term entered is a title or an author's name, a backslash, must be used to indicate that the key is the name of an author. After the backslash a few characters of the title may be entered if it appears that the author search will yield too many hits. A maximum of 26 characters

may be used for the name of the author. A comma may be put between the surname and first name of the author. Capitalization, punctuation and diacritical characters are not needed, but the apostrophe in a name should be used if it is part of the title, e.g., "dr dobbs journal" is not acceptable in place of "dr. dobb's journal".

Author names mean personal and corporate authors' names, the former being an atypical usage in serials. Nevertheless, if the editor is given an added author entry, the record will be retrieved. Strangely, however, one can retrieve the record for Library Journal by Dewey, Melvyl and Bowker, Richard but not by Cutter, Charles or Wessels, Helen even though all the four appear in the added entry-personal name fields of the record.

It is possible that only the first two added entry authors are indexed. This is not a problem with serials, but is very important to know for monograph searching. Corporate author names can be searched under either author or title, if they are main entry authors or the lead terms in the title. The 30-character limit in title searches, however, may be a problem for corporate names with long and non-distinctive leading parts, e.g., a long hit list might be expected on entering "United States Department of . . ."

### ***Qualifying a Search***

Author and title searches may result in an unwieldy hit list. The search can be limited by the

- first two letters of the author's surname
- year or range of years of publication
- exact number of pages of a publication
- type of material

Author's surname and page numbers are irrelevant for serials. The year of publication may be a good criterion, but is not always readily available in the source being consulted in the retrospective cataloging of serials. The type of materials code (S for serials) is an obviously good qualifier, if not always effective. It is of no use in solving the above problem, for example, when the title has a long generic part, as everything starting with "Journal of the American Societ" is a serial implicitly. One should be prepared for such situa-

tions with generic titles such as *Daily News*, *Today*, *El Dia*, *Noticias*. Generic titles may crop up when one least expects them. It is almost past belief that there are 61 serials with the title *Insight*.

In spite of space demands, one or two codes in the serial record should be indexed for use as qualifiers. The country code and the status code used as optional qualifiers could facilitate the narrowing of searches.

It is somewhat confusing that in the query line, the last line of the screen (see Figure 12), the label of the qualifiers must be partially overwritten when entering a code, and in the same line there are labels which may not be used as qualifiers (place and subject). These would better serve the orientation as headers for multiple hit displays on the top of the screen even if this means that one line less would be available for short record display.

At any rate, it would be convenient if the qualifier(s) entered would remain in effect until deleted. Searches of higher precision are guaranteed when the type of material code is automatically supplied with the search key unless otherwise specified by the user, as it is likely that materials of the same type will be searched in a straight in a row.

A nice feature in *Bibliofile* is that the last query can be recalled by the pressing of a single key. One certainly appreciates this when, for instance, the first character in a long search key is misspelled.

## **DISPLAYING AND PRINTING RESULTS**

After entering a search key four system responses are possible:

- “Last search returned 0 hit” message is displayed
- Unique matching record is displayed in MARC image
- Hit list of one-liner mini records are displayed
- “Request queued to Disc x” message is displayed

The underlying concept is that the index of all records is stored on the last disk, i.e., on Disk #3 for English language serials, and on Disk #5 for foreign language items, respectively. This a very clever idea, one copied frequently by other CD databases used in cataloging support. The arrangement allows one to know immedi-

ately when bibliographic records on the CDs match the search key, irrespective of their location.

If no record matches the search key, an appropriate message appears, it is possible that an error was made in entering a search key, this key should be checked. A very important feature of the software is that the search key is retained on the screen when no matching records are found. This is again something that might be taken for granted but it unfortunately does not always come up in some other CD products, which swallow the search key, and say unceremoniously "NO HIT." One might wonder if the search key, which is not visible any more, was misspelled or if there really is no hit for the item.

If the search key is matched by more than one item, a short version of the matching records is displayed. This is the hit list (see Figure 11), and for each record it contains the first 30 characters of the title, the first two characters of the author's surname (typically none for serials), the year of publication (the starting year of serials), the code of the country of publication (and the code of the state in the case of US imprints), the number of pages in the publication (typically none for serials), the broad subject category, the code of the material type, and the identifier of the disk holding the record. All of these items are useful in identifying wanted records. In cases in which multiple hits are obtained by ISSN searching, only the ISSN and the disk identifier are displayed.

The hit list is limited to 99 items, with 22 displayed at one time on the screen (somewhat fewer when the search key is an author name, which is echoed back as a separate line above titles by that author). It is possible to scroll up and down within the 99-item range. Beyond this the search key must be refined. Any record on the hit list may be selected for viewing by entering its hit list number. If the record is on the index disk, it is immediately displayed. If it is on another disk, the request is placed in a queue for that disk.

If only one record matches the search key, the system may respond by displaying either the full record (if this is on the index disk) or by automatically putting the necessary information into the queue for the particular disk. A total of 99 records may be placed in each queue. As there is a running tally of each queue, checking the watermark is easy to do. The disk may be switched at any time and

the records retrieved from the disk for which a request has been queued. Obviously disks will not be swapped for a few records, as it is preferable to wait until the disk is nearly full. It is at the same time reassuring that the queue can be saved even if it has been opened for processing. The only nuisance is that any answer except "Y" to a question as to whether or not the queue should be saved is interpreted as no. For safety, sake every answer except "N" should be interpreted as yes.

*Bibliofile* displays the full record either in a MARC image or in catalog card format, depending on system specifications made when configuring the system. Switching between the two images is just a keystroke away — an elegant solution to the need to see immediately the effect of editing the record. Records and cards can be displayed, spine and pocket labels only printed, but not displayed, though a label template can be displayed in the configuration module of *Bibliofile* when format and arrangement are specified.

It is more than impressive to see how the user can customize the arrangement, formatting and content of the cards and labels for printout. Several choices are available, for example, in such matters as printing the classification number, setting margins, printing certain parts of the card in compressed format, and omitting notes. Thirty pages of the manual are devoted to card and label customization. The imaginative user might very well devise a fake card format that would do for entries in a list of serials when printed on paper instead of on cards.

### **EDITING RECORDS**

The record retrieved is automatically put into an edit mode. This is a convenience because in modifying the records one certainly will want to add such local data as shelflist number, holdings information, barcode number, etc. It is possible also to modify or delete any element of the record.

How the record should be displayed for editing can also be customized. Any field, subfield or indicator can be set to appear automatically in record editing, and default values can be defined for certain subfields. Explanatory labels can be added to the bibliographic information to be printed on the catalog cards. This cus-

tomization also provides some validation possibility. If in the customization process one includes, for example, the indicators of a field, no invalid indicator can be entered. This works of course only with fields where indicators are standard, irrespective of field content.

The texts in front of the fields may even be modified when a record is invoked for editing. (These are only for the orientation of the cataloger, but will not appear on the cards printed as opposed to the default values of the subfields discussed above!) Hence, the English language texts of MARC tags may be translated into some other language. The "read only" nature of the compact disk is an asset because it is always possible to get back to the original LC MARC record, no matter how much damage may have been done to it while editing.

The spoiled user may wish to have a plausibility check built into *Bibliofile*. It would be good, for example, to have *Bibliofile* check to make sure when the status code of a serial is c (for current) that the end date of publication is 9999, that the end date is not smaller than the starting date, that the language and country codes are valid, that the dates are of four characters, etc. Even if one can assume—rightly or not—that LC MARC records are free of such inconsistencies, the availability of plausibility and validity checks would be helpful, as the users themselves may create and modify records.

### **EXPORTING RECORDS**

It is likely that a user will not stop at printing cards and labels, but will want to create a catalog of the records downloaded from *Bibliofile*. Such a catalog would allow more options for searching. The user may even wish to use the downloaded records for creating the item file for a serials check-in system.

Any decent library automation system must of course be able to accept records in the USMARC (ISO 2709) communications format. (It is also desirable for library software to produce records in the same format, but this is the exception rather than the rule, at least with microcomputer-based library programs). An eminent feature of *Bibliofile* is that it can both import and export such records. This is one of the batch functions of the system, one taking its input

from the file containing the records saved on floppy diskettes. Records created and or modified by the user can be saved on floppy diskettes at the push of a button. The only improvement that might be made in the process would be to make it possible to format diskettes on the fly. Converting 400 records can require less than 2 minutes.

*Bibliofile* can also import records in the USMARC communications format. Hence, if one wants to include records from other sources for titles having no match in *Bibliofile* and put these records through the same editing procedures as records retrieved from *Bibliofile*, this is possible.

## **INSTALLATION, DOCUMENTATION AND SUPPORT**

*Bibliofile* is very easy to install. It is one of few programs that can run both with or without Microsoft CD-ROM Extensions, and does not interfere with existing system files. The handling of the CD drive is built into the program. The reason for this is that when TLC introduced *Bibliofile*, there were no device drivers for this particular device and DOS could not handle files larger than 32Mbyte in size. The whole installation is menu driven, based on responses by the user to questions posed by the installation program.

*Bibliofile* can run on a floppy-based system, but a hard-disk arrangement is considered preferable. The minimum required memory is 512K, but 640K is strongly recommended. Interestingly, *Bibliofile* could run with 422Kbyte of RAM remaining, when TSR (terminate and stay resident) programs were made to coexist with *Bibliofile*. If not enough RAM is available, *Bibliofile* sends a cryptic message: No TitleRAM [AuthorRAM] available. A more unambiguous error message would be better. Other error messages are clear, such as that used when a non-*Bibliofile* disk is in the drive, or when no CD disk is mounted. This situation, which is common, sends some CD-ROM programs and one's system into hibernation.

Starting up the program is only a few keystrokes away. Here, however, *Bibliofile* is not as defensive as in the above situation. It requests confirmation of the date and time displayed, but if the ENTER key is pressed insistently instead resorting to the "Y" or "N" keys, the system will display the error message "Invalid date en-

tered, please re-enter." Nothing is accepted, and the application must be restarted. It would be good if wrong answers were intercepted and an appropriate warning issued.

The manual is in looseleaf format, is well organized, and has a very useful index, and, in an appendix, the "MARC Formats for Bibliographic Data."

Some minor reservations arise with regard to the manual. First it is not absolutely up to date. Certain screens reproduced in the manual are not identical to the ones seen in the display. There is a very useful troubleshooting chapter arranged in questions-and-answers style. It would be good, however, to have an alphabetical list of error messages included for quick information.

The manual encourages the user to contact TLC if she/he has a problem. Questions are promptly, courteously and competently answered. Not only is this service free, but so also are telephone calls made within the US.

There is an excellent demonstration program with explanations that are readily acceptable even to diehard technophobes.

## CONCLUSIONS

*Bibliofile* is a gem, and available at a bargain basement price. The yearly licence with quarterly updates is \$1070, and foreign disks, updated once a year, cost an extra \$500. *Bibliofile* is to librarians what Visicalc was to spreadsheet users. *Bibliofile* allows even libraries on the most stringent budget to undertake readily most retrospective conversion projects. Its database provides in general records of high quality, and a good collection of these, even if limited to LC cataloging. The user interface is nothing fancy, but it is menu-driven and efficient. The search engine is tailor-made for use in the known-item searching characteristic of cataloging. Improvement in the precision with which searches are carried out is desirable, but record edition and output facilities are excellent. In terms of customization, it outperforms many other library software packages for cataloging, circulation or serials control that have a price tag three times as high and do not offer as does *Bibliofile*, over 4 million bibliographic records. Installation, documentation and support score high points. *Bibliofile* is kept alive by frequent updat-

ing. It deserves credit for pioneering an application, and has inspired several similar databases.

## REFERENCES

1. Beiser, Karl. *Bibliofile Again*. *Small Computers in Libraries*. June, 1986. p. 24-26.
2. Desmarais, Norman. *Bibliofile for Retrospective Conversion*. *Small Computers in Libraries*. December, 1985. p. 24-28.
3. Kneeland, Marjorie. *Bibliofile is Coming! Help!* *CD-ROM Librarian*. March, 1988 p. 9-11.
4. Morrow, Blaine. *Library Corporation's Bibliofile*. *CD-ROM Librarian*. January, 1988. p. 25-29.

APPENDIX 1. System Diagram

