Microcomputer-based indexing and abstracting

Mary F. Tomaselli

Adjunct Lecturer at the Graduate School of Library and Information Studies at Queens College of the City University of New York; Director (1981-84) on the Board of the American Society of Indexers

A microcomputer-based indexing and abstracting project at the Graduate School of Library and Information Studies at Queens College of the City University of New York helps students gain hands-on experience in document surrogation and data base design and construction. Describes in detail a software package called 'Library Mate' for recording, storing, and retrieving bibliographic citations to a collection of informational items (books, journal articles, etc.).

The microcomputer is becoming an ever more real presence in our every-day business and academic lives. This is true of Queens College Graduate School of Library and Information Studies, which acquired a microcomputer in summer 1982 to provide hands-on experience for students taking its various technical and library services courses.

In the spring semester of 1983 I suggested to my introductory indexing and abstracting class that they might fulfil the course's term-project requirements by participating in a microcomputer-based indexing and abstracting project. Two matriculants for the Master of Library Science (MLS) degree, eager to learn about the microcomputer and to experience real computer-assisted indexing and abstracting, volunteered.

The project involved indexing and abstracting papers submitted by the School's students in completion of their requirements for the MLS degree. The papers, approximately 200 and increasing, include bibliographies with and without annotations, research reports, case studies, surveys, and state-of-the-art theses on library and information studies. They are arranged alphabetically by author in a vertical file. As each paper is added to the collection, a 3" × 5" card is typed indicating author, title, sponsoring instructor, course, and date of submission. These cards, also arranged alphabetically by author, are kept in a group at the front of the file.

Although the collection of student papers is mainly archival, it is often consulted by instructors in the department to ascertain whether and from what point of view a particular subject has already been treated. Our project was proposed and undertaken in order to facilitate this search by subject.

Hardware and software

We used the School's Apple II microprocessor with an 80-column Zenith monitor, and two Apple disk drives that hold 5¼" disks. The software, called *Library Mate*, was purchased from Mason Christner of Geosystems, Inc., Michigan, and is designed for recording, storing and retrieving bibliographic citations to a collection of informational items (books, journal articles, reports, etc.). Data entry consists of 'filling in' prelabelled bibliographic fields on the monitor screen and assigning index terms, or keywords, which may be associated with each other using Boolean operators, 'and', 'or', and 'not' to retrieve selected citations. The records created in data entry can be edited or modified as necessary.

Processing the edited citations creates various files which may all be listed to the screen or printer as output files. The output files include: all or any portion of the citations with assigned keywords in citation number order or alphabetically by author; all or any portion of the citations only; the citations by keyword; a dictionary of all unique keywords used; and two postings files, or indexes: an author index and a keyword, or subject, index.

The *Library Mate* program is a menu-driven, user-friendly package which is easy to use with the aid of its 19-page typed manual. We did, however, experience some problems with the first version we used. After we had entered approximately 40 or 50 records the program began to slow down and, sometimes, wouldn't respond to commands at all. We explained our difficulties to the supplier, who sent us an enhanced version of the software which worked quite efficiently. The information in this article is based on our experience with the enhanced version.

The project

At the start of the project, we created a worksheet (fig. 1) to record the bibliographic data, a brief abstract, and the index terms, or keywords. The students considered it easier to fill in the worksheets rather than enter data directly at the keyboard. They could do all their thinking, composing of the abstract and selection of index terms at home, and then key in the data from the

LS 745 MICROCOMPUTER PROJECT
Mary F. Tonaselli

Indexer _____

WORKSHEET

Citation # _____ ___   Indexing # _____ ___

Author__ ____ ____ ____ _____ __ _ _____

Title__ ____ ____ _____ _____ ___ ____ _____ _____

Instructor (REF) __ _____ _____ ___ _____ ___ _____

Date (DT)___ ____ ___ _____.   Course (PAGE)___ ____ __

Abstract (COM)

__ _____ ____ _____ ____ ____ _____ ____ ___ ____ ____ ___ __

__ ____ ____ ___ _____ ____ ____ ____ ____ ___ ____ ____ ___

___ _____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ___

___ ____ ____ ____ ___ ____ ____ ____ ____ ____ ___ ____ ___

___ ____ ____ ____ ____ _____ ____ ____ ___ ____ ____ ___ __

___ ____ ____ ____ ___ ____ ____ ____ ___ ____ ____ ____ ___

Keywords (Subject index terms including one document form term eg.
        Research paper, etc. and instructor's name)

__ ____ ____ ____ _____ ____ ____ ____ ____ ____ ____ ___

___ ____ ____ _____ ____ ____ ____ ____ ___ ____ ___

___ ____ ____ _____ ____ ____ ____ ____ ____ ____ __

__ ____ ____ _____ _____ ____ _____ ____ ___

__ ____ ____ _____ ____ ____ ____ ____ ____ ____ ____ ___

Figure 1. Worksheet used for the microcomputer-based project.

completed sheets, which would also serve as a hard-copy file of the data entry file.

The worksheets proved invaluable when, at an early point in data entry, we lost 16 citations because of input error or damage to the disk (we never found out which). Instead of having to recreate the data, we simply rekeyed the data from the worksheets.

The tags (REF), (DT), (PAGE), and (COM) assigned to four of the fields on the worksheet need to be explained. *Library Mate* has six fields in which to record bibliographic information: Author, Title, Reference, Date, Page, and Commentary. These basic labels cannot be changed (there is a version that allows the user to define his own tags) and appear on the screen in data entry mode for recording citations. For recording the student papers we modified the use of three of the six fields so that our citations would include: Author, Title, Sponsoring Instructor, Date of Submission, Course, and Abstract. The abbreviations for the original *Library Mate* tags were added to the worksheet to avoid confusion during data entry and editing. These field tags do not print out when citations are output to the screen or printer.

Preparing the abstract and assigning keywords

Each student, to fulfil the course requirement, had to create and enter the bibliographic and indexing information for at least 50 papers. Since the abstract field is limited to 255 characters, students were instructed to prepare succinct, terse abstracts that contained a topic sentence expressing the main idea or thesis of each paper. They were free to assign as many keywords as they deemed appropriate to characterize the subject content of the papers. In addition to subject keywords, they were required to assign one keyword that characterized document type, e.g. bibliographies, etc.; and the instructor's name.

We decided not to use a controlled vocabulary (thesaurus or subject headings list) from which to select indexing terms. The students were encouraged to create terms using the terminology of the papers as a guide, and to use direct-entry terms (Public libraries, *not* Libraries, public) and, when synonymous terms for a concept occurred to them, to consult *Sears list of subject headings* or *Library of Congress subject headings*.

The students were responsible for entering all the raw data, editing and processing the citations, and making copies of the growing bibliographic files each time they completed a session at the terminal. Because no printer was yet available at the Graduate School, we periodically went over to the Queens College Computer Center to run printouts of our output files (figs. 2–4: Keyword index, Dictionary, Citations + keywords).

Retrieving citations

Occasionally, we performed a search of our data files as an experiment or at the request of one of the faculty members. Using the Retrieve mode, the program prompted us through the operation by first asking for Primary Search Key(s) ('or' term), then for Intersect Key(s) ('and' term), Exclusion Key(s) ('not' term), and author name(s) which acts as an 'and' term. No search could be performed unless we entered at least one primary search key, but all subsequent search keys could be skipped by depressing 'Return' on the keyboard.

Library Mate permits the citations that satisfy the search parameters to be output to the screen or the printer in the form of a bibliography. Fig. 5 is a list of citations that answered our search question: Public libraries *or* Public library services.

A useful feature of *Library Mate* software is that with the Retrieve operation you can search as many volumes or data disks as you wish. During Retrieve, the citations that satisfy the search parameters are written to a 'scratch' file on the program disk. At this point the program states how many citations have been found and asks if another disk is to be searched. The program will continue to search disks and write citations to 'scratch' until told that there are no more disks to be searched. The

Figure 2. Keyword index.

Figure 3. Dictionary.

BIBLIOGRAPHY

1. BECKER, DEBORAH. INTEGRATING MEDIA INTO THE PUBLIC LIBRARY COLLECTION. (LOGSDON, LS 757.1) APRIL 1973.
ABSTRACT:PAPER DISCUSSES INADEQUACY OF NON-PRINT COLLECTIONS IN PUBLIC LIBRARIES. COVERS TYPES OF MEDIA, HISTORY, AND PROPOSES A MODEL MEDIA CENTER. PROPOSES MATERIALS SELECTION PRACTICES

2. BECKER, DEBORAH. INTEGRATING MEDIA INTO THE PUBLIC LIBRARY COLLECTION: PROPOSAL. (LOGSDON, LS 757.1) APRIL 1973.
ABSTRACT:PROPOSES TO OUTLINE A MODEL MEDIA CENTER AND COMPARE IT TO THOSE LOCATED IN THE QUEENSBOROUGH AND BROOKLYN PUBLIC LIBRARIES. ILLUSTRATES THE NEED FOR VERSATILE, BROAD MEDIA COLLECTIONS

3. BERTRAND, CYNTHIA. THE AMERICANIZATION OF THE IMMIGRANT: THE ROLE OF THE PUBLIC LIBRARY, 1900-1920. (LOGSDON, LS 790A) JULY 1973.
ABSTRACT:PAPER EMPHASIZES A CONDESCENDING AND PATERNAL AMERICAN PUBLIC LIBRARY RESPONSE TO IMMIGRANTS. EXAMINES THE PUBLIC LIBRARY'S ATTEMPTS TO INSPIRE NEWCOMERS TO AMERICAN IDEALS

4. FRANKLIN, MARLENE. ASCERTAINING WHAT PUBLIC LIBRARY SERVICES ARE MOST RELEVANT TO BLACK CHILDREN IN URBAN AREAS. (LOGSDON, LS 751.1) AUGUST, 1973.
ABSTRACT:PAPER ATTEMPTS TO DETERMINE THE RELEVANCE OF PUBLIC LIBRARY SERVICES TO BLACK DISADVANTAGED CHILDREN BY STUDYING LIBRARY-GO-ROUND AND TELL-A-TALE MOBILE SERVICES ESTABLISHED BY QUEENSBOROUGH PUBLIC LIBRARY IN JAMAICA

5. HOENIG, HELEN. CAN THE PUBLIC LIBRARY SERVE EVERYONE?. (LOGSDON, LS 757.1) MAY, 1973.
ABSTRACT:PAPER REVIEWS TRENDS IN THE HISTORY OF PUBLIC LIBRARIES IN ORDER TO DETERMINE THEIR CAPABILITY FOR PROVIDING A BROAD RANGE OF SERVICES

6. MCARDLE, CONNIE. EFFECTIVE USE OF PUBLIC RELATIONS AND PUBLICITY IN PUBLIC LIBRARIES. (LOGSDON, LS 757.1) FEBRUARY, 1974.
ABSTRACT:PAPER PRESENTS THE EFFECTIVENESS OF PUBLICITY AND STAFF INVOLVEMENT IN PUBLIC RELATIONS. FOCUSES ON THE ROLE OF MEDIA IN PUBLIC RELATIONS. DISCUSSES IMPLICATIONS FOR FUNDING

Figure 5. Bibliography.

4. ANDUJAR, MARTHA C. OUTLINE OF A PROPOSED PROJECT: RACISM AND THE BLACK FAMILY. (LOGSDON, LS 757.1) MAY 1973.
ABSTRACT:AUTHOR PROPOSES TO STUDY WHITE RACISM AND ITS EFFECTS ON BLACK FAMILIES IN URBAN CENTERS OF THE UNITED STATES. TERMS ARE DEFINED AS IS THE SCOPE OF THE STUDY. THE NEED FOR THE STUDY IS EXPLAINED
1 PROPOSALS
2 RACISM
3 BLACK FAMILIES
4 URBAN SOCIETIES
5 LOGSDON

2. ARANOV, MILDRED. THE CHANGING ENVIRONMENT IN HEALTH SCIENCES EDUCATION.... (COLBY, LS 790A) SPRING 1978.
ABSTRACT:PAPER DISCUSSES ROLE OF HEALTH SCIENCE LIBRARIANS IN RELATION TO COMMUNITY MEDICINE. DESCRIBES HISTORY OF MEDICAL EDUCATION AND GROWTH OF MEDICAL LIBRARIES IN RESPONSE TO CHANGING NEEDS OF THE PROFESSION AND THE PUBLIC
1 HEALTH SCIENCE EDUCATION
2 HEALTH SCIENCE LIBRARIES
3 COMMUNITY MEDICINE
4 CURRICULA & PROGRAMS
5 STANDARDS
6 CONTINUING EDUCATION
7 FLEXNER REPORT
8 MEDICAL EDUCATION
9 COLBY

3. ARONS, GERALDINE. BENEFACTOR INFLUENCE ON POLICY & SERVICES OF LARGE RESEARCH LIBRARIES. (LOGSDON, LS 757.1) AUGUST 1973.
ABSTRACT:PAPER STUDIES EFFECTS OF BENEFACTORS' DONATIONS TO RESEARCH LIBRARIES, AND HOW THIS AFFECTS LIBRARY POLICIES & LIBRARY USERS. SHOWS THAT THE NATURE OF GIFTS OFTEN PREDETERMINES THEIR USE
1 GIFTS
2 RESEARCH LIBRARIES
3 BENEFACTORS
4 LIBRARY POLICIES
5 LIBRARY SERVICES
6 NEW YORK PUBLIC LIBRARY
7 COLUMBIA UNIV. LIBRARY
8 LOGSDON

Figure 4. Citations + keywords.

listing of citations to 'scratch' is limited by the space available on the program disk.

Limitations of the software

We found the 255-character length for the abstract field a distinct limitation, and felt that a field length of at least 500 characters would be required to record abstracts that would truly represent the papers in the collection. Although we were aware that we could combine two or three fields to increase the length of the abstract, we did not want to 'lose' fields in which to record other important bibliographic data.

At present, *Library Mate* is available in upper-case only and editing capabilities are limited. In the Editing mode, corrections and changes must be made by rekeying the entire field, which we found time-consuming. However, a new *Library Mate* version should soon be available with upper and lower-case and text-editing features.

Through our experience we realized that we could not use the comma (,) in any keyword. When we wanted to add a personal name, we had to enter it as: Smith (John)—a rather unusual method of entry! Also, we could not underline keywords. Book or journal titles had to be distinguished from other keywords by an identifying expression following in parentheses, e.g. Black Boy (book title).

The files that can be created using *Library Mate* are limited by the volume capacity of the disk. The volume capacity is a function of the maximum citation length (which the user specifies when the disk is initialized— initial values assigned for the program) and an average number of keywords per citation. For our project, we specified 550 characters per citation, which permitted approximately 150 citations per disk. At present there is no capability to merge files on separate disks. As a result, a separate dictionary, author index, and keyword index must be consulted for each volume or disk of citations. However, provision of ability to merge such files will be considered for a future version of *Library Mate*.

We wished occasionally during the project to consult the dictionary or the keyword index. As these files grew to 200/300 listings we found it too time-consuming to start at the top of the files and space down to the particular area being sought. *Library Mate* at present does not permit users to look at a particular *portion* of the dictionary or keyword index, but requires that users start at the 'A's and space down through the file until they find the required area.

Value of the micro-based indexing and abstracting project

We concluded that the project did provide what we had hoped it would:

(1) hands-on experience in indexing and abstracting for a computer-based system. All procedures and tasks involved in document surrogation for large information storage and retrieval systems (minisystems and main frames) were touched on, if on a smaller and more limited scale;

(2) experience with designing data base record formats and output formats for printed products;

(3) familiarity with the capabilities, functions, and limitations of microsystems and micro-based software.

The Graduate School Director, Dr Richard Hyman, and myself agreed to continue the project into the following University session. Further, in an effort to give experience of computer-based indexing and abstracting to a greater number of students, we decided to incorporate several demonstrations of the system into the indexing course itself.

We are considering allowing any students in the course to complete their term-project requirement by creating their own bibliographic data base. These items could be books, journal articles, reports, films, or any kind of print or nonprint media in a specific discipline. The students would be expected to design the record format, and the output formats; enter, edit and process the bibliographic data; assign indexing terms; and maintain consistency by consulting the dictionary as it is built. Future projects could include comparing software packages similar to *Library Mate*, such as the PFS series and *Quick-Search Librarian*, a product of Interactive Microware, Inc.

There is a new version now of *Library Mate* software, intended to improve performance and speed, and further versions are planned. I told Geosystems of the software limitations we had found, and was assured that such modifications as merging files on separate disks, and providing access to portions of the dictionary and keyword index on command, would be relatively easy to make.

Another routine, File Utilities, was being added to the *Library Mate* package, and would enable the user to perform several new record and file management tasks.

Nice entry if you can find it . . .

The advanced user guide for the BBC micro by Bray, Dickens and Holmes (Cambridge Microcomputer Centre, 1983) shows a fine disregard for the working principles of index-use. The 497-pp. text has a full, 11-pp. index, with page numbers aligned in a column on the right of each page. Unfortunately, the volume is spiral-bound, and the holes punched for the wires have removed most of the page references on each verso page; a most tantalizing result.