

Information Technology and Libraries

March 1988

CONTENTS

- 3 Editorial: Standing on the Shoulders
- 7 So That's a Book . . . Advancing Technology and the Library *Raymond DeBuse*
- 19 On the Preservation of Human- and Machine-Readable Records *John C. Mallinson*
- 24 Measures of Expected Online Catalog Performance for Public Access Terminals *Raymond G. Taylor, Jr.*
- 30 Developing and Using the Online Catalog to Teach Critical Thinking *Joan Bechtel*
- 41 Laser Printing for a Variety of Library Applications *Glen J. Kelly*
- 51 Online Maintenance Features of Authority Files: Survey of Vendors and In-House Systems *Agnes M. Grady*
- 56 Bibliographic Data Management with dBASE: A Study of Secondary Key Retrieval on Multivalued Data Items *Richard Pollard*
- 67 Conventional and Knowledge-Based Information Retrieval with Prolog *William Leigh and Noemi Paz*
- 73 Library Automation with Workstations: Using Apple Macintoshes in a Special Library *Edward J. Valauskas*
- 79 Reports and Working Papers
- 79 Back to the Concept: Perspectives on Series Authorities *Mary Dabney Wilson*
- 83 Library of Congress Comment and Call for Response *Judith G. Fenly*
- 85 News and Announcements
- 91 Recent Publications
- 91 Book Reviews
- 91 Dewey, Patrick R. *101 Software Packages to Use in Your Library: Descriptions, Evaluations, and Practical Advice*, reviewed by Ronald A. Gagnon
- 91 Walton, Robert A., and Nancy Taylor. *Directory of Microcomputer Software for Libraries*, reviewed by Ronald A. Gagnon
- 92 *European Conference on Library Automation, 11-12 September, 1986, Harrogate, England*, reviewed by Tom Smith
- 93 Mandel, Carol A. *Multiple Thesauri in Online Library Bibliographic Systems: A Report Prepared for Library of Congress Processing Services*, reviewed by Karen Markey
- 95 Software Reviews
- 91 *Catalog Carder*, reviewed by Patricia A. Etter
- 91 *Librarian's Helper*, reviewed by Patricia A. Etter
- 98 Other Recent Receipts
- 101 Letters
- 103 Instructions to Authors
- 102 Index to Advertisers

BLACKWELL AUTHORITY CONTROL

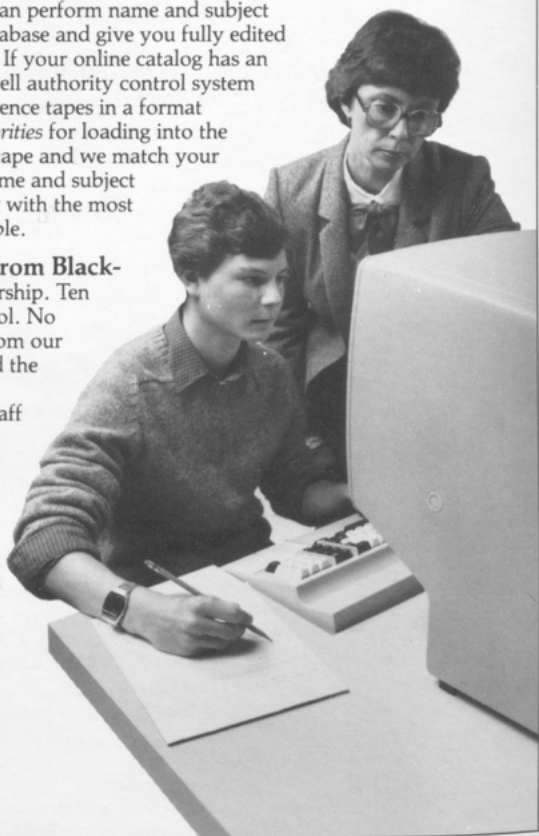
Improving patron access to your online catalog.

Loading MARC records into an online catalog? The Blackwell automated authority control system can perform name and subject authority control on your MARC database and give you fully edited catalog tapes in the MARC II format. If your online catalog has an authority control module, the Blackwell authority control system can also provide deblinded cross reference tapes in a format compatible with *LC MARC for Authorities* for loading into the system. Simply send us your catalog tape and we match your MARC records against current LC name and subject authority files to provide your library with the most up-to-date consistent headings available.

Why choose authority control from Blackwell? Quality. Sophistication. Leadership. Ten years experience with authority control. No other system equals ours. Drawing from our constantly updated authority files and the skills of our experienced editors, this unique Blackwell service frees your staff from retrospective authority control tasks.

For more information, write or call toll free, Dan Miller, Manager, Sales and Service, Technical Services Division.

**BLACKWELL
NORTH AMERICA** INC.
6024 S.W. Jean Road, Building G
Lake Oswego, OR 97034
Telephone (800) 547-6426



Information Technology and Libraries

Volume 7, Number 1: March 1988

CONTENTS

- | | | |
|-----|---|------------------------------------|
| 3 | Editorial: Standing on the Shoulders | |
| 7 | So That's a Book . . . Advancing Technology and the Library | <i>Raymond DeBuse</i> |
| 19 | On the Preservation of Human- and Machine-Readable Records | <i>John C. Mallinson</i> |
| 24 | Measures of Expected Online Catalog Performance for Public Access Terminals | <i>Raymond G. Taylor, Jr.</i> |
| 30 | Developing and Using the Online Catalog to Teach Critical Thinking | <i>Joan Bechtel</i> |
| 41 | Laser Printing for a Variety of Library Applications | <i>Glen J. Kelly</i> |
| 51 | Online Maintenance Features of Authority Files: Survey of Vendors and In-House Systems | <i>Agnes M. Grady</i> |
| 56 | Bibliographic Data Management with dBASE: A Study of Secondary Key Retrieval on Multivalued Data Items | <i>Richard Pollard</i> |
| 67 | Conventional and Knowledge-Based Information Retrieval with Prolog | <i>William Leigh and Noemi Paz</i> |
| 73 | Library Automation with Workstations: Using Apple Macintoshes in a Special Library | <i>Edward J. Valauskas</i> |
| 79 | Reports and Working Papers | |
| 79 | Back to the Concept: Perspectives on Series Authorities | <i>Mary Dabney Wilson</i> |
| 83 | Library of Congress Comment and Call for Response | <i>Judith G. Fenly</i> |
| 85 | News and Announcements | |
| 91 | Recent Publications | |
| 91 | Book Reviews | |
| 91 | Dewey, Patrick R. <i>101 Software Packages to Use in Your Library: Descriptions, Evaluations, and Practical Advice</i> , reviewed by Ronald A. Gagnon | |
| 91 | Walton, Robert A., and Nancy Taylor. <i>Directory of Microcomputer Software for Libraries</i> , reviewed by Ronald A. Gagnon | |
| 92 | <i>European Conference on Library Automation, 11-12 September, 1986, Harrogate, England</i> , reviewed by Tom Smith | |
| 93 | Mandel, Carol A. <i>Multiple Thesauri in Online Library Bibliographic Systems: A Report Prepared for Library of Congress Processing Services</i> , reviewed by Karen Markey | |
| 95 | Software Reviews | |
| 91 | <i>Catalog Carder</i> , reviewed by Patricia A. Etter | |
| 91 | <i>Librarian's Helper</i> , reviewed by Patricia A. Etter | |
| 98 | Other Recent Receipts | |
| 101 | Letters | |
| 103 | Instructions to Authors | |
| 102 | Index to Advertisers | |

Information Technology and Libraries

William Gray Potter, Editor, Arizona State University, Tempe
Pamela P. Brown, Managing Editor, Arlington Heights Memorial Library, Illinois
Karin A. Trainer, Book Review Editor, Yale University, New Haven, Connecticut
David T. Buxton, Software Review Editor, University of Arizona, Tucson
Kurt Murphy, Advertising Editor, Arizona State University, Tempe

Editorial Board:

Jane Burke, NOTIS, Evanston, Illinois
Dorothy Gregor, University of California, San Diego
Susan Harrison, New York Public Library, New York
Tom Leonhardt, University of the Pacific, Stockton, California
David R. McDonald, University of Michigan, Ann Arbor
Linda Miller, Library of Congress, Washington, D.C.
Carol A. Parkhurst, University of Nevada Library, Reno
Paul Evans Peters, New York Public Library, New York
Don L. Tolliver, Kent State University, Kent, Ohio

Information Technology and Libraries is the official publication of the Library and Information Technology Association, a division of the American Library Association, 50 E. Huron St., Chicago, IL 60611; *Executive Director*: Linda J. Knutson. The journal is issued quarterly in March, June, September, and December.

Information Technology and Libraries publishes material related to all aspects of library and information technology. Some specific topics of interest are: Automated Bibliographic Control, AV Techniques, Communications Technology, Cable Systems, Computerized Information Processing, Data Management, Facsimile Applications, File Organization, Legal and Regulatory Matters, Library Networks, Storage and Retrieval Systems, Systems Analysis, and Video Technologies. *ITAL* welcomes unsolicited manuscripts. Submissions should follow the guidelines stated under "Instructions to Authors" on page 103 of the March 1988 issue.

Manuscripts of articles, communications, and news items should be addressed to the editor: William Gray Potter, *Information Technology and Libraries*, Hayden Library, Arizona State University, Tempe, AZ 85287. Copies of books for review should be addressed to: Karin A. Trainer, *ITAL Book Reviews*, Yale University Library, 120 High St., P.O. Box 1603A, Yale Station, New Haven, CT 06520. Copies of software for review should be addressed to: David T. Buxton, *ITAL Software Reviews*, A349 Main Library, University of Arizona, Tucson, AZ 85721. Advertising arrangements should be made with Kurt Murphy, Hayden Library, Arizona State University, Tempe, AZ 85287.

Information Technology and Libraries is a prerequisite of membership in the Library and Information Technology Association. Subscription price, \$17.50, is included in membership dues. Nonmembers may subscribe for \$35 per year in U.S., Canada, Mexico, Spain, and other PUAS countries; \$40 in other foreign countries. Single copies, \$9. Second-class postage paid at Chicago, Illinois, and at additional mailing offices. *Postmaster*: Send address changes to *Information Technology and Libraries*, 50 E. Huron St., Chicago, IL 60611.

Circulation and Production: Central Production Unit (Eileen Mahoney, Director of Production and Distribution Services; Chere Elliott, CPU Supervisor; Sara Ringle Drummond, Debra M. Hall, Eloise L. Kinney, Wanda Mannery, and Paula Sedor), American Library Association, 50 E. Huron St., Chicago, IL 60611.

Publication of material in *Information Technology and Libraries* does not constitute official endorsement by the Library and Information Technology Association or the American Library Association.

Abstracted in *Computer & Information Systems*, *Computing Reviews*, *Information Science Abstracts*, *Library & Information Science Abstracts*, *Referativnyi Zhurnal*, *Nauchnaya i Tekhnicheskaya Informatsiya*, *Otdelnyi Vypusk*, and *Science Abstracts Publications*. Indexed in *CompuMath Citation Index*, *Computer Contents*, *Computer Literature Index*, *Current Contents/Health Services Administration*, *Current Contents/Social & Behavioral Sciences*, *Current Index to Journals in Education*, *Education*, *Library Literature*, *Magazine Index*, *NewSearch*, and *Social Sciences Citation Index*. Microfilm copies available to subscribers from University Microfilms, Ann Arbor, Michigan.

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984. ∞

Copyright © 1988 American Library Association. All material in this journal subject to copyright by ALA may be photocopied for the noncommercial purpose of scientific or educational advancement granted by Sections 107 and 108 of the Copyright Revision Act of 1976. For other reprinting, photocopying, or translating, address requests to the ALA Office of Rights and Permissions.

Editorial: Standing on the Shoulders

This issue marks the twentieth anniversary of the journal. The first issue of the *Journal of Library Automation (JOLA)*, the divisional journal of the Information Science and Automation Division (ISAD) of ALA, appeared in March 1968. Frederick Kilgour was the editor, and that issue included the following articles:

- "Computer Based Acquisitions System at Texas A&I University" by Ned C. Morris
- "A Book Catalog at Stanford" by Richard D. Johnson
- "Brown University Library Fund Accounting System" by Robert Wedgeworth
- "Comparative Costs of Converting Shelflist Records to Machine-Readable Form" by Richard E. Chapin and Dale H. Pretzer
- "The Development and Administration of Automated Systems in Academic Libraries" by Richard de Gennaro

Sitting on the board of editors were Susan Artandi, Joseph Becker, Verner W. Clapp, A. J. Goldwyn, Frederick H. Ruecking, and Eleanor M. Kilgour, assistant editor. (It is fascinating to contemplate this august group meeting to plan the journal in the middle of the "summer of love" in San Francisco at the 1967 ALA Annual Conference.)

Noteworthy advertisements in the March 1968 issue included a prepublication announcement for the *Encyclopedia of Library and Information Science*. The first volume was to be published in July 1968, and the set was to be completed in twelve volumes. Another ad announced that ISAD would be holding a series of ten regional institutes on LC MARC, which was then in the "experimental period."

ISAD, of course, became the Library and Information Technology Association in 1978, and *JOLA* changed its name to *Information Technology and Libraries* in 1982, a move that won the "Snake-in-the-Grass Award" for the worst title change of the year from the Serials Section of the Resources and Technical Services Division.

The articles in the March 1968 issue were more detailed, more specific, and perhaps more optimistic than the articles in the March 1988 issue that you are now holding. It is easy to feel smug about our current capabilities and accomplishments when we read about projects built around an IBM 1401 computer with 12K of main memory, four tape drives, and a 100-character printer. The work performed twenty years ago, however, forms the basis for all that we do today. Further, limitations of hardware did not limit the perception and understanding of *JOLA* contributors in 1968. Consider the following excerpt:

There has never been any tradition for research and development work in libraries—they were considered exclusively service and operational institutions. The advent of the new technology is forcing a change in this traditional attitude in some of the larger and more innovative libraries which are doing some research and a good deal of development. It is worth noting that a concomitant of research and development is a certain amount of risk but that, while there is no such thing as change without risk, standing pat is also a gamble. Not every idea will succeed and we must learn to accept failures, but the

experiments must be conducted so as to minimize the effect of failure on actual library operations. — Richard de Gennaro, *JOLA* (March 1968, p.87).

This is certainly no less true today than it was twenty years ago.

It is rewarding to look back at the articles published in 1968 and reflect upon how far we have come in the use of advanced technology in libraries and how far we still have to go. During the coming year, selected quotes from 1968 will be interspersed throughout these pages to give us some flavor of the past and some perspective on our own time. Any comments from the founding parents would be most welcome.

WGP

CD-ROM
PLUS
ONLINE ACCESS
AT NO ADDITIONAL COST!

WILSONDISC™

WILSONDISC OFFERS

■ CD-ROM Access

Search each database on a separate compact disc. And, with the touch of a button, you can update your search online—at no additional charge!

■ Online Access

For access to the most current data, subscribers to each WILSONDISC database receive unlimited online search time in that database. You pay only the nominal hourly telecommunications rate—no additional charge!

Four Search Modes for Users at All Levels

1 BROWSE Mode

The same access points as the familiar Wilson printed indexes.

2 WILSEARCH® Mode

Easy-to-follow menu options simplify search formulation.

3 WILSONLINE® Mode

The same commands available with the WILSONLINE online retrieval system.

4 EXPERT Mode

Expanded screen handling, windowing functions, and more.

WILSONDISC Databases

Each database is available as a separate compact disc updated and cumulated quarterly.

- Applied Science & Technology Index
- Art Index
- Biography Index
- Business Periodicals Index
- Cumulative Book Index
- Education Index
- General Science Index
- Humanities Index
- Index to Legal Periodicals
- Library Literature
- MLA International Bibliography
- Readers' Guide Abstracts
- Readers' Guide to Periodical Literature
- Social Sciences Index

Unlimited Online Searching with Your WILSONDISC Subscription

Subscribers to a WILSONDISC database can also take advantage of unlimited online searching in that database for the entire term of their subscription. The only additional fee is the nominal WILSONDISC telecommunications charge.

SPECIAL OFFER! Demonstration Disc Available

The specially-priced WILSONDISC Demonstration Disc contains six months of data from 16 Wilson databases, and allows all of the CD-ROM search capabilities available with WILSONDISC. The Demonstration Disc is available at \$99, deductible from your first invoice for a regular WILSONDISC database subscription.

To Order Call Toll-Free: 1-800-367-6770



In New York State call 1-800-462-6060; in Canada, call collect 212-588-8400.

THE H.W. WILSON COMPANY
950 University Ave., Bronx, NY 10452





HOW TO TAME YOUR LIBRARY.

Even the most efficient library can be a real jungle.

But no more. Data Trek can help you cut through the tangle of managerial and clerical drudgery with high-performance, fully-supported microcomputer automation software.

Regardless of your library's size, specialty or budget, Data Trek can provide everything you need to streamline your operation. Like microcomputer and LAN hardware. Comprehensive training. Technical expertise. And outstanding online support.

Data Trek makes automation very affordable, too, thanks to a unique software design which allows you to put individual modules to work, or combine them into a fully integrated system.

Still not convinced? Then consider this: our state-of-the-art Card Datalog™ System is being used by more special librarians

than any other microcomputer system in the world. And, right now, you can get a preview of exactly what it will do for you.

Just fill out the coupon below and mail it in. In return, we'll send you a free Card Datalog Sampler so that you can take your own self-running, guided tour.

If greater efficiency and productivity are what you're hunting for, check out Data Trek.

We think you'll find the difference anything but tame.

 **data trek**

We wrote the book on library automation.

621 Second Street/Encinitas, California 92024

Telephone 619/436-5055

©1987 Data Trek, Inc.



Explore The Data Trek Demo

Free Card Datalog Sampler

For your Card Datalog Sampler, fill out the coupon, attach it to your letterhead, and return to Data Trek, 621 2nd St., Encinitas, CA 92024, ATTN: Tom Beaudin. Note: IBM PC, XT or 100% compatible hardware with minimum 512K RAM required.

NAME _____

TITLE _____

LIBRARY NAME _____

ADDRESS _____

CITY _____ STATE _____ ZIP _____

() _____

PHONE _____ ITQY _____

So That's a Book . . . Advancing Technology and the Library

Raymond DeBuse

Picture yourself in any large American city 75 to 100 years ago: for most people, the electric streetcar was the key to movement around the city beyond their own neighborhoods. Street railways were big business. From the rider's standpoint the streetcar required no capital investment and almost no training or special skills and generally went where one needed to go. Cost was per trip, and a low cost it was, too. The transit entrepreneurs, on the other hand, incurred tremendous expense building their systems. Yet once tracks were laid, they could be used for decade after decade. The investment looked to be a good one.

Enter the automobile. It was only going to replace the horse and carriage, of course. At least, that's what most of its early advocates thought. It was obviously unsuited for the masses: it required a sizable capital investment, fairly high-priced operation, and special training. Yet in retrospect we can see that it met a deep-seated need: individuals could assume greater control of the responsibility for their own transportation. It became the "personal (transportation) system" of the day. The automobile went on to redefine much of our society.

We know what happened to the streetcar and other forms of mass transportation: much reduced in extent, they are massively subsidized by government. In most cities today the automobile is the major means of personal transportation, and mass transit is

tolerated as a necessary evil utilized by a minority of residents. Changes in technology reduced a vital and proud industry to a beggarly one.

Could libraries possibly suffer such a fall? Someone asks this question every time a new information technology appears: a brief debate ensues, and then it is forgotten as libraries assimilate the new technology into their own array of tools and services. But might the fancied "library of the month" on CD-ROM be different? Could those who believe in library doom be correct this time? Not to worry, you say; books are here to stay. One would be foolish to dispute that proposition, but does the existence of the book necessarily mean that the library will remain a healthy and vital institution? Consider the following:

It's the year 2002. Virtually all homes have a computer with an optical reader of some sort that will read—and write—a variety of optical media. Some people have even an Apple Dynabook. ("Dynabook" is the name Alan Kay a number of years ago gave to a hypothesized, book-sized computer and display device that would be used to interact with textual and graphic materials stored on a compact medium that could be inserted into the device.) Publishers have issued most of the materials that might be used by the average elementary or high school student on an optical medium . . . at a relatively low price. Those sit on the shelf next to the computer. The student does her or his research for term papers and the like at the computer. It is connected to a

printer on which the student prints out homework, etc. The printer is being used less and less, however, because increasingly the student's work is being submitted on magnetic or optical cards, written by the computer. If a publication is needed that is not in the student's personal collection, she or he dials up a local number to get into an information utility from which the desired text can be downloaded over a highspeed connection. Or if its need is not urgent, it will be mailed. All for a fee, of course. This same link can also be used to connect with the computers of friends for electronic mail or interactive conversation, so it is not only a work station but has some of the social aspects of the library as well. The student almost never goes to the library.¹

One of the keystones of the public library, particularly for middle-class families, is its perceived value to the education of youth. What happens to support for the library if youth no longer appear to use or even need it? What happens to the support if few of us need the library for how-to books or for answering other information needs that come up in the course of everyday living?

A simplistic view, to be sure, but not totally unrealistic. Yet I suspect that more of us are sceptics than believers. We've heard them cry wolf so often. But consider this: when Federal Express' next-day-delivery concept was first proposed several years back (before the demise of REA Express, once, as the Railway Express Agency, the premier intercity parcel delivery service in the nation) very few people in the industry believed that it could ever work. They openly predicted failure. Now use of that concept is the only way to succeed in the intercity delivery business. The critics had underestimated the capability of new technologies: air transport and computers. They saw the former only as an extension of the older transportation technologies, not as the means to revolutionize parcel delivery with new techniques.

We in the library field would be remiss if we did not carefully examine the prospect that, within the next decade, we will begin to see major changes in the way that much intellectual activity is recorded and presented. Sometimes the dreams of the dreamers come true. The following is from a publication describing the purpose and activities of the Media Laboratory at MIT, where people from many disciplines are

dedicated to turning certain kinds of dreams in the realm of electronic media to reality:

Publishing has meant printing for as long as paper has been the vehicle of mass distribution. More recently, electronic delivery systems have begun to offer competitive processes and new opportunities. The fundamental change, however, is not from ink to phosphor. It is the change effected by personalization and interactivity, wherein monologues become conversations. In electronic publishing the equivalent of the package or the binding is a process, not an object. . . .

Electronic publishing poses basic questions of authorship. It blurs, if not eliminates, the boundaries among authoring, editing, publishing, and reading. Formerly, the notion of packaging was intrinsic to publishing. Information was expected to be at a level of refinement and plural interest sufficient to justify the publishing effort itself and in some media, like film, where the sequencing of the information is all important, to be built into the packaging. In electronic publishing these constraints no longer hold. There is a continuum between thinking out loud and delivering polished thoughts. The flow and exchange of ideas are no longer destined primarily for the bookshelf or library but are a part of daily life on a global scale.²

In this paper I explore current developments that are leading in the direction that the people at the Media Lab see as our future, developments that may well redefine the concept of publishing, changes that will require libraries to adapt to new conditions as they have never done before. It's almost always foolish to make grand predictions. Few who do so prove accurate. My purpose here is not so much to make accurate predictions as to characterize a trend and try to see possible implications of the electronic technology onrush for those of us in libraries. I want to convince you that the future can plausibly bring about some enormous changes in our field. They may not be the exact changes I now envision, but they will most likely be of a similar magnitude. There is one thing I am definitely *not* about to predict, however: the demise of the book! The printed page will be with us for a long time because of its inherent advantages, just as people may still choose to walk, even when given the option to ride. Printing or publishing on paper is undergoing its own revolution, but that is a subject for another time and place.

I'd like first to take a look at some key technological developments. I will then try to chart the likely convergence of these developments in the creation of a new publishing and communication medium. Finally, I will suggest what the emergence of this medium could mean to libraries and how we as librarians might deal with it. The key developments are in six areas:

1. Powerful and portable computers
2. User interface software
3. Optical media
4. Hypertext
5. Artificial intelligence
6. Authoring systems

Each of these exists in some form today but is undergoing rapid, further development. Within ten years all of them will be mature and in common use. Let us first look at each technology individually, then explore how, in combination, they may have a profound effect upon libraries.

KEY TECHNOLOGIES

1. Powerful and portable computers. We have not seen the end of the twin spirals of data processing: the power of the machines soaring upward and their size and cost plummeting downward.

During the next decade we will finally see a truly usable portable computer. It will be powerful enough to support software of the capability that now can only be run on mainframes. Just as the PC enabled the invention of new types of software (spreadsheet, desktop publishing, etc.), these more powerful, portable machines will allow creative users to come up with applications now undreamed of that will change the way we do many things.

The new portable computer I image will be possible, in part, because of a high-resolution screen: flat, sharp, and requiring very little power. It will be about the size of a typical printed book page.

This machine, not always in portable form, will be widespread in the home, car, and office. Most people will have several, although none may be called computers—just as toasters, hair dryers, and electric ranges and irons are not called resistance heaters in common parlance. These computers will be built into a variety of appliances: the meal preparation center, the automobile navigator, and what some might

call the home knowledge center. We will know that the machines are there, but we will not call them computers. We might call the portable version of our knowledge center a "Dynabook." The simplest model will have a limited number of controls; all will have the ability to read from and write to a highly dense storage medium, no doubt optical.

There is much more that could be said about the hardware, but that is really the least interesting area of development. We can be certain that we will see a device within a few years that approximates the brief description above.

2. User interface software. There certainly won't be a single user interface for all these machines, anymore than your stove and car share a single standard for the way in which you interact with them. Yet there are de facto standards for kitchen ranges and for automobiles, and a manufacturer assumes some risk if it departs too radically from those standards. The same is becoming true in regard to interaction with computers (despite efforts to enforce limits on such de facto standardization through legal action). We are moving toward some kind of graphic interface with a pointing device (finger, mouse, etc.) Some like to call an interface such as that of the Macintosh (toward which IBM-oriented programming is moving) "intuitive." It may well be, in the sense that pointing is a primitive behavior and that we don't seem to process our thoughts in terms of language but in abstractions that may have more of a visual basis. The latter point may be argued, but it you have seen Alan Kay's videotape of his accountant's twenty-one-month-old child using Macpaint on a Macintosh, you might have to agree that if not intuitive, the interface is certainly easy to learn. For my part, I chose to use a Mac because of that interface. I've been laboring to remember command languages for computer systems for years, and I am delighted with a machine that is more transparent, so that I can concentrate on the intellectual content of writing, spreadsheets, etc., rather than on the process of recording and modifying that content.

An electronic device will not be successful in the mass market, it is said, until it is designed in such a way that a child can op-

erate it. The audiocassette player, the VCR, and now the CD player all have met this criterion. The reading device for our new electronic publications will need to be as simple . . . remember the student with the Dynabook I described earlier. But just as one can buy stereo systems with differing levels of sophistication in terms of control, the same will be true of our electronic reader or Dynabook. While sharing the same basic user interface, some will have but a few simple buttons for controlling the basic operation of reading. Others will allow for input of data rather than just instruction; some will even have an auxiliary keyboard or voice input device.

What is important here is that the underlying concept of the interface be shared industrywide. Each manufacturer will put its own mark on its version, but among the different versions, there must be more in common than not. If there are competing interfaces (as there are now), it is likely that there will be a painful thinning of vendors until one is ultimately dominant.

A footnote to the issue of interfaces: we are already seeing software that will allow you to define your own interface or to adopt the one with which you are already familiar. The power of our new computer will be such that these ideosyncracies can easily be supported, so that if, in ten years, you still hate the Macintosh-type graphic interface offered by all of the vendors, you will be able to have your own.

3. Optical storage, the publishing medium. You are probably tired of reading predictions and hype about optical disks, CD-ROM, Laserdiscs, and the like by now, but, needless to say, in some form such compact data storage is essential for the distribution of the new publications. I doubt that the medium used will be the present CD-ROM, even though that will probably be a standard for optical publishing during the next several years; the follow-on medium may be a write-once or even a read-write disc, in order to meet some of the needs that I will describe. It will hold more than 550 megabytes of data as well (or 500 typical, large books without illustrations). It's even possible that there may be two optical media in some versions of the reader: the disc, of whatever sort becomes the standard, carrying the large, in-

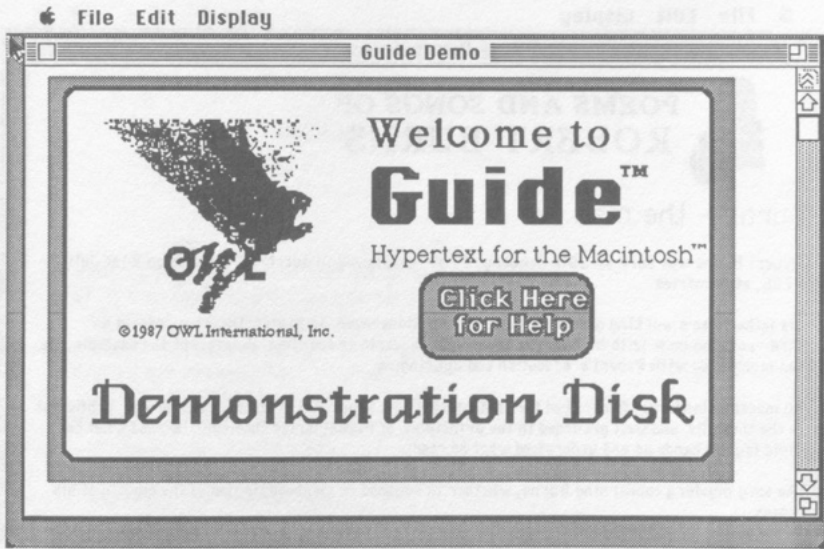
tegrated databases, and an optical card for individual works (analogous to pocket books, perhaps?) and other input and output tasks.

So now we have a portable, powerful, book-sized computer with a massive storage medium and an intuitive user interface. Are we going to use it to replace books? Not at all; we'll use it to interact with a new intellectual form that transcends the concept of the book. For want of a better term, we'll adopt the word *hypermedia*, the fourth technology.

4. Hypermedia, a term generalized from *hypertext*. That, in turn, is a word that has become popular recently but was coined in the mid-sixties to represent a fairly rigorously defined form of a concept that dates from the mid-forties. Ted Nelson, following the work of Doug Englebart and others, refined the relationships among the knowledge elements that would be stored in a computerized version of Vannevar Bush's hypothesized Memex.³ Few paid any attention to this work until CD-ROM came upon us and made it possible to distribute huge databases inexpensively. Hypertext has become a promising means of organizing and providing access to such databases. It is a particular form of nonlinear writing (or illustration, for that matter—thus the generalization to hypermedia). In its simplest form, hypertext involves linkages between different but related ideas and expansions of text, much like online catalogs. In the most advanced of these products, one can navigate their syndetic structure by means of automated linkage of references. In even the earliest versions one can expand displays from brief to full records. Hypertext uses similar techniques but in databases of full text.

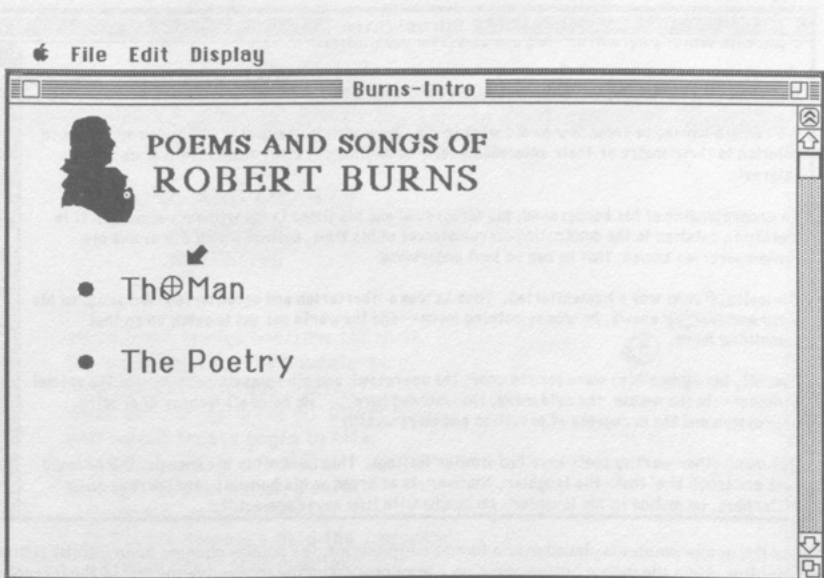
Hypertext and the broader hypermedia, then, are a way of creating and representing multiple levels and relationships of and among intellectual concepts. Hypermedia is interactive; in even its most basic form readers in some sense participate in the writing, because they may organize certain aspects of the presentation that they read. It is possible that readers may even be able to create their own references and replacements and thus participate directly in authorship.

Hypermedia allows writers to write for



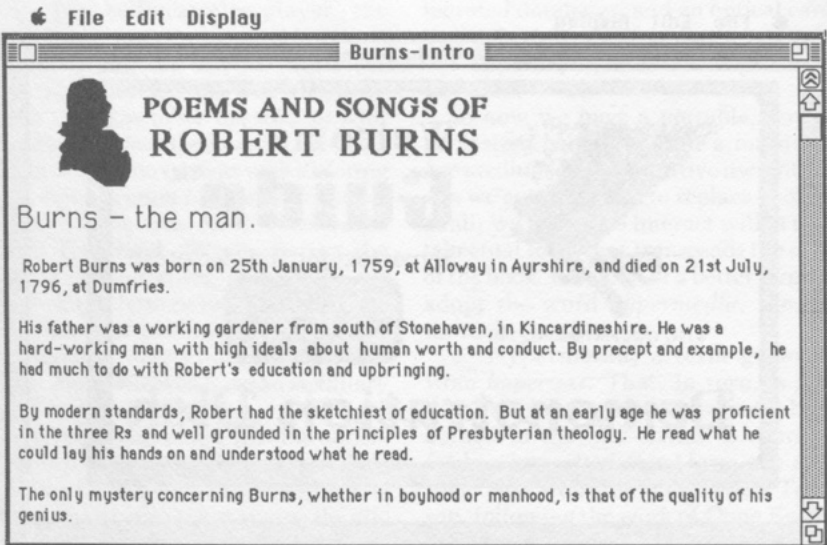
Opening screen of a demonstration of GUIDE for the Macintosh, a hypertext program published by OWL International.

Fig. 1. Opening Screen of GUIDE Demonstration



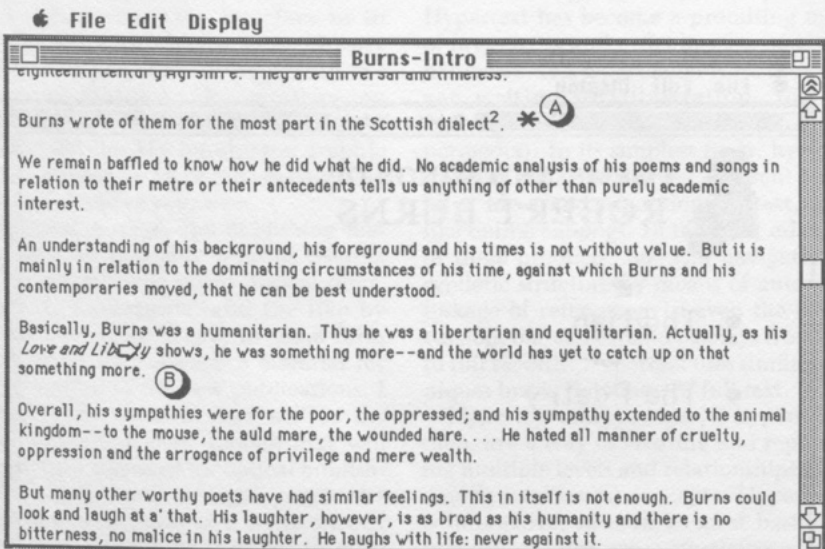
A screen from within the demonstration of GUIDE. When the mouse-driven pointer is placed on text or a graphic that is set up to be replaced with an expansion or substitution, the pointer shape changes to a bull's-eye (see arrow). Clicking the mouse then activates the replacement, shown in figure 3.

Fig. 2. Screen within the GUIDE Demonstration



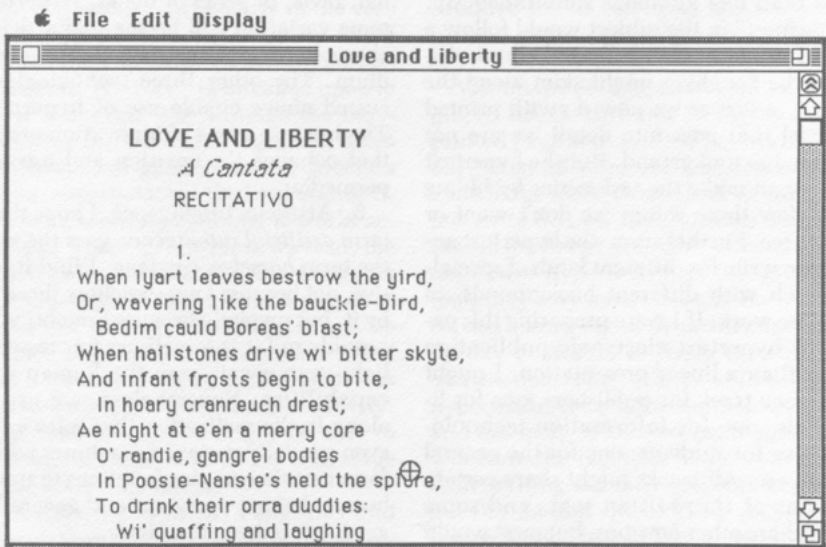
The first page of text of the expanded "The Man" portion of the GUIDE document. It can be scrolled in the normal manner.

Fig. 3. First Page of "The Man"



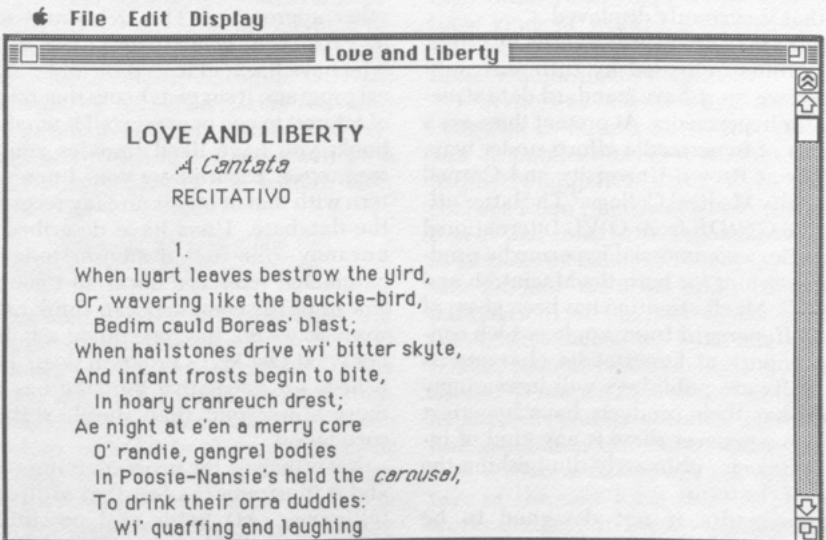
When the mouse pointer is placed over a footnote superscript, the pointer changes to an asterisk (point A). Holding down the mouse button opens up a temporary window containing the text of the footnote (not shown here). Release of the mouse closes the window. In a separate action (point B), linkages between files are indicated by a change in the shape of the pointer to an arrow. Clicking on "Love and Liberty" in this example will bring up a display of a poem itself, which could be stored in a separate file (see figure 5).

Fig. 4. The Mouse Pointer



The poem, which is linked to its name in the text of the previous example. By simply clicking the pointer wherever it is in the shape of a square, one returns to the point of reference in that text. Or, it would be possible to link other files (criticism, derivative works, etc.), creating a fabric of linked concepts. A simple but useful replacement is shown here: when the pointer is placed on a word in the Scottish dialect, such as *splore*, clicking the mouse replaces the word with its normal English equivalent (see figure 6).

Fig. 5. A Fabric of Linked Concepts



The Burns poem with the word *splore* replaced by its normal English equivalent. Clicking on the mouse again restores the original word to the poem.

Fig. 6. Mouse Restores Original Word

more than one audience simultaneously. The "expert" in the subject would follow a work through all of its technical labyrinths, while the neophyte might skim along the top . . . much as we now do with printed material that goes into detail we are not prepared to understand. But the hypertext author can make the task easier by hiding from view those things we don't want or need to see. Furthermore, the hypertext author can write for different kinds of specialists, each with different backgrounds, in the same work. If I were preparing this paper for hypertext electronic publication rather than a linear presentation, I might create one track for publishers, one for librarians, one for information technologists, one for students, one for the general public, etc. All tracks might share certain portions of the written text, and some might share other portions, but most would have some sections dedicated solely to readers in that track.

Hypertext is a form of access that can overcome some of the disadvantages of Boolean logic searching, although it cannot replace text and string searching with or without Boolean functions. It is used in searching to assist the reader in finding and quickly moving to material that is related in one or more specified ways to the material that is currently displayed.

If we are to link works written by different authors or issued by different publishers, we must have standard data structures for hypermedia. At present there are a number of hypermedia efforts under way, notably at Brown University and Cornell University Medical College.⁴ The latter utilizes the *GUIDE* from OWL International of Seattle, a commercial hypermedia product in versions for both the Macintosh and IBM PC. Much attention has been given of late to *Hypercard* from Apple, which contains important hypermedia characteristics. Software publishers will increasingly claim that their products have hypertext features whenever there is any kind of interfile linkage, ultimately diminishing the value of the term.

Hypermedia is not designed to be printed. Its publication is of necessity interactive, to be mediated by a computer. Once it is printed, it is no longer hypermedia but becomes just another linear article, jour-

nal, book, or series of books. Hypertext or some variation will be the major defining characteristic of the new publishing medium. The other three technologies discussed above enable use of hypermedia. The last two areas of innovation are those that enhance the creation and use of hypermedia.

5. Artificial Intelligence. I hope that the term *artificial intelligence* goes the way of the term *horseless carriage*. I find it offensive, not because I am somehow threatened by it, but precisely because I'm not: what is possible so far is sometimes impressive but light-years away from the human mind's capabilities. Nevertheless, we are well along in the realization that software and even specially designed machines will perform certain complex functions that we, as human beings, carry out and generally regard as intellectual activity.

Despite this initial progress, many of those active in the field admit that there is "much more smoke than fire" in artificial intelligence research. The most successful practical applications have been in the area of "expert systems." Examples often cited are used in oil exploration and the configuration of computer equipment. While most expert systems are rule-based and mimic the thinking of a human expert, there are other approaches. The *Del Mar Bookseller's Assistant*, given other names by those who have licensed it, is basically a statistical program: it suggests books that might be of interest to you once you tell it what other books you have liked (movies you have seen, etc.). It correlates your interest pattern with that of others already recorded in the database. Users have described it as uncanny—the fiction advisor today, the automated reference librarian tomorrow? Not until the computer can think associatively as we do. But, one might ask, is that not what *Del Mar's* program does? I don't believe so; associative thinking has many more dimensions than simple statistical correlation.

Nevertheless, the program is impressive, and it illustrates the fact that artificial intelligence (AI) holds real possibilities. What role will AI play in the publications of the future? It will make large hypertext constructions manageable. Hypertext, while simple in concept, can become un-

wieldy in practice, particularly in the very large databases that are foreseen. With writers linking to other writers' hypertext publications, etc., AI will be called upon to help us wend our way through the labyrinths. Our software alter egos will learn our interests and guide us in the direction these interests dictate at the level appropriate to our expertise (assuming it works according to "specs"). The more we read, the broader its experience and the better its ability to predict what we want to pursue in a hypertext session. Obviously it can be overridden or turned off if we so desire, as I would hope that we often would, since this program could stifle intellectual creativity. On the other hand, we might be able to shift it into a mode that would encourage serendipitous encounters.

6. Authoring systems. Programs and hardware that will permit the publication, distribution, and use of nonlinear intellectual works have been discussed. There also must be support for their creation. Just as authors today are increasingly using word processors and very soon, perhaps, their *Microsoft Bookshelf*, the hypermedia authors of tomorrow will be using sophisticated hypertext and hyperimage processors. Their precursors on the market today are known as "authoring systems for interactive publications." Most are oriented toward the creation of interactive instructional publications on videodisc. *GUIDE*, which I mentioned earlier, is relatively simple, hypertext authoring software. Hypermedia authoring software will be a standard component of the writer's electronic tool kit.

IMPLICATIONS OF THE NEW TECHNOLOGIES

These six interrelated technologies will combine to form the basis for what will eventually be a monumental shift in the way that knowledge is recorded and communicated. We are now using electronic technology to assist in the more efficient production and distribution of knowledge in the traditional mode, taking advantage of word processors, sophisticated imaging and illustrating programs, satellite communication, computer-driven typesetting, and desktop publishing. In the future we will use these tools even more heavily to

produce linear text and illustrations on printed pages faster and more economically.

Within the decade, however, we will begin to break through the limits of the printed page and start to take advantage of our potential to record and communicate knowledge more dynamically, using technology to do new things rather than to do the old job better or faster. Authors will not only have hypertext software as a major tool; they may have virtually the entire recorded knowledge of their discipline immediately available to them through their workstation. They will be able to develop hypermedia links between their own work and this body of knowledge. Some of them will have the facilities to incorporate into their writing not only graphics and sound but full-motion video. These facilities will not be limited to scholars. Indeed, scholars may be far from the first to use them.

Some obvious early adopters of the new medium:

- Developers of educational textbooks. We see hypermedia principles being applied in the work of the Visual Courseware Lab at MIT and elementary or embryonic hypermedia in some of the training material being developed for interactive videodisc. Industry will support the creation of these kinds of works for educational and training purposes, and the larger educational establishment will follow.

- Writers of how-to books. I would be delighted to have a series of hypermedia how-to publications that would allow me to quickly jump from a section containing a concept I don't understand to an explanation of that concept and then back, as well as graphics on which to zoom in or even put into motion. We'll see many stand-alone publications of this sort in the next few years, well before we have our portable reader, standard user interface, and AI assistant. Writers will become familiar with hypermedia and its usefulness on their current generation of home computer.

- Writers of fiction will also take advantage of the new form in order to create works designed to involve readers not just as readers but as creators themselves: interactive fiction. We've seen some of this already through online services and tend to scoff at its lack of sophistication. Most of us

see it as a form of video game. Just wait: interactive fiction and poetry will become a major art form by the turn of the century.

The publishing industry will assume responsibility for primary distribution of the new media. That distribution may well be done on demand, however, rather than in press runs (against paper, plastic or metal, as the case may be). Publishers could then operate what we might call (heaven forbid!) "information utilities," supplying the local retailer and, we hope, the library. The retailer might be somewhat like the Source, Dialog, the New York Public Library, and a corner video store all rolled into one. There, at the "neighborhood knowledge center," a stock of hypermedia titles is held on disk or card for sale, rental, or transfer. Since shelf space is not a concern, there is little limitation on what is stocked. Purchasing and stocking might be online through a high-speed connection to the publishers. You could probably dial in to read the publications or download them yourself if you wanted, bypassing the local retailer, but it would likely be more expensive to do so. Once an item is acquired by the local center, only copies of it are made, so it never need be restocked (assuming it is not lost or stolen, of course). Seven-Eleven Food Stores might become a major outlet. You could fill the tank of your car, eat a fast-food snack, and pick up bread, milk, and a complete, annotated library of Russian literature on the way home from work!

Once we begin to accustom ourselves to the efficiencies of hypermedia, paper publication will in fact decline in certain fields. We will see it first in the "sci-tech" and medical areas. Is the decline not already under way in one sense even without the introduction of the new technologies? To be sure, the total amount of printed material on the market in these fields is increasing, but how much more might it be increasing if there were no online information sources? CD-ROM will likely make a further impact, but increased publication in hypermedia will accelerate the trend for the simple reason that hypermedia can't be printed on paper and still retain its unique characteristics and advantages.

Once the market for hypermedia publications has been established, the new hardware specifically designed for it will ap-

pear. The basic, stripped-down Dynabook (should we call it the Hyperbook instead?) will allow one to read, view, and otherwise interact with hypermedia. Other versions will permit one to create new linkages of one's own choosing, personalize the publication, or relate it to other publications. Such devices may also let one annotate the work, insert electronic bookmarks, and, in the most complete version, create new hypertext publications. Of course there will be add-ons: a high-speed modem, additional optical memory drives or readers, a digital TV module—so that the Dynabook doubles as a portable TV and possibly supports the integration of broadcast video into one's hypermedia database. Wouldn't you really love to have a book you could read at the beach without the wind flipping pages? And that would allow you to look in on the afternoon baseball game every now and then?

Eventually we will use this device to organize and transform published and private knowledge into our own, personal, electronic libraries and then to access the components of that library through a variety of searching and hypermedia programs. Some would describe this library as a "seamless" collection of knowledge. There will be seams, but they will be soft, not the hard seams physically imposed by the format of the book. They will be logical seams, imposed by the data structure onto which the knowledge is imposed—or, rather, that is imposed upon the knowledge. The owners of the libraries, with the help of AI assistants, will create the threads that tie these data structures together. They will define, through hypermedia techniques, those relationships that are meaningful to them. The reader becomes in a real sense an extension of the author. A new relationship to recorded knowledge will have come into being.

A word to the skeptics: in this scheme of things it is true that a machine must be used to mediate between the "knowledge container"—in this case a disk or card encoded optically—and the reader. One advantage of a book is that it requires no such mediation . . . it can be read directly. The machine, it is argued, is inconvenient, cumbersome, and expensive. Yet have we not come to terms with machine mediation

in the visual and aural realm? We derive benefit from the use of machines that will play recorded music, for example, and so we have the Walkman. Few of us are able to perform the music we desire to hear; we are dependent upon someone else's performance. Back in the heyday of the trolley, before the creation of the recording industry, if we were very wealthy we might have been able to hire musicians to perform for our personal benefit. We heard music in church or in concerts, but otherwise what we now call folk music *was* music for most of us. It was what we, our family, and friends could perform and as such was usually played with relatively simple, if any, instruments. Now there are few limits to what we can hear. The machine has given us the potential to range widely among styles, forms, and interpretations of countless musical genres. Have our lives not been enriched? We accept machine mediation for sound and video, because without the machine, we would not perceive the performance. We tolerate the machine because of the benefit it brings us. We use the machine because of its ease and low cost.

Now if all we wanted or needed was to read in a linear fashion, as I am writing now, there is probably little likelihood that a Hyperbook would ever be needed. But we will develop and use the new form, the interlinked, nonlinear publications. We will do so because of the increasing complexity of knowledge and the limitations of our own minds, in order to find our way through this mass of complexity, extract what we need, and make the connections we must make in order to solve the problems with which we are confronted in present-day life. The Hyperbook will offer us a way to obtain those benefits. Its high-quality display, low cost, and intuitive user interface will make it functionally acceptable. Need I repeat that an entire generation is increasingly interacting with ideas and images through machines, beginning with computers and video games, going through interactive video training, and maybe extending soon to CD-I. They'll not only tolerate the Hyperbook, they'll love it.

LIBRARIES AND THE NEW TECHNOLOGIES

Librarians too will embrace the new me-

dium, with much less reluctance than they have accepted microfilm and the audio and visual media introduced in recent decades. The reason is that hypermedia inherently contains a means of helping librarians increase their ability to achieve two of the profession's major goals: the organization of and access to recorded knowledge. Hypermedia is not just another medium, like print or film or video; it is a means of intellectual organization and access to knowledge. Our focus will not be on the technology but on the body of knowledge to which this technology can be applied. We will eventually convert even our catalogs to hypertext, making them an integral part of civilization's store of knowledge, not something standing outside of it and pointing from a distance. The catalog can be linked to the works it describes.

Some librarians will become hypertext engineers, if I may suggest such a term. They will provide intellectual connections between the works of different authors or convert linear publications to hypermedia publications. Our subject analysis and indexing expertise will be put to the ultimate test. There will be even greater need for librarians with solid subject expertise.

I see a new form of interlibrary cooperation: the sharing of hypermedia links. LITA will have an Interest Group on Hypermedia Templates. Such templates will be the hypermedia structure and relationships that can be imposed upon copyrighted, linear works (requiring specified editions, of course, because the template will be keyed to particular locations in memory on the optical disks or cards containing the works). These templates will be supplied to our patrons; they aren't going to want to create all of the linkages themselves, although they will often want to modify them as they read and explore.

Libraries may become publishers of full hypermedia as well. They hold much material in the public domain that could be preserved and more widely distributed if put on an optical medium. While current experimental work revolves around image storage, rather than text storage, the two can be combined, and new generations of optical scanners may make it possible for us to preserve both the visual form and, separately, the intellectual content in

character-manipulative form. The connection between the two would be another example of the kind of linkage I've described earlier. And once converted to character form, of course, we could then go on to integrate that work with the rest of our hypermedia knowledge store. I should point out that librarians need not wait for the arrival of hypermedia techniques to begin publishing material in their collections in optical form. There may well be a market for particular collections transferred in linear form to CD-ROM.

Librarians will continue to do what they do now as well: collect, organize, and provide access to intellectual records in a wide variety of formats. Hypermedia will be collected, organized, and circulated as well. Despite what I've said earlier, I don't believe that the whole process of hypermedia distribution and access will necessarily be commercialized. Bookstores haven't put libraries out of business. Libraries themselves can become an information utility of the sort I described earlier. We who run libraries know about local outlets, telecommunications, and information. We've built OCLC and RLG, why not the information utility?

I believe that we librarians will extend our mission, not only to collect and offer our publics the new media but to provide even greater guidance to those who enter our doors or log onto our systems. Through the use of computer technology we have been able to offer more and more "hits" that may or may not satisfy a user's need. We will offer hypermedia reviews that will help our users select among all those hits and find what best suits their requirements, reviews linked to the catalog that is linked to the work. And why not make it possible for the patrons to add their personal reviews to the catalog?—monitored by the

catalog editor, of course. There is so much that the library can do that the private sector will have no interest in doing because it will be unable to make a profit doing it. Advancing technology will offer new opportunities that no one else will be able to pursue. We should seize those that are consistent with our mission at the earliest feasible time.

One of those opportunities is about to present itself to us soon. Products based upon optical technologies are coming into the market. We should find the best, acquire them, use them, and, when appropriate, lend them. We should adopt this new technology as soon as possible in order to identify the library with it. Such an identification in the public mind will make it easier in the future to advance the library as the base for hypermedia. We will have much to do to build that base, and we need to start now.

Such efforts can help to assure that the library will not go the way of the Railway Express Agency or the streetcar. We have kept the library a vital institution with the assistance of new technologies in the past. While the future may be quite different from what I hypothesize here, there is no doubt that new forms of publication are emerging that will require the use of new technologies. We must pay attempt to them and integrate those that are successful into our collections and our services. The library, with its mission of preserving and making available what has been recorded in the past, along with its technological experience, is in a unique position to bridge the old and the new. We will find ways to help students to advance their knowledge through the use of the Hyperbook and the traditional printed book alike. To do so we must neither forget the book nor ignore the direction of technology.

REFERENCES

1. Stewart Brand, *The Media Lab* (New York: Viking, 1987).
2. Massachusetts Institute of Technology, Media Laboratory, untitled publicity brochure (Cambridge, Mass.: MIT Pr., n.d.), p.5.
3. Vannevar Bush, "As We May Think," *Atlantic Monthly*, July 1945, p.101-8.
4. "Cornell Med School Uses Hypertext Net," *Infoworld* 9, no. 43 (Oct. 26, 1987). ■■

On the Preservation of Human- and Machine-Readable Records

John C. Mallinson

Since 1979 the author has been a member of the Committee on Preservation of the National Archives and Record Service (NARS), which is now an independent agency (NARA). This committee was subdivided into three subcommittees charged with advising NARS on its paper, photographic, and machine-readable (mostly magnetic tape) records, respectively. The last, subcommittee C, has the following membership:

John C. Mallinson, Chairman, Ampex Corp., now UC-San Diego;
 John C. Davis, National Security Agency;
 Walter O'Neill, IBM Corp.;
 Andrew M. Persoon, MMM Corp., now retired;
 Leslie S. Smith, National Bureau of Standards;
 Richard G. Zech, McGraw-Hill, now Rochchild Associates.

Initially subcommittee C considered its mission limited to advising NARS how best to expedite acquisition of 1/2-inch computer tapes. This remains a critical issue, since NARA has only accessioned an extremely small fraction, about ten thousand reels out of an estimated twenty million reels in U.S. government use alone.

It subsequently became clear to subcommittee C that there is not only a profound misunderstanding in the archival community concerning the nature of machine-readable records but also an unfounded belief that some unidentified future

technology will emerge to resolve most of the present and future machine-readable record problems.

In July 1984 the subcommittee delivered to the archivist of the United States a white paper, "Strategic Technology Considerations Relative to the Preservation and Storage of Human and Machine Readable Records," which recommended that, for the reasons expounded below, NARS formulate a policy whereby in the main, the majority of the holdings be on human-readable microfilm.¹ Discussions of these findings have appeared in recent publications.²⁻³ In one paper, a NARA representative stated that the findings were not acceptable to the agency.⁴ Beyond this, NARA has not yet responded to the white paper.

DEFINITIONS

In machine-readable records, such as magnetic computer, audio and video tape, magnetic and optical disks, and phonograph records, it is understood that the recorded information can only be recovered, in a practical sense, by converting it to human-readable form, such as a paper text, photograph, or video terminal display. In analog recordings this conversion requires appropriate hardware, and in digital recordings, hardware, software, and documentation.

The information in human-readable records is, on the other hand, comprehensible simply by visual inspection of the record or through a magnified image of it. Only sim-

John C. Mallinson is Director, Center for Magnetic Recording Research, University of California-San Diego.

ple optical hardware, such as microscopes and projectors whose design principles can never change, are needed to read the record completely.

INFORMATION AGE

Modern civilizations are now considered to be entering the "Information Age," during which the vast majority of the societies' information and records is stored, manipulated, and disseminated by electronic means such as computer networks, earth satellite relays, and television broadcasting. It appears to subcommittee C that the archival community tends to forget that the principal motive for these technologies is their speed of access, with the concomitant ability to perform electronic data processing rapidly, and that this speed is only achieved at an extremely high cost. The machines themselves—computers, satellites, and TV receivers, for example—are rarely expected to have a useful life in excess of ten years. The machine-readable records are operated at ever-increasing information storage densities, a trend that is surely inimical to long-term archival preserva-

tion, in order not only to store more information but also to decrease access times. Since 1956 no less than ten differing, incompatible videotape formats of increasing storage density have emerged, and since 1952, eight differing computer tape formats have been used; these are listed in table 1. Each format typically mandates a different machine with its unique set of demodulators, decoders, reformatters, etc. This proliferation of incompatible systems is the root cause of the archivist's dilemma with machine-readable records. The speed of access and electronic data-processing abilities are indeed attractive, but the records and their associated hardware will become obsolete within a couple of decades! Since the information and communication industries are most definitely not driven by long-term archival considerations, it seems to be futile to expect technology to resolve this problem. Technology continues to cause the "machine-readable" problem and will not solve it.

ARCHIVAL HIERARCHIES

Subcommittee C believes that the fol-

Table 1. Video and Computer Tape Formats

Product	Current Status
Video tape formats since 1956 (analog, frequency modulation)	
2-inch quadruplex	Obsolete
2-inch quadruplex, double density	Obsolete
1-inch helical, type A	Obsolete
1-inch helical, type B	
1-inch helical, type C	
3/4-inch helical, U-Matic	
1/2-inch helical, Betamax	
1/2-inch helical, VHS	
8mm helical	
4mm helical	
1/2-inch computer tape formats since 1956 (digital, various channel codes)	
7-track NRZI, 100 BPI	Obsolete
7-track NRZI, 200 BPI	Obsolete
7-track NRZI, 556 BPI	Obsolete
7-track NRZI, 800 BPI	Obsolete
9-track NRZI, 800 BPI	
9-track PE, 1,600 BPI	
9-track GCR, 6,250 BPI	
18-track RLL, 19,000 BPI	

Code key: NRZI = Non-Return to Zero Inhibit; PE = Phase Encoding; GCR = Group Code Recording; RLL = Run Length Limited

lowing hierarchy of priorities applies to NARA and other major long-term archives.

Top Priority: To accession, to preserve valuable records indefinitely, and to make them available to the public.

Note that this first priority has no inherent need for instantaneous electronic access; indefinitely long preservation is the principal objective.

Secondary Priority: To make the archived records available to the public in a timely, inexpensive manner.

For the majority of records in national, state, and county archives *timely* probably means "within a few tens of minutes."

Lowest Priority: To provide the archived information in an electronic form virtually immediately.

This priority applies only to a small minority of the archives and permits, for example, multiple, time-sharing video terminal access and the ability to perform electronic data processing.

It will be noted that, if this hierarchy is valid, there is no justification for an endeavor to operate NARA as a computer-driven operation. It is acknowledged that other organizations follow different hierarchies. For example, the Library of Congress Optical Disk Pilot Program addresses the lowest priority (above) and coincidentally achieves short-term preservation goals by restricting the public's access to the original documents, photographs, and books. Similarly, the Smithsonian Optical Disk Project is primarily a broadcasting or publishing operation following different priorities.

ARCHIVAL PROPERTIES OF MAGNETIC AND OPTICAL RECORDING MEDIA

Magnetic tape has three components: the base film, which is a plastic (Mylar, Estar, Celanar, etc.) of exceptionally good stability; the particular binder system, which is usually of polyesterurethane; and the magnetic particles themselves. The binder system is subject to hydrolysis, much like gelatin in photographic film, but under normal conditions (65-75 degrees F., 40 ± 5 per-

cent R.H.) reaches a satisfactory equilibrium state. The most common material used for the magnetic particles is iron oxide, which is known to be stable. Other materials, for example, chromium dioxide and elemental iron, are vulnerable to deterioration at elevated temperatures and humidity. Generally, however, it is believed that magnetic tapes, properly stored, will outlast the hardware (computer drives, videocassette recorders, etc.). When rapid angular acceleration of the tape is expected, as in most electronic data processing operations, periodic (five years) rewinding is needed in order to maintain adequate layer-to-layer tape pressures.

In the future the inexorable trend toward higher storage densities will lead to magnetic tapes and disks that use very thin metallic layers. It is believed that not only will such tapes approach, within a factor of two, the storage densities of optical and magneto-optical disks but that the physical properties of the metallic storage layers will be very similar. Thus, magnetic disks with a 200-A-thick layer of a Co-Ni alloy and magnetic optical disks having a 150-A-thick layer of a Co-Fe-Tb alloy will be in widespread use. The archival properties of metallic disks, be they magnetic or optical, are not presently known; estimates of ten to thirty-year lifetimes are current. Again, the salient point is that the recording medium may well outlast the hardware. Silver halide photographic film, processed according to ANSI and federal standards, is, of course, a certified archival medium whose long-term storage properties are known and acceptable.

ARCHIVAL PROPERTIES OF SOFTWARE AND DOCUMENTATION

In a certain class of machine-readable records, namely those employing computer or digital technologies, a further archival problem arises. The mere recovery of the digital data is not possible without some software. The proper operating system must be available at the time of data recovery. Sadly for the archivist, the operating systems today are changing even more rapidly than the hardware. For example,

Western Electric's UNIX operating system has been offered in over thirty versions in one decade. Offsetting this serious problem there are, of course, many potentially attractive reasons for using digital recording techniques: the compatibility with the computer environment and the ability to perform perfect error detection and correction are prime examples.

Given the proper operating system, so that the digital data can be read out, a further requirement arises. Appropriate documentation must be at hand to provide the necessary information on the digital codes used, the organization or format of the record, and several other minor but critical details.

The operating systems are usually resident upon computer tapes or floppy disks, thus compounding an already difficult archival problem. The documentation may be in machine-readable or human-readable form but, given the human species' well-known tendency to procrastination, may well be incomplete or missing.

ARCHIVAL PROPERTIES OF HARDWARE

The fact that most electronic hardware is expected to function for no more than ten to twenty years raises very serious problems for long-term archival preservation. Given that the operating system and documentation problems have been dealt with somehow, what is the archivist to do when the machine manufacturer declares the hardware obsolete or simply goes out of business? Will there be an IBM or a Sony in 2200 A.D.? If they still exist, will they maintain a 1980- to 1990-vintage machine? Moreover, it must be realized that no archival organization can hope realistically to maintain the hardware itself; integrated circuits, thin film heads, and laser diodes cannot be repaired, nor can they be fabricated except in multimillion-dollar factories. The inescapable conclusion is that if an institution like NARA acquires machine-readable records, it will be committed eternally to "file conversion" (i.e., copying records in the old, obsolete format into the new, current format) approximately every ten to twenty years. Repetitive file conversion is, of course, an extremely expensive

preservation strategy for an archive since it leads, within ten to twenty years, to the situation where almost the total archival holding is locked into an eternal file conversion cycle. Precisely such file conversions take place all the time, of course, in today's computer facilities, but there the important and critical difference is that records are continually being retired—perhaps, indeed, to be sent to NARA!

SUBCOMMITTEE C'S RECOMMENDATIONS

Subcommittee C has recommended that NARA store and preserve most of its future accessions in a human-readable microfilm mass memory. The advantages of a microfilm-based archival approach are as follows:

(a) It permits a complex, evolving, electronics technology-driven problem to be understood and solved in a records management and preservation context;

(b) The records are stored upon a certifiable archival medium (silver halide film processed to ANSI and federal standards);

(c) It employs a single, basic technology that is not likely to change significantly in the future;

(d) Because the record is human-readable, the information will be preserved independently of the ever-evolving electronics technology;

(e) It is compatible with future electronic information-handling technology: CCD scanners, automatic filing, storage and retrieval, systems under computer control, and remote video terminals can all be added or changed in the future;

(f) It is an essentially risk-free solution requiring no research and development breakthroughs.

Subcommittee C's recommendations may be perceived by some as anachronistic or inappropriate in today's electronic information age. The subcommittee members, all senior technologists in their respective high-technology organizations, have unanimously come to believe that these recommendations are the only solution that is logically defensible today. While risking being characterized as "technological Philistines," they, nevertheless, welcome debate from any quarter.

REFERENCES

1. The complete white paper, by Richard Zech, will be provided by the author upon request.
2. *Preservation of Historical Records* (Washington, D.C.: National Academy Pr., 1986), chap. 6 and 7.
3. John C. Mallinson, "Preserving Machine-Readable Archival Records for the Millenia," *Archivaria* 22:147-52 (Summer 1986).
4. Sue Gavrel, "Preserving Machine-Readable Records: A Reply to John Mallinson," *Archivaria* 22:153-55 (Summer 1986). ■■

Twenty Years Ago in JOLA

Subscription to *Journal of Library Automation* is included in membership dues to Information Science and Automation Division, American Library Association. Other subscriptions are \$10.00 per year. Single copies are available at \$3.50.

From the masthead, *JOLA* (Mar. 1968).

Measures of Expected Online Catalog Performance for Public Access Terminals

Raymond G. Taylor, Jr.

In recent years the library literature has included numerous references to queuing models for determining the minimum number of online catalog terminals needed to provide a specified level of service for public access. The entire July 1981 issue of *College & Research Libraries*—as well as works by Tolle;¹ Tolle, Sanders, and Kaske;² Knox and Miller;³ Borgman and Kaske;⁴ Taylor;^{5,6} and Gebbie and others⁷—have provided a reasonably strong basis for understanding the queuing method and its utility for online catalog systems.

DETAILED QUEUING INFORMATION

Missing from the literature, however, is an indication of the additional power of the queuing equations to provide detailed, expected performance measures for online systems. Once the arrival rate, service rate, and service policy are determined, the minimum number of terminals is given by the queuing model;⁵ thereafter, the model also has the capacity to answer many important questions and to specify levels of expected service. For example, the following types of questions could be answered:

1. How seriously will the system degrade if 1 terminal is "down," or 2, or 3?

2. If the service policy of the library is that "ninety-nine percent of the time patrons shall wait less than one minute for an available terminal," what is the chance that they will wait 2 minutes, 3, or even 5?

3. How many patrons on the average will be using the system?

4. If the flow of patrons could be partially controlled (such as by scheduling classes or demonstrations), what is the maximum number of patrons that could be served?

5. If there is space for only 6 patrons to wait, how often will that number be exceeded?

6. If a demonstration is scheduled for 5 persons during a busy time and 5 terminals are being used, what effect will that demonstration have on the waiting time of other patrons?

The above questions pertain to system performance as a function of the number of terminals available. Similar questions can be asked using arrival and service rates as variables. For example, consider the following questions pertaining to arrivals:

1. During a period of initial curiosity, the new online system may be used by a larger number of patrons than anticipated; how will wait times be affected by a 25 percent increase in demand?

2. Once the system is accessible through home computers using modems, the arrival rate at the public access terminals is expected to diminish; if the rate drops from 2.5 patrons per minutes to 2.2, how many public terminals can be relocated into the staff's work area?

3. If campus enrollments increase by 20 percent over the next three years and are re-

flected in proportionately higher library use, how many additional terminals will be needed; or, if no additional terminals are provided, how will the performance of the system change?

Similar questions regarding service rate can also be answered from a detailed report of the queuing calculations:

1. When searching for known items, patrons presumably use the terminal for a shorter length of time than when searching by subject; when subject searching comes online, how much system degradation can be expected for each 10-second increment in average search-session time?

2. Search-session time for new and inexperienced users is very high, contributing substantially to the average total search time; if through an improved orientation program and staff assistance, new users could quickly become competent as searchers and if the mean service rate thereby improved, how would system performance respond?

These, and many other questions, can be answered from the queuing model. If arrival and service rates are properly measured, and if the library adopts a specific service policy,⁶ very little additional work is needed in order to use the queuing model to determine the recommended number of terminals and to obtain the sort of detail cited above. Librarians in universities may wish to consult their operations research department; or, readers may request a complimentary copy of the microcomputer program written by the author for this article (North Carolina State University, Box 7801, Raleigh, NC 27695). Some information regarding this program is provided in a technical note included here.⁷

ILLUSTRATION

The tables that follow are merely illustrative. They are based on specific arrival and service rates found empirically at NCSU (2.5 patrons per minute, .18 patrons served per minute per terminal).⁸ Further, they are based on the particular service policy adopted by NCSU—the same policy used as an illustration above. The kinds of information given in the table below, however, can be found for any meaningful

combination of arrival and service rates and policy.⁹

Table 1 provides a summary of the input data, information provided by the program, and results of a few very basic calculations. The *effective service rate*, for example, indicates how many patrons could be served per minute if they arrived in a smooth flow and if they all used the terminals at the given service rate (.18). This is the theoretical limit of the system's performance, given the specified arrival and service rates and number of terminals; no amount of administration or control can improve the system beyond this limit.

Note in the first table that *steady state probabilities* are requested for up to thirty persons in the system. *Steady state*, in this context, means that the initial effects of starting up the system (at the beginning of a day or a peak period) have worn off and that the system has settled into its typical behavior.

In the parlance of queuing theory, *number in the system* refers to the number of patrons in the queue plus the number of patrons being serviced; if there are 22 terminals and 12 patrons *in the system*, then there are 10 vacant terminals, and no one is waiting.

Finally, the first table also indicates that *waiting time probabilities* are requested for 1, 2, 3, 4, and 5 minutes. In each case, an analysis is being made of the likelihood that a patron will wait in the queue more than the specified amount of time. The times can be modified if the library researcher is interested in very short or very long waits; the times selected for the present illustration

Table 1. Cycle Identification and Summary Statistics

Usable Terminals = 22
Mean Arrival Rate, Patrons per Minute = 2.5
Mean Service Rate, Patrons Served per Terminal per Minute = .18
Average Time Individual Patron Spends on Terminal = 5.56 Minutes
Effective Service Rate, All Terminals Combined = 3.96
Steady State Probabilities Provided for N from 0 to 30
Waiting Time Probabilities Provided from 1 to 5 Minutes

bracket those mentioned in the literature.

Table 2 contains a great deal of information that can be used to answer many questions. The number (N) given in the left column refers to patrons *in the system*, as defined above. The second column indicates the probability of *exactly* that number being in the system, no more and no less. The third column indicates the likelihood of the system containing the number specified *or fewer*, and the last column, the likelihood of the system containing *more* than the number of patrons specified.

Consider the following kinds of inquiries that could be made of table 2. (The queuing model solution is typically based on arrival-service rates from peak periods, so the reader should assume that all given inquiries and uses of the tables refer to peak periods.) How often will no one be using the system? (so rarely that, for practical purposes, the answer is "never"; $P[X = <0] = .000$). How of-

ten will there be 10 or fewer users? (about 18 percent of the time; $P[X = <10] = .182$). How often will 1 or more persons be waiting? (about 2 percent of the time; $P[X >22] = .019$). How often will exactly 1 patron be waiting? (less than 1 percent of the time; $P[N = 23] = .007$).

Table 3 indicates the likelihood of waiting in the queue more than 1 minute, 2 minutes, and so on. It also provides four performance averages: number in the system, number in the queue, time in the system, and time in the queue. Here the policy can be verified; note that the likelihood of being in the queue more than 1 minute is less than 1 percent, which matches the specified policy that 99 percent of the time patrons shall wait 1 minute or less. The chance of waiting 2 minutes or more is only 2 in 1,000 ($P[WT > 2] = .002$), and the chance of waiting 3 minutes or more is nil. In the present illustration, if the library ad-

Table 2. Steady State Probabilities

Number (N)	$P(X = N)$	$P(X = <N)$	$P(X < N)$	
0	0.000	0.000	1.000	
1	0.000	0.000	1.000	
2	0.000	0.000	1.000	
3	0.000	0.001	0.999	
4	0.001	0.002	0.998	
5	0.004	0.006	0.994	
6	0.009	0.015	0.985	
7	0.018	0.033	0.967	
8	0.032	0.065	0.935	
9	0.049	0.114	0.886	
10	0.068	0.182	0.818	
11	0.086	0.268	0.732	
12	0.100	0.368	0.632	
13	0.106	0.474	0.526	
14	0.106	0.580	0.420	
15	0.098	0.678	0.322	
16	0.085	0.762	0.238	
17	0.069	0.832	0.168	
18	0.053	0.885	0.115	
19	0.039	0.924	0.076	
20	0.027	0.951	0.049	
21	0.018	0.969	0.031	
22	0.011	0.981	0.019	Being Served
23	0.007	0.988	0.012	Waiting
24	0.005	0.992	0.008	
25	0.003	0.995	0.005	
26	0.002	0.997	0.003	
27	0.001	0.998	0.002	
28	0.001	0.999	0.001	
29	0.000	0.999	0.001	
30	0.000	1.000	0.000	

Table 3. Probability of Waiting Time in the Queue Exceeding T

T	P(WT>T)
1	0.007
2	0.002
3	0.000
4	0.000
5	0.000

Average Number in System	= 13.94
Average Number in Queue	= 0.05
Average Time in System	= 5.58
Average Time in Queue	= 0.02

ministrator walked by the public access area at random times during peak periods, he or she should expect to see about 14 terminals being used and 6 vacant (average number in system = 13.94). The average length of time that a patron will be in the system, that is, waiting for and then using a terminal, is about 5.6 minutes (average time in system = 5.58 minutes).

Although this information may at first appear academic in nature, examine what happens when 1 or more terminals are "down" or when the arrival-service rates change.

In the author's program, tables 1, 2, and 3 are regenerated or recycled under user control for various modifications of terminals available, service rate, and arrival rate. The successive cycles renumber the tables with an incremented suffix; for example, table 1 becomes table 4. In order to illustrate the value of these secondary cycles, table 4, 5, and 6 are reprinted here and explained. The second cycle was given the following characteristics: 22 terminals is the normal complement, but 4 are out of use; the peak arrival rate is normally 2.5 patrons per minute, but a surge of 3.0 is being experienced; the normal service rate of .18 is being maintained.

In the illustrated second cycle the degradation of the system is apparent in tables 5 and 6. In a previous inquiry, for example, it was determined that at least 1 person would be waiting 2 percent of the time. Under the modified condition, patrons would wait 6 percent of the time (table 5, $P[X > 18] = .619$). In a previous inquiry, it was noted that the likelihood of waiting up to 1 minute was less than 1 percent; under the modified condition that likelihood is greatly increased to more than 50 percent

(table 6, $P[WT > 1] = .526$). Similarly, it was noted in a previous observation that when all terminals are in use and the arrival rate is normal, there is virtually no chance of having to wait 5 minutes; under the revised conditions that chance is 1 in 5 (table 6, $P[WT > 5] = .201$).

As expected, if the system is properly sized, it can tolerate minor changes in terminal availability and arrival-service rates, but if two or more of these variables change in the same direction, or if any one changes substantially, the system may fail to meet the specified service policy by a wide margin.

IMPORTANCE OF THIS ANALYSIS

There are two important groups of reasons why library administrators should have detailed measures of expected online catalog system performance for public access terminals.

The first set has to do with the technical considerations of planning and control. Anticipated changes in equipment availability and/or user demand are reflected in predictable changes in system performance. Using such predictions, budgeting, staff changes, purchase of equipment, staff and user development programs, and other preparatory measures can be taken to offset the effects, or avoid the occurrence, of system performance fluctuations.

The second set of reasons for obtaining detailed measures of expected system performance pertains to the political value of maintaining proper expectations for both the user and the governing/funding agency. Online catalogs represent a major change in the way most patrons are accus-

Table 4. Cycle Identification and Summary Statistics

Usable Terminals = 18
Mean Arrival Rate, Patrons per Minute = 3
Mean Service Rate, Patrons Served per Terminal per Minute = .18
Average Time Individual Patron Spends on Terminal = 5.56 Minutes
Effective Service Rate, All Terminals Combined = 3.24
Steady State Probabilities Provided for N from 0 to 26
Waiting Time Probabilities Provided from 1 to 5 Minutes

Table 5. Steady State Probabilities

Number (N)	$P(X = N)$	$P(X < N)$	$P(X > N)$	
0	0.000	0.000	1.000	
1	0.000	0.000	1.000	
2	0.000	0.000	1.000	
3	0.000	0.000	1.000	
4	0.000	0.000	1.000	
5	0.000	0.000	1.000	
6	0.001	0.001	0.999	
7	0.002	0.004	0.996	
8	0.005	0.008	0.992	
9	0.009	0.017	0.983	
10	0.015	0.032	0.968	
11	0.022	0.054	0.946	
12	0.031	0.085	0.915	
13	0.040	0.125	0.875	
14	0.047	0.172	0.828	
15	0.052	0.224	0.776	
16	0.055	0.279	0.721	
17	0.053	0.332	0.668	
18	0.049	0.381	0.619	Being Served
19	0.046	0.427	0.573	Waiting
20	0.042	0.470	0.530	
21	0.039	0.509	0.491	
22	0.036	0.545	0.455	
23	0.034	0.579	0.421	
24	0.031	0.610	0.390	
25	0.029	0.639	0.361	
26	0.027	0.666	0.334	

Table 6. Probability of Waiting Time in the Queue Exceeding T

T	$P(WT > T)$
1	0.526
2	0.413
3	0.325
4	0.256
5	0.201

Average Number in System = 25.02
 Average Number in Queue = 8.35
 Average Time in System = 8.34
 Average Time in Queue = 2.78

tomed to accessing library materials. Many of them do not willingly embrace change and are skeptical of its value and hostile to its presence. In the face of such resistance,

it is important that library personnel neither make promises they cannot deliver nor create expectations that are unrealistic. System librarians need to make certain that library administrators understand the vulnerability of the system's performance to changes in certain key variables so that, if and when such changes conspire to degrade performance, no one in authority feels misguided. Such cautions are difficult to provide, however, if no one within a particular library has studied the relationships between the system's performance and the key variables measured for the institution. The writer hopes that this article will encourage and aid such study.

REFERENCES AND NOTES

1. John E. Tolle, *Public Access Terminals: Determining Quantity Requirements* (Dublin, Ohio: OCLC, 1984).
2. John E. Tolle, Nancy P. Sanders, and Neal K. Kaske, "Determining the Number of Online Catalog Terminals," *Information Technology and Libraries* 3:261-64 (Sept. 1983).
3. A. Whitney Knox and Bruce A. Miller, "Predicting the Number of Public Computer Terminals Needed for an Online Catalog: A

- Queuing Theory Approach," *Library Research* 2:95-100 (Spring 1981).
4. Christine L. Borgman and Neal K. Kaske, "Determining the Number of Terminals Required for an Online Catalog through Queuing Analysis of Catalog Traffic Data," as in *Public Access to Library Automation; 17th Clinic on Library Applications of Data Processing*, J. L. Divilbiss, ed. (Champaign-Urbana: Univ. of Illinois, 1981).
 5. Raymond G. Taylor, "Determining the Minimum Number of Online Terminals Needed to Meet Various Library Service Policies," *Information Technology and Libraries* 6:197-204 (Sept. 1987).
 6. Raymond G. Taylor, "Incremental Costs of Library Service Policies for Online Catalog Access," *Information Technology and Libraries* 6:305-10 (Dec. 1987).
 7. The program is written in BASIC for an IBM or compatible PC. As the program completes each group of calculations, it prints the results and releases memory for the next series of calculations. Therefore, memory requirements are not stringent. No peripheral equipment is required except a standard printer. Because the program contains many long calculations, however, a PC-AT may be preferred. The program assumes the user has a working knowledge of the queuing model; it requests a range of service and arrival rates and a beginning number of terminals. Terminals are then cycled from the beginning number down to five below that number. Other program parameters can be changed internally; they are identified by REM(arks). The program flags unreasonable combinations of parameter values, but it has not been thoroughly tested against such user errors.
 8. J. Gebbie and others, "Determining the Minimum Number of Terminals for an Online Catalog at a University Research Library," *Journal of Library Administration* (in press).
 9. Some combinations of arrival rate, service rate, and policy require an extremely large number of terminals. Further, some combinations are "impossible"; e.g., when arrivals exceed terminals times service rate. Throughout this article such combinations are referred to as "not meaningful." ■ ■

Twenty Years Ago in JOLA

Devising a plan for automating a library is not entirely unlike formulating a program for a library building. While there are general types of building best suited to the requirements of different types of library, each library is unique in some respects and requires a building which is especially designed for its own particular needs and situation. As there are no canned library building programs, so there are no canned library automation programs, at least at this stage of development; therefore, the first task of a library administration is to formulate an approach to automation based on a realistic assessment of the institution's needs and resources.

Richard DeGennaro, "The Development and Administration of Automated Systems in Academic Libraries," *JOLA* 1:1 (Mar. 1968), p.76.

Developing and Using the Online Catalog to Teach Critical Thinking

Joan Bechtel

Students are understandably bored with learning online catalog use when the focus is on the technicalities of the catalog itself rather than upon issues central to the educational experience. AutoCat, the online public access catalog developed at Dickinson College, was designed to force students to make informed choices as the search progresses. Teaching use of the catalog, therefore, focuses on the critical-thinking process necessary for formulating and researching a topic, an issue that is central to undergraduate education.

In the planning stages since 1982, AutoCat, Dickinson College's online public access catalog, is intended to serve the needs of faculty and students as they engage in the teaching/learning process. Rather than merely taking advantage of recent advances in computer technology and assisting librarians in organizing and making library materials more widely available, the central concern of the AutoCat team from the beginning has been serving the central educational purposes of the college.

EARLY DEVELOPMENT

In the mid-1970s librarians at Dickinson were feeling discomfort when they taught students to use *Library of Congress Subject Headings* as the key to unlocking the treasures behind the doors of traditional library organizational techniques and practices. The headings in the "red books" were increasingly out-of-date, and the Dickinson subject catalog, like those in undergraduate libraries across the country, defied their best efforts to update antiquated terminology. Consequently, the librarians turned to computer technology for assistance in pro-

viding improved subject access to the library's holdings.

In order to keep the focus sharply on the needs of students and faculty, the initial step in the development of the catalog was the articulation of a set of guiding principles, a set of benchmarks for testing the design as it developed. The operating principles defined in 1982 were that the system would be

- Simple to use and understand,
- Consistent in its command language and in computer responses,
- Flexible enough to provide both experienced and inexperienced users with appropriate freedom and control, and
- Helpful to all users no matter where they might be in their research or what types of problems they might encounter.

The goal of designing an online catalog to fit the needs of Dickinson undergraduates coincided with the mission of the library, articulated and refined over the past ten years. In order to serve the educational goals of the college, the eight librarians, organized collegially, have all shared in the work of technical and public services. Each

has worked in liaison with a group of four or five academic departments and has provided bibliographic instruction for the students in those departments. At the beginning of the online catalog project we knew very little about the intricacies of the MARC record and even less about computers. We did know well, however, the needs of faculty and students as they engaged in teaching and learning. Therefore, we began working on AutoCat at the point that we knew and understood best and cared about most, that is, we began with the information needs of undergraduates and not primarily those of librarians struggling to keep up with changing cataloging rules and geometrically increasing quantities of information. The initial principles of simplicity, consistency, flexibility, and helpfulness were articulated with Dickinson students and faculty specifically in mind.

CONCERN FOR BIBLIOGRAPHIC INSTRUCTION

While continuing their normal responsibilities in the library and serving on the AutoCat team, the librarians became more and more restless with traditional bibliographic instruction. We began to realize that improving access to library materials by automatic updating of LC subject headings and by adding access points for searching in the online catalog would be inadequate. Even those all-important guiding principles, faithfully used at every step in AutoCat development, were clearly not going to guarantee improved use of information by undergraduate students. In short, it became increasingly evident that the impediments to students' scholarly growth had more to do with their underdeveloped, flawed thinking processes than with their inability to find library materials.

Student papers were not magically improved by introducing the full range of library materials and teaching students to use the card catalog, bibliographies, and indexes more skillfully. The awful truth dawning on us was that the quantities of energy and money spent on online catalog development were going to miss the mark even if AutoCat were deemed successful in

the worlds of library/computer technology. To meet undergraduate students' needs, AutoCat would be required to encourage, perhaps even force, student growth in critical thinking. If AutoCat were to become an important educational tool, it would need to enhance the critical use of information rather than the mere finding of it.

PROBLEMS IN TEACHING UNDERGRADUATES TO USE EXISTING ONLINE CATALOGS

The need for an online catalog designed specifically for the undergraduate users is clearly reflected in two articles by Betsy Baker and one by Brian Nielsen, librarians at Northwestern University.¹ Both have been deeply involved in a project funded by the Council on Library Resources to develop a model for teaching the use of online catalogs, a model that might be useful to a wide range of libraries as they introduce automated catalogs to their users. The underlying assumption of the undertaking was that online catalogs have been and will be developed or purchased in order to improve access to library holdings. The task of librarians, then, will be to teach the capabilities and limitations of the new technology as well as the nitty-gritty of the keyboard manipulation by which it is operated. Concluding from their work that students and others at Northwestern University do not find learning to use the online catalog an "interesting endeavor in itself,"² Baker pointed to the need to couch instruction in conceptual rather than mechanical terms. Additionally she called for the design of "instructional programs that provide incentive for learning as well as providing needed skills for an automated society."³ Her solution to the double-sided problem of boredom and the need to communicate a conceptual framework was to expand the context of instruction beyond the image of the card catalog to the general principles of database management systems, record file structure, and other "building blocks for information retrieval systems,"⁴ because the online catalog would then "serve as an example of a particular implementation of general principles."⁵

On the other hand, Brian Nielsen stated, "it is important that in our teaching we not fixate on the online catalog at all as the object of instruction."⁶ Baker's approach, rather than turning from primary focus on the catalog itself, as her colleague Nielsen urged, forces the user more deeply into the inner workings of the online catalog. It is like asking a student who is bored with instruction in use of the card catalog to learn the detailed, multitudinous ALA filing rules because they may also assist in the use of phone books, encyclopedias, zip code directories, etc. While Baker's incentive may work for people who are fascinated by computer technology, it misses her own target: development of "a stronger educational stance that integrates online catalog instruction with other educational goals."⁷ It does so precisely because it continues to focus on the library tool—albeit now housed in shiny, new electronic technology—rather than upon the central educational process of the college or university.

Nielsen went on to suggest that in the future, librarians "transfer some of the energies away from teaching to get more directly involved in systems design."⁸ The clear message is that the difficulties in teaching online catalog use lie principally in the design and implementation of the catalog itself and not in the techniques of teaching its use. The early pioneers in online catalog development understandably concentrated on using the new technology to replicate the card catalog in order to improve bibliographic access; the library and academic worlds, therefore, are deeply in their debt. Second-generation online catalogs, however, must go beyond the models offered by traditional library tools and concepts. In their development the focus now must be on the needs of the users that are to be served. The work of Baker and Nielsen and the Dickinson experience make clear that it is no longer sufficient to invent new access points, faster machines, and ever more clever uses for keyword searching, Boolean logic, and authority control. In the academic setting the conceptual framework for development, and for teaching the use of the resulting catalog, must grow primarily out of the needs of students and faculty engaged in learning and research. In the case of undergraduate students, the

overriding need is for assistance with initiating and carrying out the critical-thinking processes required for addressing problems and issues during their college years and beyond.

AUTOCAT DESIGN

AutoCat has been designed with this particular goal in mind. Conceived with the strong sense that the old card catalog was hopelessly inadequate for providing subject access to the library's materials and shaped from the developing concern about enhancing students' critical facilities, AutoCat, Prototype III, was introduced to the campus in spring 1987. At that time the database contained MARC bibliographic records for 99 percent of the library's 200,000 books, scores, and sound recordings, as well as authority records in the MARC format for names, subjects, and phrases in subjects.

Based on the notion that conscious, enlightened choice is a major factor in thinking critically, AutoCat was designed to offer, even force, the user to make informed choices throughout the search process. Authority control makes possible initial review screens that reflect the variety of names or subjects appropriate to the keywords entered for both name and subject searches. Users must choose which particular name(s) or subject(s) they want before any specific books are presented. Furthermore, call number information is withheld until users decide exactly which materials they desire, based upon author, subject, and title information as well as the age and format of the material. Knowing full well that some students, pressed for time, settle for something, anything, on a topic and that the more careful students search for everything on the topic, the librarians wanted the system to encourage discriminating, thoughtful choice as a first step in thinking critically about any given topic. This process of decision making is illustrated in the following searches performed in response to typical undergraduate term-paper assignments.

USING AUTOCAT TO TEACH CRITICAL THINKING

In the first instance the students in an introductory geology course were asked to

write a fifteen to twenty-page paper on the storage of waste materials. As is often the case with undergraduate term paper assignments, the topic was unfamiliar to the students and was actually, in scope, more a topic area than a specific term paper topic. The first task of the students, therefore, even though they may not have known it, was to narrow their focus to a topic that could be treated adequately in a paper of the assigned length and in the time normally expected for such an assignment.

In this situation, the student did not know exactly which terms would be used in subject headings appropriate for the topic: storage of waste materials. The librarian, therefore, suggested that she search for words from the topic in the title index, in the hope that there might be a book with those keywords in the title. If so, the student would then have not only one book related to the topic area but also could see on the screen the correct subject heading or headings that would lead to further material. This approach to subject searching combines the advantages of citation indexing, based upon current terminology that researchers are actually using, and the full power of controlled terms used in subject headings.⁹

Initially the student used the command *TI* (for title index) and the keywords *storage* and *waste* (see figure 1). The search resulted in one book, *Nuclear Waste: Socio-*

economic Dimensions of Long-term Storage (see figure 2). Glancing at the subject headings, the student discovered that the subject heading terms appropriate for a book on the storage of waste materials are *waste disposal* rather than *waste storage* (see figure 3).

Having discovered the appropriate terms, the student returned to the search request screen and entered *SU* (for subject heading index) and the keywords *waste* and *disposal*; the Boolean *AND* is assumed.¹⁰ The result was an authority review displaying all the subject headings in the database that contain the words *waste disposal* along with the number of bibliographic records attached to each heading (see figures 4 and 5).

Reading the full array of headings that appeared in response to the request for information concerning waste disposal, the student discovered that the topic is very broad indeed, since 39 percent of the database produced sixteen subject headings! Clearly, she could write about the economic or environmental aspects of storing radioactive waste. She could investigate storage of radioactive wastes in particular geographic areas or geologic formations. Furthermore, she could choose to research wastes from the chemical, food, or nuclear industries. In its subject authority review, *AutoCat* clearly demonstrated that the student could not possibly treat the storage of waste materials adequately in one term paper; a particular as-

```

-----AutoCat Prototype-----
                          Command Search Screen

Enter your search request below and press RETURN.  Example: AU JOHN STEINBECK

Search request:

su waste disposal

Primary indexes      Secondary indexes      Valid booleans
AU - Author          BL - Branch Library    AND - narrows search
CO - Contents        FO - Format             BUTNOT - excludes
SE - Series          LA - Language          OR - Expands search
SU - Subject
TI - Title

Useful keys:  PF2 - provides help
              Up-arrow - recalls previous search request
              CTRL/U - erases current search request
              PF3 - stops the search in progress
  
```

Fig. 1. Command Search Screen

```

-----AutoCat Prototype-----
                          Review Display

-----Author-----      -----Title-----      Date Form
1. [Multiple authors]   Nuclear waste : socioeconomic dimensions  1983 Bk

```

Please select one or more line numbers: 1

Fig. 2. Review Display

```

-----AutoCat Prototype-----
                          Full Display

1. 363.728 N9642

Nuclear waste : socioeconomic dimensions of long-term storage / edited by
Steve H. Murdock, F. Larry Leistritz, Rita R. Hamm. Boulder, Colo. :
Westview Press, 1983.
  xxi, 343 p. : ill. ; 24 cm.
Includes index.
SERIES: Westview special studies in science, technology, and public
policy/society

Subjects: 1. Nuclear industry--Waste disposal--Government
policy--United States
2. Nuclear industry--Waste disposal--Economic
aspects--United States
3. Nuclear industry--Waste disposal--Environmental
aspects--United States
4. Radioactive waste disposal in the ground--
Government policy--United States

Press PF3 to return to the review display or any other key to continue

```

Fig. 3. Full Display

```

-----AutoCat Prototype-----
                          Command Search Screen

Enter your search request below and press RETURN. Example: AU JOHN STEINBECK

Search request
vi storage waste

Primary indexes      Secondary indexes      Valid booleans
AU - Author          BL - Branch Library   AND - narrows search
CO - Contents        FO - Format            BUTNOT - excludes
SE - Series          LA - Language         OR - Expands search
SU - Subject
TI - Title

Useful keys: PF2 - provides help
              Up-arrow - recalls previous search request
              CTRL/U - erases current search request
              PF3 - stops the search in progress

```

Fig. 4. Command Search Screen

-----Autocat Prototype-----		
	Subject Heading Review	Percent searched 39
Cnt	-----Subject Heading-----	
1.	1 Atomic power-plants--South Carolina--Aiken--Waste disposal	
2.	1 Atomic power-plants--Sweden--Waste disposal	
3.	1 Atomic power-plants--Waste disposal	
4.	1 Chemical industries--United States--Waste disposal	
5.	1 Chemical plants--Waste disposal--Hygienic aspects--New York	
6.	1 Chemical plants--Waste disposal--Social aspects--New York	
7.	1 Chemicals--Manufacture and industry--Waste disposal--Handbo	
8.	1 Food industry and trade--Waste disposal	
9.	1 Nuclear facilities--Waste disposal--Congresses	
10.	1 Nuclear industry--Waste disposal--Economic aspects--United	
11.	1 Nuclear industry--Waste disposal--Environmental aspects--Un	
12.	1 Nuclear industry--Waste disposal--Government policy--United	
13.	3 Radioactive waste disposal in rivers, lakes, etc.	
14.	3 Radioactive waste disposal in the ocean	
15.	7 Radioactive waste disposal	
16.	1 Radioactive waste disposal in the ground	
17.	1 Radioactive waste disposal in the ocean--Bibliography	

Press PF3 to stop search or any other key to continue

Fig. 5. Subject Heading Review

pect of the topic needed to be chosen. Furthermore, the subject heading review screen provided the information needed to make such a choice. This particular student decided that storage of waste materials in the water, that is, in rivers, streams, and oceans, interested her most. Using the Boolean OR, she added the terms *rivers*, *ocean*, *water* to the search so that the results would more exactly match this newly refined topic (see figures 6 and 7).

The student now had a much shorter list from which to choose; and, the books were all related to the topic she had chosen. The search process thus far had done three things for her: it assisted in narrowing the focus; produced several monographs; and, in the bibliographies, provided the tools for further search. The student next decided to secure the call number for the bibliography covering radioactive waste disposal in the ocean (see figure 8). Returning to the subject heading review display, she subsequently chose to look at the three books with the same heading (see figures 9 and 10).

The resulting display contained author, title, date, and format information for each book, making it very clear that the student had retrieved one book containing fairly recent material as well as two books that, in the field of nuclear energy and research, were quite old. If the student wished to include information on the history of the

topic, she could use the older material. If, on the other hand, she decided to focus on the most recent developments in the field, the older material would have been irrelevant. With this information at hand, the student had the necessary knowledge for making an informed decision about which material she wanted.

The second student search was performed in response to the syllabus for a comparative civilizations course that specified the final assignment in the course: a twenty-five-page research paper on China. In the past students have typically gone to the *C* drawer in the subject catalog to find information on such a topic. In this example, the student turned to AutoCat, entered the subject term *China*, and began to get the list of subject headings containing the term (see figures 11 and 12).

The student stopped the search when sixteen headings, all beginning with *A*, appeared and only 1 percent of the database had been searched. The AutoCat subject review display revealed that there are hundreds of subject headings containing the term *China*, making clear that the student had to refine the focus of the paper. He could do so by adding terms to the search, using the Boolean operators or secondary indexes. In fact, scanning just these few headings suggested several possible term paper topics to the student. Foreign rela-

```

-----AutoCat Prototype-----
                          Command Search Screen

Enter your search request below and press RETURN.  Example: AU JOHN STEINBECK

Search request:
SU rivers or ocean or water and waste disposal

Primary indexes      Secondary indexes      Valid booleans
AU - Author          BL - Branch Library   AND - narrows search
CO - Contents        FO - Format            BUTNOT - excludes
SE - Series          LA - Language         OR - Expands search
SU - Subject
TI - Title

Useful keys:  PF2      - provides help
               Up-arrow - recalls previous search request
               CTRL/U   - erases current search request
               PF3      - stops the search in progress

```

Fig. 6. *Command Search Screen*

```

-----AutoCat Prototype-----
                          Subject Heading Review      Percent searched 100

Cnt  -----Subject Heading-----
1.   3  Radioactive waste disposal in rivers, lakes, etc
2.   3  Radioactive waste disposal in the ocean
3.   1  Radioactive waste disposal in the ocean--Bibliography
4.   1  Radioactive waste disposal in rivers, lakes, etc.--Bibliogra
5.   2  Radioactive waste disposal in the ocean--Congresses
6.   1  Radioactive waste disposal in rivers, lakes, etc.--Congresse
7.   3  Waste disposal in the ocean--Congresses

```

Please select one or more line numbers: 3

Fig. 7. *Subject Heading Review*

```

-----AutoCat Prototype -----
                          Full Display

3.  621.4838 I61      Physics

International Atomic Energy Agency.
  Disposal of radioactive wastes into marine and fresh waters.  Vienna :
International Atomic Energy Agency, 1962.
  365 p. ; 24 cm.
  Includes index.
  SERIES: Bibliographical series - International Atomic Energy Agency; no. 5

  Subjects: 1. Radioactive waste disposal in the ocean--
             Bibliography
            2. Radioactive waste disposal in rivers, lakes,
             etc.--Bibliography

```

Press PF3 to return to the review display or any other key to continue

Fig. 8. *Full Display*

```

-----AutoCat Prototype-----
                          Subject Heading Review          Percent searched 100
-----Subject Heading-----
Cnt -----
1.  3 Radioactive waste disposal in rivers, lakes, etc
2.  3 Radioactive waste disposal in the ocean
3.  1 Radioactive waste disposal in the ocean--Bibliography
4.  1 Radioactive waste disposal in rivers, lakes, etc.--Bibliogra
5.  2 Radioactive waste disposal in the ocean--Congresses
6.  1 Radioactive waste disposal in rivers, lakes, etc.--Congresse
7.  3 Waste disposal in the ocean--Congresses
    
```

Please select one or more line numbers: 2

Fig. 9. Subject Heading Review

```

-----AutoCat Prototype-----
                          Review Display
-----Author-----Title----- Date Form
1. [Multiple authors] Radioactive wastes and the ocean / edit 1983 Bk
2. Commission of the Eur Applications of a derived formula for t 1973 Bk
3. Commission of the Eur Principles and methods for the derivati 1971 Bk
    
```

Please select one or more line numbers:

Fig. 10. Review Display

```

-----AutoCat Prototype-----
                          Command Search Screen
-----
Enter your search request below and press RETURN. Example: AU JOHN STEINBECK

Search request
Su China

Primary indexes      Secondary indexes      Valid booleans
AU - Author          BL - Branch Library    AND - narrows search
CO - Contents        FO - Format             BUTNOT - excludes
SE - Series          LA - Language          OR - Expands search
SU - Subject
TI - Title

Useful keys: PF2 - provides help
              Up-arrow - recalls previous search request
              CTRL/U - erases current search request
              PF3 - stops the search in progress
    
```

Fig. 11. Command Search Screen

-----AutoCat Prototype-----		Percent searched 01
Subject Heading Review		
Cnt	-----Subject Heading-----	
1.	1 Adoption--China	
2.	1 Africa--Foreign relations--China	
3.	1 Africa--Relations (general) with China	
4.	1 Aged--China--Shen-yang shih--Social conditions	
5.	1 Aged--China--Social conditions	
6.	1 Aged--Government policy--China	
7.	8 Agriculture--China	
8.	1 Agriculture--China--History	
9.	1 Agriculture--China--History--Sources	
10.	11 Agriculture--Economic aspects--China	
11.	1 Agriculture--Economic aspects--China--Szechwan	
12.	1 Agriculture--Economic aspects--China--Yunnan (Province)	
13.	4 Agriculture and state--China	
14.	1 Agriculture and states--China	
15.	1 Agriculture and state--China (People's Republic of China, 1949-)	
16.	1 Agriculture, Cooperative--China--Case studies	

Please select one or more line numbers: 11

Fig. 12. Subject Heading Review

-----AutoCat Prototype-----			
Review Display			
-----Author-----	-----Title-----	Date	Form
1. [Multiple authors]	Institutional reform and economic develop	1984	Bk
2. Chao, Kang, 1929-	Agricultural production in Communist Chin	1970-	Bk
3. Ch'en, Hang-sheng,	The present agrarian problem in China / b	1933	Bk
4. Ch'en, Hang-sheng,	Landlord and peasant in China ; a study o	1936	Bk
5. Dawson, Owen L.	Communist China's agriculture; its develo	1970	Bk
6. Hsu, Cho-yun	Han agriculture : the formation of early	1980	Bk
7. Institute of Pacifi	Agrarian China : selected source material	1978-	Bk
8. Lippit, Victor D	Land reform and economic development in C	1974	Bk
9. Perkins, Dwight Hea	Agricultural development in China, 1368-1	1969	Bk
10. Wittfogel, Karl Aug	Agriculture: a key to the understanding o	1971	Bk
11. Wong, John.	Land reform in the People's Republic of C	1973	Bk

Please select one or more line numbers:

Fig. 13. Review Display

tions, aging in China, and aspects of agriculture in China were all represented in this display. The student was interested in farming and consequently requested more information on the eleven books listed under Agriculture—Economic aspects—China. The result is shown in figure 13.

Note that while no call numbers were presented, the student did have the information needed to make some important decisions. The first item listed showed [Multiple authors] in the author column, indicating that the book probably contains a variety of viewpoints, an issue of major importance when doing a piece of research.

Also there were at least three, perhaps four, Chinese names listed, indicating that there may be some important, primary source material listed here. Title 3, *The Present Agrarian Problem in China*, might have caught the eye of the student and suggested a possible topic. However, a quick glance at the date of the book serves notice that the title, appropriate in the 1930s, is misleading now. Reviewing the titles and dates listed, the student decided that he could write a paper comparing the agrarian situation before and after the revolution. The titles, dates, and authors listed on this one screen appeared to indicate sufficient ma-

terial, varying viewpoints, primary sources, and a range of publication dates appropriate for such a topic.

As exemplified in these searches based on typical undergraduate assignments, AutoCat is designed to encourage students to use the search process itself as a tool for improving the quality of their thinking on any given topic. As their faculty mentors already know, developing one's own position on an issue is dependent upon shaping an appropriate question and upon discovering and evaluating the variety of perspectives held by others. To grow in their critical abilities students must begin to ask questions about who holds which opinion based on what data gathered at what period of time. In its displays and computer responses, AutoCat is designed to encourage students to choose from a multitude of related topics and subtopics within a subject area, to discriminate between primary and secondary source material, to discover multiple viewpoints, and to make initial decisions about the appropriateness of materials based upon knowledge of authors, subject, date, and format. Teaching its use, therefore, necessarily focuses on issues cen-

tral to the educational process and not merely on the mechanics of using the online catalog.

CONCLUSION

It is not surprising that librarians at Northwestern University and elsewhere have found that students are bored with learning the technicalities of online catalog use. Students have always been less than excited about learning how to use the card catalog, periodical indexes, etc. when library organization was taught as an end in itself. Such learning is not only tedious but also distracting and time-wasting, as it draws attention away from the central concerns of students and faculty engaged in education. The development team believes that students will find the incentive to understand AutoCat so deeply embedded in its contribution to the thinking, reasoning, and learning processes that mastering the details of its use will become a nonissue. The highest hopes of the development team will have been met when students find their intellectual growth stimulated by their use of AutoCat.

REFERENCES AND NOTES

1. Betsy Baker, "A Conceptual Framework for Teaching Online Catalog Use," *Journal of Academic Librarianship* 12, no.2:90-96 (1986); "A New Direction for Online Catalog Instruction," *Information Technology and Libraries* 5, no.1:35-41 (Mar. 1986); Brian Nielsen, "What They Say and What They Do: Assessing Online Catalog Use Instruction through Transaction Monitoring," *Information Technology and Libraries* 5, no.1:28-34 (Mar. 1986).
2. Baker, "A Conceptual Framework," p.94.
3. *Ibid.*
4. Baker, "A New Direction," p.40.
5. *Ibid.*
6. Nielsen, p.33.
7. Baker, "A New Direction," p.39.
8. Nielsen, p.34.
9. In the November 1986 issue of the *Journal of Academic Librarianship*, Alexis J. Jamieson, Elizabeth Dolan, and Luc Declerck state that "there is a need for detailed studies investigating the relative value of keyword searching versus authority control" (p.277). They conclude that keyword searching, although very powerful, does not by itself satisfy the needs of users. "A catalog without a cross-reference structure for variant forms of names and subject headings will give its users inferior service" (p.283). We could not agree more with the writers' conclusions. The phrasing of the question in terms of the "relative value of keyword searching versus authority control," however, suggests that the choice in an online catalog is either keyword searching or authority control. If one settles for keyword searching only, the traditional gathering and finding capacities of the card catalog are lost. If, on the other hand, authority control without keyword searching is provided, the online catalog is little better than cards for locating information. Computer technology is fully capable of providing both keyword searching and a syndetic structure. The notion that one must choose either keyword searching or authority control because they are functionally or intellectually incompatible is false. Furthermore, choosing one or the other in order to conserve resources is false economy.
10. The full AutoCat, expected within the aca-

demarc year, will include a "term" search that combines in one index keywords from the title and all subject headings associated with a sin-

gle work. It will then also be possible to initiate a new search at any point, so that it will be unnecessary to return to the search screen. ■■

Twenty Years Ago in JOLA

Through a program of mutual education, the librarian learned of the computer and what it could do and what it could not do; and systems and computer personnel learned of the library's requirements and desires. There evolved the basic design for a system capable of being implemented on the equipment at hand and acceptable to the library.

Richard D. Johnson, "A Book Catalog at Stanford," *JOLA* 1:1 (Mar. 1968), p.15.

Laser Printing for a Variety of Library Applications

Glen J. Kelly

Laser printers currently range in price from \$1,995 to \$166,000, but many libraries will find that lower-priced models in the \$3,000 to \$6,000 range are suitable for almost all printed products. Most library software packages running on either a personal computer or a mainframe can be modified to accept laser printers, which are similar in several ways to photocopy machines, in that they both use single sheets of bond paper in either 8½-by-11-inch or 8½-by-14-inch sizes. The paper stock can be preprinted, and call number and bar code label stock are available for laser printers. Libraries requiring special, preprinted form stock can now design their own from regular paper stock. The high quality, printing speed, low noise, affordable maintenance costs, and ease of use make laser printers ideal for many library printing applications from call number and bar code labels to purchase orders and book catalogs. The printer's inability to produce multiple-copy forms is a weakness that may no longer be a concern for libraries that store their files in machine-readable form and don't wish to maintain additional printed copies in manual files. Files can be printed or reprinted on demand, as need dictates, making laser printing a wise choice.

Laser printing technology, formerly affordable to only the largest of library systems or consortia, has now come within the price range of most libraries. The consistent high quality, low noise levels, reliability, and ease of both use and maintenance are features of laser printing that will provide most libraries an attractive alternative to the current generation of dot matrix and daisy wheel printers. There are some weaknesses, however, of which a library should be aware before purchasing a laser printer. The attempt in my analysis is to cover the following topics: current trade literature on laser printers, hardware and technical considerations, software and product considerations, measurement and operational

considerations, ease of use and maintenance, and capital and operating costs.

CURRENT TRADE LITERATURE ON LASER PRINTERS

This literature search is highly selective, as I have chosen only current information that I felt would be pertinent to this article. In researching the trade literature, I located fifty-one laser printer manufacturers, and the fairly comprehensive listing includes price, printing speed, and other technical information on eighty-seven models.¹ Prices range from \$1,995 for 6-to-8-page-per-minute printers such as the QMS KISS and Laser Pro Express models to

Glen J. Kelly is Chairman, Library Technical Services and Automation Project Coordinator, Laurentian University Library, Sudbury, Ontario.

\$166,000 for faster, 125-page-per-minute machines such as the Vision System from Anser Technology. Machines on the market that are designed specifically for, say, Apple or Digital Corporation's own equipment are excluded. The listing does include laser printers that can be used, or adapted for use, not only with IBM hardware but also with other brands of personal and mainframe computers.

An interesting review of four LaserJet Plus compatibles can be found in the June 1987 issue of *PC World*. The authors compare and rate the machines on a variety of factors including the quality of text printing, ability to print both bit-mapped and object-oriented graphics, paper handling, and overall value. The models tested were the Oasys Laser Pro 810-R, Mannesmann Tally MT910, Kyocera F1010, and Quadram QuadLaser, all of which emulate the popular LaserJet. With over 250,000 LaserJets sold, Hewlett-Packard has rapidly gained the largest portion of the lower-priced desktop laser printer market; most software programs have print drivers that include it. The models tested were acceptable alternatives to the LaserJet Plus and had some features that were more attractive; however, the authors concluded that none of the compatibles excelled in every category and that buyers should be aware of the differences before purchasing a laser printer.²

Although there are minor differences in ease of use, technical specifications, maintenance routines, and price, most features and operational characteristics are common to a wide variety of desktop laser printers that emulate the LaserJet. The remainder of this article will focus on the specific features of the Hewlett-Packard LaserJet that was tested.

HARDWARE AND TECHNICAL CONSIDERATIONS

Laser printers are currently available in desktop size (28 by 20 by 11 inches in height). They resemble and possess many of the same physical and operational characteristics of small desktop photocopiers. Laser printers use toner cartridges, paper trays, fusing assemblies, and transfer corona wires, features similar to most dry

photocopy machines, but there is one notable difference: no visible light source can be seen when a laser printer produces a printed page. A photocopier takes a picture of a document, which is then transferred to the corona wire and fusing assembly, whereas a laser printer receives from a computer a digital signal, which is transferred to the corona wire and fusing assembly that produces the printed page. Laser printers are capable of producing high-resolution, dot matrix patterns of more than 300 by 300 dots per inch with the same consistent, high-quality resolution for 3,000 or more pages, the average life cycle of a toner cartridge.

During the printing cycle almost no noise can be heard over the normal background noise of the laser printer's fan motor, an important consideration for many libraries. The noise level during the printing cycle is low enough that no special acoustical damping covers are required. The speed of laser printers is measured in *pages per minute* and varies with the model purchased. The average laser printer is rated at 8 to 10 pages per minute, or *300 characters per second*, assuming 1 page equals 2,000 characters. This printing speed is almost twice as fast as a dot matrix printer rated at 150 characters per second in letter-quality mode and ten times as fast as the current generation of daisy wheel printers. The resident memory buffer in the printer tested was 512 kilobytes of internal memory, or approximately 10 pages of printing. This can sometimes be a significant factor during large print jobs that can tie up a computer as it waits for the printer to finish before sending more output.

Laser printers can communicate with either a mainframe or a personal computer through a standard RS 232 serial interface or a parallel printer interface. The model I tested on a Digital VAX 8530 mainframe was a standard LaserJet, which provides configuration instructions for a wide variety of personal computers including DEC Rainbows, Apples, IBM PCs and IBM compatibles. The model tested could print either horizontally from left to right across the width of a page or vertically from left to right down the length of a page. Four standard paper cassette trays were available for

the following sizes of stock: 8½ by 11 inches; 8½ by 14 inches; A4 (210 by 297mm); and B5 (192 × 257mm). The cassettes on the model tested held approximately 100 sheets at a time. Other models of laser printers are available with optional dual paper cassettes for increased capacity and sorting features that stack the output face up or face down (correct order for list printouts). Adjusting the darkness of the printed output is simply a matter of turning a dial, instead of having to replace a ribbon as would be the case with an impact printer.

The laser printer tested had two resident fonts; in addition, twenty-one font cartridges are available, each having a variety of styles and sizes. The font cartridge tested included OCR; Code 39 for bar code labels; 14 point, which was used for call number spine labels; and 9.5 point, for printing applications requiring smaller-sized characters.

A font cartridge costs \$350 and can easily be inserted or removed from the front panel of the laser printer. The special fonts contained on a cartridge must be accessed by special commands in the software being used; otherwise, all print jobs default to the resident printer font.

SOFTWARE AND PRODUCT CONSIDERATIONS

After a font cartridge is inserted, special printing commands from the software override the resident fonts. The setup string of printer escape sequences required to turn certain fonts on or off are completely controlled by the application program. Although examples of printer escape sequences contained in the instruction manuals are straightforward and easy to copy, some expertise is required to set a nonresident font via the application program being used. This procedure is *not* for the amateur or the faint at heart and requires a good knowledge of the application program's printer escape sequence routines. Once the procedure is completed, and as long as no one changes the printer escape codes in the application program, no further expertise is needed to choose the desired font from the program.

One rather annoying problem may arise

when using a mainframe that will require special assistance from an operating system programmer.³ The resident font in the LaserJet printer, known as Roman-8 character set, or USA ASCII plus Roman extension, contains a foreign character set that works well with most personal computers. Foreign characters are printed on the basis of a character map that is set for personal computers.

The mapping of the foreign character set did not match the ASCII code used by our mainframe. For example, the VAX 8530 hexadecimal code for a lowercase *e* acute accent (*é*) is E9; however, the laser printer recognized E9 as an uppercase *O* tilde and printed *Ö* instead of *é* (see figure 1).⁴ To overcome this problem, one could choose to

1. purchase the same brand of laser printer as the output-producing computer, whose manufacturer would have the correct map of the foreign character set built into a chip inside its brand of laser printer;
2. modify the application software output codes;
3. write a translation program in the print routine for the ASCII foreign characters; or
4. change the print symbiont modification routine in the mainframe's operating system software.

In searching for a solution to the foreign-character mapping problem, a translation program was written that was called from the print routine each time a product requiring a foreign character set had to be printed. This solution was not desirable, as it required extensive additional modifications to the application program's print subroutines. The "translate" program had to be accessed before printing a report, and building this procedure into the application software would have required extensive programming effort.

Modifying the print symbiont routines supplied in the latest versions of Digital Corporation's VAX/VMS operating system software required less programming effort and provided an ideal solution to the problem. The operating system software contains print symbiont routines controlling all output sent to the various printers connected to the VAX 8530. Some earlier versions of the VAX/VMS operating system

Roman-8																		
	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F		
0	NUL	DLE	SP	0	@	P	'	p					-	á	À	Á	Þ	
1	SOH	DC1	!	1	A	Q	a	q					Â	ê	Ë	ä	þ	
2	STX	DC2	"	2	B	R	b	r					Ã	ë	Ü	å		
3	ETX	DC3	#	3	C	S	c	s					È	°	ú	Æ	Ð	
4	EOT	DC4	\$	4	D	T	d	t					É	Ç	á	à	ð	
5	ENQ	NAK	%	5	E	U	e	u					Ê	ç	é	í	ï	
6	ACK	SYN	&	6	F	V	f	v					Ë	Ñ	ó	ø	ì	-
7	BEL	ETB	'	7	G	W	g	w					Ì	ñ	ú	æ	Ó	¼
8	BS	CAN	(8	H	X	h	x					Í	ï	à	Ä	Ö	½
9	HT	EM)	9	I	Y	i	y					Î	è	ì	Õ	*	
A	LF	SUB	*	:	J	Z	j	z					Ï	é	ò	Ö	ó	²
B	VT	ECS	+	;	K	[k	{					Ï	ù	Û	Š	«	
C	FF	FS	,	<	L	\	l						Ï	ÿ	ä	É	š	■
D	CR	GS	-	=	M]	m	}					Û	š	ë	Ï	Ú	»
E	SO	RS	.	>	N	^	n	~					Û	ſ	ö	ß	ÿ	±
F	SI	US	/	?	O	_	o	DEL					ÿ	ç	ü	Ô	ÿ	

Fig. 1. Roman-8 Symbol Set

*** m u l t i L I S ***

87-03-30 Page: 4

Villefranche
Author Catalogue

- Villeneuve, Jocelyne, 1941-
La saison des papillons : recueil de poèmes rédigés à la
façon des haïkai : suivi de Propos sur le haïkai /
Jocelyne Villeneuve.
Sherbrooke, Qué. : Editions Naaman., 1980.
PS9593I43S35 UCC
- Villeneuve, Jocelyne, 1941-
La princesse à la mante verte / Jocelyne Villeneuve ;
[dessins de Luc Robert].
Sudbury, Ont. : Prise de parole., 1983.
PS9593I43P7 UCC
- Villeneuve, Jocelyne, 1941-
Nanna Bijou : le géant endormi / Jocelyne Villeneuve.
Sudbury, Ont. : Prise de parole., 1981.
PS9593I43N3 UCC
- Zaslow, Morris.
The opening of the Canadian North, 1870-1914 / Morris
Zaslow.
Toronto, Ont. : McClelland and Stewart., 1971.
FC3217Z38 UCC

Fig. 2. Example of Printed List with Foreign Characters

software do not allow modification of the print symbiont routines. A macro program was written, modifying the print symbiont routine in a manner whereby the foreign characters are translated into the correct hexadecimal codes *before* the output is sent to the printer. By changing the print symbiont routine before the output is sent to the printer, the correct hexadecimal codes for the foreign character set are always sent in a form that the LaserJet will recognize. Figure 2 is an example of a laser-printed list containing foreign characters.

MEASUREMENT AND OPERATIONAL CONSIDERATIONS

Pitch and point size are two measurements that all printing output has in common. Generally, librarians are familiar with pitch, namely, the number of characters in a *horizontal* inch of text. Most fonts are fixed pitch: the higher the number, the more characters in an inch of text. Point size is a measurement unit equal to $1/72$ inch in height. Point size is the *vertical* height of a character. The higher the unit of measurement in point size, the larger the character, and conversely, the lower the unit of measurement, the smaller the character. What is important to remember is the difference between the units of measurement because they signify exactly the opposite: 14-point characters cannot be printed in 14 pitch because they are too wide; however, they can be printed in 10 pitch. Point and pitch are important considerations when program output is required on preprinted forms or label stock.

Another operational consideration when using a laser printer is the high temperature produced during the fusing process. Colored paper, preprinted forms, and letterhead can be used as long as they were produced with heat-resistant inks that withstand temperatures of 190° C (374° F) without melting.⁵ Care must be taken with certain types of envelopes, as their adhesive flaps may seal during the fusing process. Also, some envelopes may wrinkle or crease and cause feeding problems in a laser printer.

CALL NUMBER SPINE LABELS

The aforementioned aspects of measurement and high temperatures were particularly important in finding a laser-printed solution to producing call number spine labels. Formerly, a daisy wheel printer had been used, but the slow speed, noise, quality of the spine labels, and number of printer ribbons that had to be changed were factors that made the decision to purchase a laser printer easier.

Most libraries were, and to my knowledge still are, producing call number spine labels on impact printers, either daisy wheel or dot matrix. A few libraries are using laser printers to produce their own bar code labels, and the very large libraries and consortia are using high-speed laser printers in the price range of \$60,000 to \$166,000 to produce cards and other products, including spine labels.

From the outset we asked ourselves whether it would be possible to use *one* printer to produce spine labels, purchase order forms, *and* regular printed lists. Call number spine labels had to be produced in 14-point characters large enough to be read easily on the shelves, and excellent printing resolution was needed. A special font cartridge was purchased, capable of printing 14-point characters.

Obtaining call number label stock in required dimensions (1 by 1½ inches), on single sheets that could feed easily through the laser printer's feed mechanism, was not a simple task.⁶ Spine labels on the former label stock were arranged on a continuous, tractor-feed carrier sheet made of wax paper, rather than regular bond, to allow for the easy removal of labels. In attempting to use some of the former label stock by cutting sheets to the correct size (8½ by 11 inches), we found that the sheets did not feed properly, and we were afraid to damage the laser printer's drum in the fusing process, as the wax-paper carrier sheet began to melt.

Laser printers require that all labels be arranged on a single 8½-by-11 or 8½-by-14-inch carrier sheet made from bleached sulfate stock and silicone-coated for easy removal.⁷ Adhesive labels must be arranged

on the carrier sheet so that they cover the entire page with no spaces in between. Labels are arranged with a margin around the outside edges of the paper so that the labels will not peel off during the printing process.

Only acrylic-based adhesives are used, because they are stable at the high temperatures encountered during the fusing process. The adhesive properties of the labels tested seemed to be as good as the former label stock; both types of stock must be covered with clear tape to prevent the labels from coming off certain types of book bindings. The final suitability test for any label or other preprinted stock should always involve a sample in the laser printer; if it jams the feeding mechanism, begins to burn the ink, or melts the carrier sheet, a different paper product or label stock obviously must

be used. Figure 3 illustrates a single sheet of call number labels produced with the laser printer tested.

PURCHASE ORDER FORMS

Purchase order forms were formerly preprinted commercially and produced in six-part, continuous, tractor-feed form stock. Five copies were used to maintain manual files, which were discontinued when we purchased a fully integrated online system with ordering, cataloging, circulation, and public access modules. Only one copy of the purchase order form is now produced and sent to the vendor. Claim forms are produced in a separate print run. We approached our major vendors and asked whether they preferred an individual purchase order form for each title or a printed list of ordered titles. They indicated that,

BF 199 D45 1984	D 1055 E795 1984	D 421 S277 1972	D 779 C2 W37 1985	DA 426 F73 1975	DA 566.9 C5 M26 1984	DA 585 C5 D54 1984	DA 714 J585 1984
DF 234.2 F38 1985	DG 552 A2 058	DL 380 G34 A37	DP 143 P34 1973	DS 779.15 C476	DS 889 A986 1972B	DT 87 L35 1985	E 99 M87 S38
F 2385 H66 1985	FC 149 B45 1986	FC 3095 N67 P64 1985	FC 620 G72 1986	G 419 B76 1983	G 73 R25	HC 104 M38 1975	HD 2932 M845 1986B
HD 4165 L47 1985	HG 229 B6 183	HJ 2056 N6 G67 1984/85	HJ 7784 D44 1983	HN 110 Q7 O42 1986	HV 568 B64 1978	HV 8203 F345 1985	HX 44 B73
JN 1129 T7 P8 1985	JN 2961 Y76 196	JN 5448 V47 1970	JX 1393 N54 V39 1983	KE 3098 S63 1986	KE 3108.5 L33 1986	KE 3113 W45 1986	KE 8808.5 S36 1985

Fig. 3. Call Number Spine Labels From a Laser Printer

manually, it would be easier for them to match a single purchase order form with a book than to match many books with a purchase order list.

Figure 4 illustrates a purchase order sheet produced using the laser printer. The form is printed in 9.5 point and designed in such a manner that the vendor's address appears on the one-part form and can be seen in a window envelope without having to re-address the envelope. Two purchase order forms are printed on each page of regular bond paper. They are cut along the broken line, using a paper cutter, and inserted into a window envelope for mailing.

EASE OF USE AND MAINTENANCE

Font cartridges are easily installed, and

disk-based soft fonts can be downloaded from most popular brands of personal computers to some models of laser printers. A program is available that permits production of one's own bar code labels or replicate copies of lost bar codes, using a laser printer, a personal computer, and the correct bar code label stock.⁸

The procedures involved in the normal maintenance routine for a laser printer are easy to learn, even for an inexperienced operator. The procedures involve replacing toner cartridges and separation belts and cleaning the fusing roller and transfer corona wires as required. Paper jams are cleared by simply lifting one release lever, which allows complete access to the interior of the machine. One word of caution, however: the fusing assembly is extremely

PURCHASE ORDER - BON DE COMMANDE
Laurentian University Library
Bibliothèque de l'Université Laurentienne
Sudbury, Ontario
Canada P3E 2C6

Order no - no de commande : 104
Quantity - Quantité : 1

Title - Titre : Casebook series.
Publisher - Éditeur : London : Arnold.
Year - Année : 1986.

Blackwell's, B.H. Limited,
Broad Street,
Oxford OX1 3BQ,
England.

ATTENTION : 1 vol.(1 ex.)

Account No. 29501377/500

PURCHASE ORDER - BON DE COMMANDE
Laurentian University Library
Bibliothèque de l'Université Laurentienne
Sudbury, Ontario
Canada P3E 2C6

Order no - no de commande : 106
Quantity - Quantité : 1

Title - Titre : Letters of Charles Dickens.
Publisher - Éditeur : Oxford : Clarendon Press

Blackwell's, B.H. Limited,
Broad Street,
Oxford OX1 3BQ,
England.

Subscription beginning with 1986-
Editor's adress : REF. PR 4581 A3 H6

Account No. 29501377/500

Fig. 4. Two Purchase Orders Printed on Regular Paper

hot, and staff opening the printer should be warned of the possibility of accidentally burning themselves. The operator's control panel is easy to understand and use; the status display gives information and error codes to keep the operator informed of the printer's operating state at all times. Changing paper is a simple procedure: filling a tray with the appropriate stock and inserting it into the printer, which will wait until the paper tray is refilled and will resume printing, even the next day, unless the power to the printer is turned off. Changing paper trays from label stock to plain bond paper for lists and purchase orders is much easier than changing the three different styles of continuous, tractor-feed paper stock previously used.

CAPITAL AND OPERATING COSTS OF LASER PRINTERS

The toner cartridge on the tested LaserJet is rated for more than 3,000 pages of 2,000 characters per page, and as the toner cartridge empties, one can redistribute the toner by simply removing the cartridge and shaking it. This procedure is sometimes useful in times of high humidity or when the print resolution fades at the extreme edges of the printed product. As the toner cartridge empties, redistributing the toner in this manner will often restore the printing resolution without the need to replace the cartridge itself. The toner costs approximately \$65, or 2¢ per page. One may con-

sider this expensive; however, if you must regularly replace nylon or carbon ribbons on a dot matrix or daisy wheel printer to obtain similar high-resolution printing quality, the laser printer is cost-effective.

The higher cost of continuous, tractor-feed paper versus regular bond paper stock is another significant price factor in favor of laser printing. The significant savings in call number label stock amounted to 91¢ per label. Regular bond paper is approximately 3.5¢ less than tractor-feed paper, depending on the volume purchased.

COMPARISON OF CAPITAL AND OPERATING COSTS FOR A THREE-YEAR PERIOD

Table 1 compares the capital and operating costs of a Hewlett-Packard LaserJet to an Epson LQ 2500 with similar high-resolution print quality. Operating costs are based on the same printing output of 15,000 pages annually, sufficient for producing approximately 10,000 purchase orders; 10,000 call number spine labels; and 10,000 pages of lists, budget reports, etc. All the prices are given in U.S. dollars, based on actual wholesale prices quoted in the literature or paid at the time of this writing, and do not include shipping, handling, or local sales taxes.

The prices quoted for the printers and the toner cartridge were taken from advertisements contained in *PC World*.⁹ I re-

Table 1. Three-Year Cost Comparison of a Laser and an Impact Printer

	Laser		Impact
	Hewlett-Packard LaserJet Series II		Epson LQ 2500
1. Printers Chosen for Comparison			
2. Capital Costs			
a. Wholesale List Prices	\$1,699		\$990
b. Font Cartridge	350		
c. Acoustical Cover			257
Total Capital Costs	\$2,049		\$1,265
3. Annual Operating Costs			
a. Toner	\$325		
b. Ribbons			\$1,250
c. Difference in Paper Supplies Cost			\$700
e. Difference in Call Number Spine Labels Cost			\$91
Total Difference in Annual Operating Costs:	\$1,716		

The total net savings, including capital and operating costs and based on the above-mentioned items, amounts to \$4,364 over a three-year period in favor of the LaserJet printer.

cently purchased a black ribbon for my Epson LQ 2500 and paid \$32.55 in Canada; the cost was converted to \$25 U.S. in calculating table 1. In order to equal the price of a laser toner cartridge, ribbons for the LQ 2500 would need to be reduced to \$6.50. Toner cartridges for the LaserJet are available from vendors at discount prices for \$82.95 and are rated for 3,000 to 5,000 pages, or 2¢ to 3¢ per page. I have used \$65 for 3,000 pages as the average cost for a toner cartridge.

The average life of an LQ 2500 ribbon having comparable print quality to the LaserJet was rated at 300 pages. Beyond 300 pages, either new or reinked ribbons had to be used in order to maintain an acceptable print resolution for the output. Differences in label and paper supplies costs are based on what was actually paid to suppliers. The differences amounted to 3.5¢ more per page and .91¢ per label for the tractor-feed stock that was used in the impact printer tested. Printing speed, in my opinion, precludes the use of a sheet feeder for the impact printer, as the delay (waiting for longer reports to be printed) would become a substantial cost factor. When tested with a sheet feeder, the Epson LQ 2500 took 17 seconds to feed a single sheet of regular bond paper and 60 seconds to print 52 lines of text. The LaserJet, in the same time interval, was able to print 10 pages containing 52 lines of text, ten times the printing speed of the Epson LQ 2500.

Even if the paper supplies and label stock were identical for both printers, the laser would clearly demonstrate cost savings over the impact printer tested. The quality of the resolution and printing speed are major cost factors, influencing to some degree the other supplies and operating costs expe-

rienced, and clearly favor the laser over the impact printer tested.

DISADVANTAGES OF LASER PRINTING

The major disadvantage is the higher initial purchase price. The printer tested was \$1,059 more expensive, including the font cartridge, than a comparable impact printer priced to produce spine labels. If one were to add the cost of providing an acoustical cover for the impact printer, the total difference in capital costs between laser and impact printers would be \$784.

The only other disadvantage that deserves attention is that a laser is limited to printing only one copy at a time, while an impact can print forms with up to eight copies at the same time. This may or may not be a major concern, depending on printing requirements.

CONCLUSION

Laser printing is affordable and suitable for a wide range of library-printed products. There are significant advantages in laser printing that cannot be found in the current generation of impact printers, either dot matrix or daisy wheel. The overall capital costs of purchasing a laser are not significantly higher than those experienced for impact printers if one adds the cost of office noise-abatement equipment. The real savings in laser printing for libraries will be found in the operational expenses, printer supplies, and paper stock.

The ease of use and maintenance, overall performance, printing speed, and superior quality of the finished product gives laser printing an unqualified edge over the competition for many library applications.

REFERENCES AND NOTES

1. *PC Tech Journal* 4, no.13:108-10 (1987).
2. Marjorie McCloy and Ted Nace, "Would-Be LaserJets," *PC World* 5, no.6:218-25 (June 1987).
3. I would like to acknowledge the technical assistance provided by Bill James of our Computer Services Department in writing the macro programs and changing the print symbiont routines, allowing us to use the laser printer with our application program.
4. Figure 1 is reprinted with the permission of Hewlett-Packard and illustrates the Roman-8 symbol set used by their LaserJet printer. Table 1 is reprinted from their technical manual "HP LaserJet Printer: Using the HP 92286 W Bar code 3-of-9/OCR-A Font Cartridge," June 1985, p.12.
5. Hewlett-Packard, "LaserJet and LaserJet +

- Printers Operator's Reference Manual," May 1986, p.D13.
6. When first contacted in December 1986, library suppliers of spine labels did not have any stock suitable for laser printers and considered our request a custom order. After almost a month of searching, a supplier who had been producing bar code-label stock for laser printers—and who was also willing, for a onetime setup charge of \$150, to produce custom spine-label stock in the dimensions mentioned—was located. What was astonishing is that money was actually saved in the process! Including the special setup charge, the cost per label was 91¢ less than what was paid to our former supplier, who recently announced label stock for laser printers at a slight premium over the regular stock for impact printers! Spine-label stock was obtained for our laser printer from Imaging Products, 12696 Rockhaven Rd., Chesterland, OH 44026.
 7. Additional information on using adhesive labels and other special paper stock can be found in Hewlett-Packard, "LaserJet and LaserJet + Printers Operator's Reference Manual," p.D8-D14.
 8. The program, Polaris Labelmaker, is available from Polaris Software, 310 Via Vera Cruz, Ste. 310, San Marcos, CA 92069. It is specifically designed for Hewlett-Packard LaserJet printers, will run on IBM PCs or 100 percent compatible machines, and costs \$124.
 9. *PC World* 5, no.6:352, 388 (June 1987). ■■

Twenty Years Ago in JOLA

Organization of papers. Whenever possible, material should be presented in classical form: an introduction, which should set the scene for the report by referring to pertinent related literature; methods and materials; results; discussion; and conclusion or summary.

An indicative abstract of 100 words or less must accompany the manuscript on a separate sheet of paper. It should include purpose, method, results and conclusions, and any specialized information. It should be intelligible without reference to the article, while assuming that the reader has a general knowledge of the subject. Clarity is essential, but complete sentences are not necessary. There shall be no footnotes.

From the instructions to authors, *JOLA* (Mar. 1968).

Online Maintenance Features of Authority Files: Survey of Vendors and In-House Systems

Agnes M. Grady

At the 1986 American Library Association (ALA) Midwinter Meeting in Chicago, during a meeting of the Ad Hoc Committee on Online Maintenance of the ALA LITA/RTSD CCS Interest Group on Authority Control in the Online Environment, it was suggested that authority control capabilities, particularly online maintenance features of present systems, be examined. It was felt that the best method would be a survey of libraries having authority control and using authority control vendors and of the vendors themselves. A small working group¹ developed and tested a survey instrument, after which the survey was mailed and the responses received and tabulated. This is a report on the results.

SURVEY INSTRUMENT

The individual members of the working group began by drafting questions relating to the concepts that they felt were appropriate for the survey. After consultation with other librarians involved with online systems, a compiled draft of forty-nine questions was organized into sections that dealt with global update features; uniqueness, hierarchical, and reciprocity checks; human intervention; validation; and sources of authority files. Several revisions that combined, changed, and eliminated questions were made before sending the questionnaire to four test sites: three university libraries and one vendor.

Based upon these responses, additional revisions were made. To ensure consistent responses, a glossary of nine terms was appended. Space was provided at the end for adding comments, questions, and clarifications. A cover letter explaining the reason for the survey and asking for cooperation in completing the questionnaire was also added.

The survey instrument was mailed in December 1986 to several institutions and companies whose names were included either in a listing from the April 1985 *Library Resources & Technical Services* article on network and vendor authority control (p.195-205) or a list used by OCLC, or, in the case of some libraries, were known to have online authority control. The initial response was eighteen, two of which were returned blank with the information that the institutions had no online authority control but would be interested in knowing the results of the survey.

During the January 1987 ALA Meeting, a preliminary report, based on these responses, was given. Afterwards, however, a reminder was mailed that allowed an extension of time for completing the questionnaire, since December is not the most auspicious month to conduct a survey. Five more responses were received. After mailing a total of seventy-eight survey instruments, twenty-two were returned with enough of the questions answered to count,

Agnes M. Grady is Head, Cataloging Department, University of Tennessee, Knoxville. A version of this paper was presented at a meeting of the LITA/RTSD Authority Control Interest Group at the ALA Annual Conference in June 1987.

but every question was not answered by all respondents.

SURVEY RESPONSES

First, some general comments about those responding. There is a clear distinction between vendors having a system and libraries developing their own systems: a vendor system is generic, while a library system is customized. There are also differences within these two groups. Some vendors have a system for providing libraries with authority control, and some have a system for enabling libraries to provide authority control. The response from some in the latter group indicated at appropriate places in the survey that local options would prevail. In other words, the system may be able to provide some feature of authority control, but the vendor could not say whether the library having that system would implement that authority-control feature. Nine of the respondents can clearly be categorized as vendors, two of which provide batch authority processing of bibliographic records.

The degrees of implementation varied. One response noted many possible features in the system that "are unused or rarely used by operational forces of this library." The libraries who have or are developing their own systems usually offered long informational comments, pointing up their systems' unique features. The responses indicated the extremes of either a short or long period of experience with online authority maintenance.

The survey instrument began with this statement: "This survey addresses only online maintenance of authority files. The questions do not address maintenance of bibliographic files unless the question specifically states such." Only one question dealt with bibliographic maintenance. Please keep this statement in mind while reading the survey questions, followed by the summarized results and a brief discussion. The definitions that were given in a glossary at the end of the questionnaire have been included at the beginning of each relevant section.

Global Update Features

Global update: capability to use one

command to change all representations of a heading from one form to another form.

Partial field replacement: capability to automatically change an access field to a new form if either the entire character string or a portion of it matches.

1. Does your system have global update capability within the authority file?

18 (85.7 percent) Yes; 3 No

2. Global updates are accomplished through:

12 update online

12 update in batch mode

8 allow for human intervention on each occurrence of a match

7 allow for human intervention as an option

4 not accomplished

A main concern regarding online authority maintenance is how one makes changes. The abilities to update online and in batch mode were both used in several systems. Additionally, some systems had provisions that also allow human intervention, either on each occurrence of a match or as an option. Some comments indicated that procedures have been established for mandating human review prior to a global update that affects many authority records.

3. Does your system have automatic partial field replacement? [E.g., could your system change "United Kingdom" to "Great Britain" where that phrase was a candidate for change in a field that was indexed?]

9 (45 percent) Yes; 10 No

4. Your system has global update feature on fields that are indexed:

13 (62%) all [i.e., 1xx, 4xx, 5xx fields]

4 (19%) only 1xx fields

0 only 1xx and 4xx fields

1 only 1xx and 5xx fields

3 no fields

5. Does your system automatically update bibliographic records when a change occurs in an authority record 1xx field?

19 Yes; 3 No

This is an 86.3 percent rate for the only question in the survey that dealt with bibliographic record maintenance.

Uniqueness Checks

Uniqueness checks: determines matches between two identical character strings in access fields.

Normalization: a computer edit designed to eliminate all but the essential characters of the heading for the purpose of comparison. For example, all diacritics and most punctuation could be removed, all letters converted to uppercase, and all modified letters converted to their unmodified equivalents.

6. Your system performs uniqueness checks between:

- 15 1xx fields in *different* records
- 16 headings (1xx field) and *see* references (4xx fields)
- 13 headings (1xx fields) and *see also* references (5xx fields)
- 2 no fields

Some of the responses included checks for all three choices, and some checked two.

7. Does your system use normalization?
16 (72.7 percent) Yes 6; No
If Yes, do you use a standard table or a locally developed table?

Most responses followed the example in the definition, and many indicated that they used a standard table.

Describe your matching algorithm.

A few respondents listed such differences as providing locally specified diacritics processing, regarding or disregarding death dates, and changing the date to lowercase alphabetic characters to provide chronological arrangement.

Hierarchical Checks

Hierarchical checks: determines if the parent bodies of an entry established subordinatedly are represented in the authority file.

8. Does your system perform hierarchical checks on corporate headings?

For example, when the entry UNITED STATES. CONGRESS. HOUSE. COMMITTEE ON FOREIGN AFFAIRS. is checked hierarchically, the following entries would have to have been previously established:

UNITED STATES. CONGRESS.
HOUSE.

UNITED STATES. CONGRESS.
UNITED STATES.

2 (9 percent) Yes; 20 No

9. Does your system perform hierarchical checks on:

- (a) series 2 (9 percent) Yes; 19 No
- (b) subjects 5 (23.8 percent) Yes; 16 No

10. Does it check name/title headings by checking for presence of an authority record for the name when entering a name/title access point?

9 (42.8 percent) Yes; 12 No

The returned surveys clearly demonstrated that the ability to perform all possible hierarchical checks is missing from most current authority maintenance systems.

Reciprocity Checks

Reciprocity checks: determines if a corresponding record bearing a matching 1xx field is present when a *see also* reference (5xx) is created.

11. Does your system perform reciprocity checks for presence of authority record with 1xx heading which would match each 5xx (*see also* reference) input?

12 (60 percent) Yes; 8 No

If the answer was yes, respondents were asked to explain what happened when either a match or no match was found. Most stated that the result was the generation of a *see also* reference. Remember we are talking about authority records; this does not mean that any 5xx headings were created.

The above features concentrated more or less on machine capability, but authority control is founded on intellectual decision making. We were therefore interested in the human intervention relative to online maintenance of authority files.

Human Intervention

12. Does your system provide notification of changed authority records?

10 (52.6 percent); Yes; 9 No

13. Does your system provide notification of records to be deleted?

4 (22 percent) Yes; 14 No

14. What percentage of authority maintenance work (e.g., global updates, uniqueness checks, hierarchical checks, reciprocity checks) does your system do automatically?

3 (0 percent); 0 (20-59 percent)

3 (60 percent)

7 (80 percent) 3 (100 percent)

7 No response/not applicable/not known

Some respondents qualified the 100 percent category, one by changing it to 90-100 percent and another with the comment "after global updates initiated." At least two of the 80 percent responses were qualified by limiting the features included, such as "not counting hierarchical checks." One vendor providing batch authority work stated that "it varies depending on the bib. file source and general state of the file. Can be as high as 80%."

15. Does your system notify the cataloger of "near matches," e.g., headings with nearly identical dates, common spelling errors, etc.?

9 (40 percent) Yes; 13 No

Validation

Dependent field validation: capability to cross-check between fields, usually between fixed and variable fields. E.g., in a series authority record, byte 16 should be coded in accord with the presence of 64x fields.

Content designators: tags, indicators, subfield codes, control field values.

16. Does your system have dependent field validation?

6 (27 percent) Yes; 14 (64 percent) No

2 (9 percent) No answer

17. Does your system validate content designators in the record against a table of valid codes (e.g., USMARC format or other authorities format?)

15 (68 percent) Yes; 7 No

Appended to one affirmative answer was the comment "to the extent they are present in the master authority record." This response was almost the opposite of that for dependent field validation.

Other Features

18. What does your system do with superseded versions of an authority record?

3 keeps 3 merges 15 deletes

Three respondents checking the "keeps" category added comments: "user option for merging and deleting" and "user notified of change, identified as superseded." One re-

spondent indicated that it was up to the cataloger to decide which action was the most appropriate to take. Another stated that superseded records were handled manually, in order to decide if conversion to a *see* reference is needed.

19. Does your system permit co-existence of authority records from multiple sources, such as Library of Congress (LC), National Library of Medicine (NLM), and National Agricultural (NAL), in the same database?

16 (76 percent) Yes; 5 No

If yes, in the same or separate authority file?

4 same; 12 separate authority files

One system provides the retention of records in either the same or separate files as a user option.

Authority Files

Authority file: a file containing authorized headings (name, subject, series titles), related references (*see* and *see also* references), and information upon which the decision for the authorized heading was based.

20. Does your system accommodate USMARC authority format?

16 (76 percent) Yes; 5 No

1 No answer

The final questions (numbers 21 and 22) dealt with the various authority files to which the institutions had access and how often these files were updated. These questions were divided between types of files (names, subjects, series, uniform titles) and whether the records for these files were provided through a network, by a vendor, and/or locally (which could also include extracting headings from bibliographic records). Other sources of authority records in addition to Library of Congress and local input included NLM, NAL, and Catholic, and one vendor said "any." At the time the surveys were completed, only three institutions and vendors were taking advantage of LC Subject Authorities tape loading. One respondent reported that only authority records for headings requiring references or notes for the catalog or for recording local series decisions were retained.

The answers to how often updating of

authority files occurred ranged from "instantaneously" to "rarely." More precisely, one response indicated monthly updating, four weekly, and four daily. Other designations included "every six months" and "irregularly," while automation nomenclature was expressed by the phrases "online" and "real time." One comment appended to a weekly check was that new headings were added weekly and references added only irregularly. Another respondent updated names irregularly, subjects continuously, and series and uniform titles rarely.

MISCELLANEOUS SURVEY INFORMATION

Comments from those responding indicated not only the current status of authority control systems but also future enhancements. Several respondents who replied no to questions concerning such features as notification of changed records also stated that these were under development and appended the expected date for acquiring the capability. One respondent answered the entire questionnaire on the basis of an online authority system in development but not yet implemented; this response was not included in the numbers reported throughout this report.

There were a few cases where the response was a clear indication that a term was not being used in the way it was defined in the glossary. A few answers showed confusion between *authority* and *bibliographic* maintenance. In these instances, the answers were not included in the tally.

CONCLUSION

In summary, this was a very small survey of a subject that is currently getting the attention it deserves because of such factors as increases in library implementation of online catalogs and vendors' attempts to provide the software for implementing online authority control. If this survey were repeated in one year, could we expect questions on any different features? Would the response be more of the same, except in greater numbers? From the discussion on authority control systems heard during the 1987 ALA Annual Conference in San Francisco, it does not appear that other features will soon be considered. Several libraries, however, will be implementing authority control systems within the next year. With more libraries gaining practical experience, today's theoretical base will need expansion tomorrow.

REFERENCES AND NOTES

1. The members of the working group were Judith G. Fenly, Library of Congress (chair); Agnes M. Grady, University of Tennessee, Knoxville; and Andrea Stamm, Northwestern

University. Editorial assistance was provided by George Gibbs, University of California-Los Angeles, and Barbara Tillet, University of California-San Diego. ■■

Bibliographic Data Management with dBASE: A Study of Secondary Key Retrieval on Multivalued Data Items

Richard Pollard

The microcomputer database management system dBASE is being used in a wide range of library-related activities involving bibliographic data. In this article we examine the ability of dBASE to perform secondary key retrieval on multivalued data items. If the data set is large and the number of matches small, dBASE can provide an adequate level of performance. If, however, queries are such that a large number of matches occur, retrieval performance may be unsatisfactory. In these circumstances a hierarchical database management system or a file management system based on inverted files might be more appropriate.

The widespread acceptance of microcomputers by libraries has led to a growing interest in the local creation and management of bibliographic databases. Reports appearing in the library and information science literature suggest that microcomputer database management systems (DBMS) are being widely used for this purpose.

Of the large number of DBMS products available for microcomputers, Ashton Tate's dBASE appears to have found particular favor. Beiser, in a recent book, proposes the use of dBASE III Plus in a wide range of library-related activities.¹ Suggested applications specifically involving bibliographic data include management of bibliographies, newspaper indexing, production of a serials union list, maintenance of acquisition files, serials control, and even the online catalog. In addition, the journal literature contains reports of dBASE being

used in a variety of library applications including management of a subject headings list,² a photocopy invoicing system,³ a subject index to statistical resources,⁴ a reserve collection,⁵ circulation files,⁶ and an interlibrary loan journal request file.⁷

While a general-purpose microcomputer DBMS such as dBASE can certainly be made to function in a wide range of applications involving bibliographic data, it may not always be the most appropriate tool to use. This author contends that dBASE has significant shortcomings with respect to the management of bibliographic data. First, dBASE records and fields have a predominantly fixed-length orientation, whereas bibliographic data is mostly of variable length. Second, and of more consequence, dBASE index structures are more suited to primary key searches than to secondary key searches. In this article we examine dBASE's ability to perform simple

ISBN: 0913312789
 Title: GEOREF Thesaurus and Guide to Indexing
 Author: Palmer, Crystal S., Ed.
 Edition: 4th., 1986
 Subjects: Geology, Geochemistry; Geophysics; Hydrology

Fig. 1. *Bibliographic Record*

secondary key retrieval on multivalued data items.

MULTIVALUED DATA ITEMS

Data items that occur in bibliographic records often contain more than one distinct value. For example, in the abbreviated bibliographic record shown in figure 1 the *Subjects* data item has four individual values: *Geology*, *Geochemistry*, *Geophysics*, and *Hydrology*. In this article we use the term *multivalued* to refer to such data items.

Multivalued data items can be accommodated by general-purpose database management systems such as dBASE, in which a file consists of one or more records and each record contains one or more fields. Note, however, that as dBASE does not support subfields (a structural device for accommodating multivalued data items), we must explicitly arrange such data items so that individual values may be used as keys in an indexed search. In this article we refer to such an explicit arrangement as a *mapping*.

Single-File Mapping

The format of the bibliographic record shown in figure 1 suggests a "natural" mapping in which the multivalued data item is placed in a single field and individual values are separated with punctuation or some

special character. Figure 2 shows the example bibliographic record under this mapping. The multivalued data item *Subjects* appears in the field SU with individual values separated by semicolons. We will refer to this arrangement as a single-file mapping.

Those familiar with database theory will protest that the single-file mapping should not be considered, as it is an example of poor practice. While we are fully aware that this mapping does indeed represent poor practice, we nonetheless consider it because it is such common practice in the microcomputer environment.

A significant problem with the single-file mapping becomes apparent as soon as we attempt to use individual values of a multivalued data item as keys in a dBASE indexed search. If we were to use the key value *Geology* in an indexed search of the file containing the record shown in figure 2, the search would identify the record as a match. A similar search with the value *Hydrology* as key would, however, fail to identify the record as a match.

The search fails to identify the record as a match, because "substring or partial key searches work only if the search expression watches [*sic*] the index key expression, starting with the leftmost character."⁸ Thus, an indexed search of the single-file mapping using an individual value of a

BN	0913312789
TI	GEOREF Thesaurus and Guide to Indexing
AU	Palmer, Crystal S., Ed.
ED	4th., 1986
SU	Geology; Geochemistry; Geophysics; Hydrology

Fig. 2. *Bibliographic Record under the Single-file Mapping*

multivalued data item as key will work only if that value happens to be the first in the field. If we wish to use values other than the first as search keys we have to resort to a sequential search of the file.

Fortunately, there are alternatives to the single-file mapping. One of these takes advantage of the relational capability of dBASE.

Multifile Mapping

A characteristic of dBASE that sets it apart from "flat-file managers" such as PFS Professional File and Rapidfile is its ability to have several files open at the same time. The record pointer for each open data file usually moves independently of the others. It is possible, however, to link two files so that whenever we change the position of the record pointer in one file, the pointer in the other file moves automatically to a corresponding record. The link between two dBASE files can be based on record numbers or on a field that is common to both files. This ability to link files on the basis of a common field gives dBASE its relational capability.

We can use this relational capability to form a multifile alternative to the single-

file mapping discussed above. The multifile mapping is formed by assigning the multivalued data item to the primary or "controlling" file and the remaining data items to the related or "linked" file. Each record in the controlling file contains a single value from the multivalued data item and a common field value that uniquely identifies the corresponding record in the linked file. In the parlance of the relational data model we have removed a repeating group from the unnormalized relation represented by the record shown in figure 1. Further details of the relational model as it applies to bibliographic data can be found in articles by Koenig,⁹ Brooks,¹⁰ Caswell,¹¹ and Crawford.¹²

Figure 3 shows the example bibliographic record under the multifile mapping. The SU field of the controlling file contains individual values from the multivalued *Subjects* data item, and the linked file contains data items TI, AU, ED, and BN, where BN acts as the link between the two files. An indexed search of the controlling file using the value *Geology* as key will cause the pointer in the linked file to be positioned to the record with a BN value of 0913312789, identifying this record as a

Controlling file	
SU	Geology
BN	0913312789
SU	Geochemistry
BN	0913312789
SU	Geophysics
BN	0913312789
SU	Hydrology
BN	0913312789

Linked file	
BN	0913312789
TI	GEOREF Thesaurus and Guide to Indexing
AU	Palmer, Crystal S., Ed.
ED	4th., 1986

Fig. 3. *Bibliographic Record under the Multifile Mapping*

match. A similar search with the value *Hydrology* as key would also identify the record as a match.

In summary, it is possible to accommodate multivalued data items in dBASE such that individual values can be used as keys in an indexed search. This can not be accomplished simply by placing all values of the multivalued data item in a single field. A multifile mapping consisting of two files linked on a common field may, however, be used.

SECONDARY KEY RETRIEVAL

An individual value of a multivalued field often occurs in more than one record in a particular file. For example, a value from the subject descriptor field of one bibliographic record may also appear in the subject descriptor field of many other records. In this article we use the term *secondary key* to refer to an attribute value that occurs in more than one record. For the sake of simplicity we restrict our discussion to simple queries, that is, those which involve a specific value of a specific attribute.¹³ A discussion of the use of dBASE for Boolean queries can be found elsewhere.¹⁴

From a conceptual point of view, each entry in a dBASE index can be envisioned as containing an attribute value and the address of a single record with that value. A simple secondary key search in dBASE can not, therefore, be satisfied simply by examining a single index entry: every matching index entry must be examined.

Fortunately, all the index entries for a particular attribute value are grouped together in one logical area of the index. Once the first matching entry for a particular secondary key has been identified, the remaining matches are located by stepping sequentially through the index until a nonmatching entry is found. Figure 4, for example, shows a fragment of a dBASE index in which all the entries for the attribute value *Engineering* are grouped together. Matching records would be identified by stepping through these entries until the value *Environment* is encountered.

From the foregoing it is clear that secondary key retrieval is possible in dBASE. Such queries can, however, only be satis-

Attribute value	Record number
Engineering	7
Engineering	26
Engineering	79
Engineering	85
Environment	26
Environment	74
Environment	98

Fig. 4. Logical View of dBASE Index Entries

fied by examining every matching entry in the index.

RELATIVE TIME COST OF SECONDARY KEY RETRIEVAL IN DBASE

In this section we estimate the time cost of dBASE secondary key retrieval relative to the number of matching records. We start by examining the dBASE implementation of the multifile and single-file mapping.

Multifile Mapping

The multifile mapping consists of two files linked on the basis of a common field. If two dBASE files are to be linked, one of them has to be designated as the controlling file, the other as the linked file. In addition, the linked file has to be indexed on the common field. Whenever the record pointer in the controlling file is moved, the new common field value is used as the key in an indexed search of the linked file. As a result, the pointer in the linked file moves to the record that is associated with the current record in the controlling file.

The dBASE implementation of the multifile mapping consists of four files. The controlling file contains individual values from the multivalued data item and an associated index file provides indexed access to these values. The linked file contains the remaining data items from the bibliographic record. An index to the linked file is required to implement the linkage between the data files.

A secondary key search of the multifile mapping starts with an indexed search of

the controlling file. If a match is found, an indexed search takes place on the linked file. The record pointer in the controlling file is then moved to the next record in indexed order. If this record matches the key, another indexed search takes place on the linked file. This process continues until the record pointer in the controlling file is moved to a record that does not match the search key. In summary, the first match requires two indexed searches: one of the controlling file, the other of the linked file. Each additional match requires one movement of the record pointer and one indexed search of the linked file.

The time cost of an indexed search depends on the structure of the index and the number of records in the indexed file. In dBASE, an index file is organized as a B^+ tree data structure.¹⁵ A B^+ tree consists of an index set organized as a B tree and a sequence set organized as a linked list. In a B tree of order d for a file of n records, the time cost of a primary key search is proportional to $\log_d n$ in the worst case.¹⁶ As an example of how reasonable this cost can be, Comer indicates that a B tree of order fifty that indexes a file of one million records can be searched with only four disk accesses in the worst case.¹⁷ A further desirable property of the B^+ tree is that it requires at most one disk access to locate the next key in indexed order.¹⁸

To summarize, the time cost of a dBASE secondary key search on the multifile mapping is proportional to the number of matching records in the data set. The time cost of locating each match is proportional to the logarithm of the number of records in the data set.

Single-file Mapping

Secondary key retrieval on the single-file mapping uses a sequential search. The sequential search starts at the beginning of the file and examines every record in turn until a match is found, or the end of the file is reached. If a match is found, the search examines each remaining record until another match is found. This process continues until the end of the file is reached. Every record in the file is examined, so that no matter how many records match the search key, the time cost of a secondary key

search on this mapping is not dependent on the number of matches that occur but is proportional to the number of records in the data set.

ABSOLUTE TIME COST OF SECONDARY KEY RETRIEVAL IN DBASE

Our informal analysis of the relative time cost of dBASE secondary key retrieval suggests that the multifile mapping is superior to the single-file mapping when the number of matching records is small. It also suggests that, for a data set of fixed size, the difference in performance between the two mappings becomes smaller as the number of matches becomes larger. The analysis does not, however, provide any indication of the absolute time cost of dBASE secondary key retrieval.

The library and information science literature contains many descriptive reports on the use of a single-file mapping to represent multivalued data items in dBASE. At least two of these studies also give an indication of the absolute time cost of secondary key retrieval on this mapping. Armstrong, reporting on the use of dBASE II to manage a small collection of offprints, photocopies, and journals, states that "a complete [sequential] search of 300 records takes well under 30 seconds."¹⁹ Sullivan used dBASE II to retrieve citations from a reprint file and found that a sequential search on a 100-character keyword field in a file of 500 records takes about 30 seconds.²⁰

The literature also contains some reports on the use of the multifile mapping to represent multivalued data items in dBASE. Bordwell, for example, uses a multifile mapping to represent a trade literature collection in which a single company may have several products.²¹ O'Brien uses a multifile mapping to provide indexed access to a field in the SIR database that may contain up to 7 values.²² Vaughan, using a data set of 10,000 records represented in dBASE III by a multifile mapping, reports that a simple query yielding 49 matches takes 15 seconds.²³ By way of comparison she notes that the same query takes 5 minutes, 30 seconds if a sequential search is used.

We were unable to locate any studies that compare the absolute time cost of secondary key retrieval for the multifile mapping and the single-file mapping with increasing numbers of matching records. In the remainder of this article we describe an experimental determination of these costs.

DETERMINATION OF THE TIME COST OF SECONDARY KEY RETRIEVAL IN DBASE

To determine the time cost of secondary key retrieval we represented a set of bibliographic data in dBASE using a single-file and a multifile mapping. For each mapping we measured the time required for a series of queries yielding increasing numbers of matching records.

We chose as data set a collection of bibliographic citations to journal articles. Each record in this collection contained a multivalued subject descriptor field on which the secondary key searches were to be performed. We limited the size of the data set to 1,000 records, as studies reported in the literature suggest that dBASE takes approximately 60 seconds to sequentially search a file of this size. We decided, somewhat arbitrarily, to measure the time required to perform secondary key searches yielding 0 to 300 matches in increments of 50 matches. As the data set did not support searches yielding exactly the required number of matches, we added subject descriptors to make such searches possible.

Single-file Mapping

The structure of the dBASE file used to represent the single-file mapping is shown in figure 5. The AN field contains an acces-

```
Structure for database: s1map.dbf
Number of data records: 1000
Date of last update   : 02/20/87
Field  Field Name  Type      Width  Dec
  1  AN             Character  7
  2  TI             Character 160
  3  AU             Character  80
  4  JW             Character  80
  5  PY             Character  15
  6  DE             Character 254
** Total **                597
```

Fig. 5. dBASE File for the Single-file Mapping

sion number that uniquely identifies each record. Fields TI, AU, JN, and PY contain respectively, the title, author, journal name, and publication date. Field DE contains the subject descriptors.

Figure 6 shows the main loop of the dBASE program used to perform a secondary key search on the DE field of this file. The command *LOCATE FOR key\$de* performs a sequential search for a record for which the logical expression *key\$de* evaluates to TRUE. If no matching record is found, *EOF()* evaluates to TRUE, and the search stops. If a matching record is found, it is processed. The *CONTINUE* command searches forward from the current matching record for the next match. If another matching record is found, it is processed. When the end of the file is reached, *EOF()* evaluates to TRUE, and the search stops. The logical expression *key\$de* evaluates to TRUE if the character string in the variable *key* is contained in the DE field of the current record.

Multifile Mapping

The data files used to represent the multifile mapping in dBASE are shown in figure 7. Each record in the file *citation.dbf* (CITATION) contains the fields AN, TI, AU, JN, and PY as described for the single-file mapping. Note, however, that this file does not have a DE field. Each record in the file *subjects.dbf* (SUBJECTS) has just two fields: field SU contains a single value from the subject descriptor data item of the original bibliographic record, and field AN contains the record's accession number.

The multifile mapping also has two dBASE index files associated with it. The file *su.ndx* provides indexed access to the subject descriptors contained in the SU field of SUBJECTS. The file *an.ndx*, the index to the AN field of CITATION, is required by dBASE to link files SUBJECTS and CITATION. The mapping is completed by linking the data files on the common field AN so that SUBJECTS is the controlling file and CITATION is the linked file. Figure 8 shows the dBASE commands used to link the two files. A summary of the multifile mapping, obtained from the dBASE *DISPLAY STATUS* command, is shown in figure 9.

```

LOCATE FOR key$de
DO WHILE .NOT. EOF()
    * Statements to process a matching record
    CONTINUE
ENDDO

```

Fig. 6. dBASE Search Program for Single-file Mapping

```

Structure for database: citation.dbf
Number of data records: 1000
Date of last update : 02/20/87

```

Field	Field Name	Type	Width	Dec
1	AN	Character	7	
2	TI	Character	160	
3	AU	Character	80	
4	JN	Character	80	
5	PY	Character	15	
** Total **			343	

```

Structure for database: subjects.dbf
Number of data records: 6972
Date of last update : 02/20/87

```

Field	Field Name	Type	Width	Dec
1	AN	Character	6	
2	SU	Character	56	
** Total **			63	

Fig. 7. dBASE Data Files for Multifile Mapping

```

SELECT 1
USE citation INDEX an
SELECT 2
USE subjects INDEX su
SET RELATION TO an INTO citation

```

Fig. 8. dBASE Commands to Link CITATION and SUBJECTS on Field AN

```
Select area: 1, Database in Use: citation.dbf Alias: CITATION
Master index file: an.ndx Key: an
```

Currently Selected Database:

```
Select area: 2, Database in Use: subjects.dbf Alias: SUBJECTS
Master index file: su.ndx Key: su
Related into: CITATION
Relation: an
```

Fig. 9. Summary of the Multifile Mapping

```
SEEK key
DO WHILE su = key .AND. .NOT. EOF()
    * Statements to process a matching record
    SKIP
ENDDO
```

Fig. 10. dBASE Search Program for Multifile Mapping

The main loop of the dBASE program used to perform a secondary key search on the multifile mapping is shown in figure 10. The command *SEEK key* performs an indexed search on the SU field of SUBJECTS for the value contained in the variable *key*. If a matching record is found, it is processed. If no matching record is found, EOF() evaluates to TRUE, and the search stops. The SKIP command moves the pointer to the next record in indexed order. If the SU field of this record contains the same value as the variable *key*, the record is processed. The search stops when SU no longer matches the value in *key*.

Method

The tests were performed using dBASE III Plus version 1.1 running under PC-DOS 3.2 on a PC/XT compatible computer. A suite of dBASE programs was used to import the data set, add subject descriptors, and build the files used to represent the mappings. We examined the disk layout of the resulting files and eliminated fragmentation, if present.

For each mapping we measured the time required to perform a series of queries with a predetermined number of matches. Although DOS reports time in hours, min-

utes, seconds, and hundredths of seconds, we only recorded search times to one-tenth of a second. The PC's system clock is, in fact, only changed about 18 times a second so the hundredths of a second figure is not accurate.²⁴ The results of the tests are shown in figure 11, in graphic form, and in figure 12.

Discussion of Results

The informal analysis presented earlier suggests that the time cost of secondary key retrieval on the single-file mapping is independent of the number of matching records. The results of our test show that, in practice, the time cost does increase somewhat as the number of matches becomes larger.

Examination of the search program in figure 6 shows that in an unsuccessful search, execution of the LOCATE FOR command sets the end-of-file condition to TRUE, and the DO WHILE loop is not invoked. In a successful search, however, the DO WHILE loop is invoked once for each matching record. The additional time needed to interpret and execute the loop, and the commands contained within it, accounts for the increase in search time with increasing numbers of matches. The in-

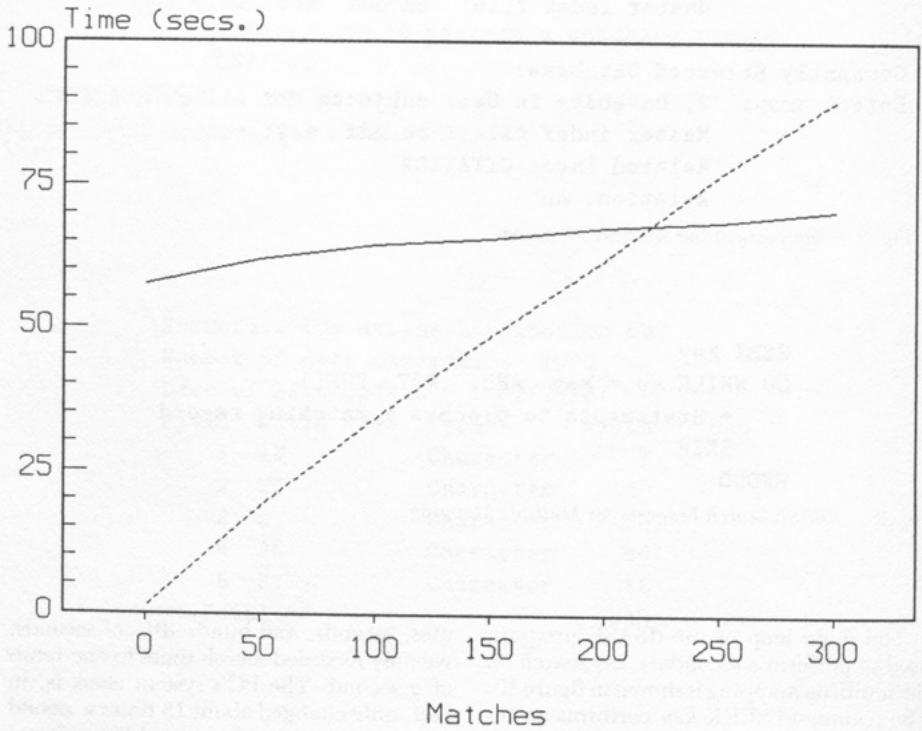


Fig. 11. *Retrieval Performance*

Matches	Single-file mapping (secs.)	Multi-file mapping (secs.)
0	57.4	0.7
50	61.9	18.6
100	64.3	34.1
150	65.6	48.4
200	67.6	61.9
250	68.5	76.7
300	70.5	89.6

Fig. 12. *Time Cost for dBASE Secondary Key Retrieval*

crease is moderate, amounting to about 2 seconds for every additional 50 matches. The minimum time cost for a secondary key search on a 1,000-record data set represented as a single-file mapping was about 57 seconds.

By contrast, the minimum time cost of a secondary key search on the same data set represented by a multifile mapping was less than 1 second. However, as predicted, the time cost increases in proportion to the number of matching records. The increase is substantial, amounting to about 14 seconds for each additional 50 records. The size of this increase can be attributed to the fact that for every matching record in the controlling file, an indexed search of the linked file is performed, and the DO WHILE loop in the search program is interpreted and executed. Nonetheless, the performance of the multifile mapping is superior to that of the single-file mapping for searches yielding up to about 225 matches. Beyond this point the multifile mapping is inferior to its single-file counterpart.

These performance differences notwithstanding, the secondary key retrieval times obtained in our trial are hardly commendable. Fortunately, there are a number of ways of improving these results. We could, for example, run the dBASE programs on a faster computer than a PC/XT compatible. As dBASE is an interpreted language, a faster processor would increase the speed of execution of the search programs. A faster hard disk would also improve performance, as the search programs make frequent accesses to disk storage.

The execution speed of the search programs could also be improved by using more efficient dBASE coding techniques. The DO WHILE loop in the search program for the multifile mapping could be replaced with a dBASE command such as LIST that takes a WHILE condition clause. Dunlop suggests that the execution time for such a construction is approximately half that of an equivalent DO WHILE loop.²⁵ The LIST WHILE command cannot, however, be used in the search program for the single-file mapping, as it requires matching records to be grouped together. Finally, the results could also be improved by using compiled ver-

sions of the dBASE search programs.

CONCLUSION

If one is committed to using dBASE for secondary key retrieval, one must decide whether to represent the data set as a single-file or a multifile mapping. Factors influencing this decision include retrieval time, file creation and maintenance costs, and storage costs.

If the number of records in the data set is small and queries are such that the number of matching records is large, the difference between retrieval times for the two approaches may not be significant. Conversely, if the data set contains a large number of records and queries are such that the number of matches is small, then the multifile mapping will yield significantly faster retrieval times. Small numbers of matches are, however, more likely to occur with Boolean queries than with the simple queries considered in this article.

If the time required to create and maintain the data set representations is of paramount importance, the single-file mapping is preferable to the multifile mapping. Similarly, if storage space is at a premium, the single-file mapping would be the better choice.

The performance attainable from dBASE is ultimately limited by the fact that the time cost of a secondary key search of the single-file mapping is proportional to the number of records in the data set, while the time cost of a corresponding search on the multifile mapping is proportional to the number of matching records. If a satisfactory level of performance can not be obtained from dBASE, an alternative approach is indicated.

Should one wish to continue using a general purpose DBMS for bibliographic retrieval, a program which supports the hierarchical data model (e.g., PC/FOCUS) might be an appropriate alternative to dBASE. If one is not committed to using a DBMS program, a file management program based on the inverted file structure might be an appropriate choice. Programs of this ilk (e.g., INMAGIC) are well suited to secondary key retrieval and also support variable-length fields and records.

ACKNOWLEDGMENT

The research reported in this article was

supported by a University of Kentucky Summer Research Fellowship.

REFERENCES

1. Karl Beiser, *Essential Guide to dBASE III + in Libraries* (Westport, Conn.: Meckler, 1987).
2. Natalie O. Sturr, "Writing a Subject Headings List Program Using dBASE III," *Library Software Review* 5:275-79 (Sept.-Oct. 1986).
3. Janet Evans, "Using dBASE III to Create a Photocopy Invoicing System," *Library Software Review* 5:263-74 (Sept.-Oct. 1986).
4. Cindy Hepfer and Amy R. Loucks-DiMatteo, "Using dBASE III to Prepare a Subject Index to Statistical Resources," *Library Software Review* 5:284-87 (Sept.-Oct. 1986).
5. Philomena Miller and Janette Cochrane, "Administration of a Reserve Collection at Paisley College Using dBASE II," *Program* 19:262-70 (July 1985).
6. Charlotte A. Wagner, "dBASE III for Library Circulation," *Library Software Review* 6:184-89 (July-Aug. 1987).
7. Susan Eichelberger, "Using dBASE III for Interlibrary Loan Journal Request File," *Library Software Review* 6:178-79 (July-Aug. 1987).
8. *Using dBASE III Plus* (Culver City, Calif.: Ashton-Tate, 1985), p.5-149.
9. M. E. D. Koenig, "Data Relationships: Bibliographic Information Retrieval Systems and Database Management Systems," *Information Technology and Libraries* 4:247-72 (Sept. 1985).
10. Terrence A. Brooks, "Relational Database Design in Information Science Education," *Journal of Education for Library and Information Science* 26:3-15 (Summer 1985).
11. Jerry V. Caswell, "Normalization: A Method for Structured File Design," *Information Technology and Libraries* 3:293-96 (Sept. 1984).
12. Robert G. Crawford, "The Relational Model in Information Retrieval," *Journal of the American Society for Information Science* 32:51-64 (Jan. 1981).
13. Donald E. Knuth, *The Art of Computer Programming, Vol. 3: Sorting and Searching* (Reading, Mass.: Addison-Wesley, 1973), p.550.
14. Sheila Vaughan, "Multi-value Fields and dBase III," *Program* 21:183-88 (Apr. 1987).
15. Luis Castro, Jay Hanson, and Tom Rettig, *Advanced Programmer's Guide* (Culver City, Calif.: Ashton-Tate, 1985), p.82.
16. Douglas Comer, "The Ubiquitous B-Tree," *Computing Surveys* 11:121-37 (June 1979).
17. *Ibid.*, p.127.
18. *Ibid.*, p.130.
19. C. J. Armstrong, "The Use of a Commercial Microcomputer Database Management System as the Basis for Bibliographic Information Retrieval," *Journal of Information Science* 8:197-201 (June 1984).
20. Jeanette Sullivan, "Using dBASE II for Bibliographic Files," *Online* 9:46-51 (Jan. 1985).
21. Stephen Bordwell, "dBASE II—Library Use of a Microcomputer Database Management System," *Program* 18:157-65 (Apr. 1984).
22. Sean O'Brien, "Retrieval with dBASE II," *Program* 19:67-71 (Jan. 1985).
23. Vaughan, "Multi-value Fields and dBASE III," *op. cit.*
24. Peter Norton, *Inside the IBM PC* (Bowie, Md.: Robert J. Brady, 1983), p.72.
25. Neil Dunlop, *dBASE III Plus: Things the Manual Didn't Tell You* (Cleveland: Weber Systems, 1986), p.97. ■■

Conventional and Knowledge-Based Information Retrieval with Prolog

William Leigh and Noemi Paz

Prolog is a programming language especially suited to information retrieval applications. Boolean and deductive retrieval are easily implemented; the exciting opportunity, however, is knowledge-based retrieval, which implies the translation into machine-processable logic of the actual knowledge contained in a document. Queries could then be directed against the substance of the documents rather than against keywords.

Prolog is a powerful tool for implementing conventional information retrieval systems, which are based on index and citation information and support Boolean querying. Such a system, if developed in Prolog, would support deductive querying as well. In this section, we illustrate Prolog programming with these applications.

Simple facts are expressed in Prolog as predicates with arguments. For example, the relationship of authorship might be expressed

```
author_of (Kowalski, Logic for Problem Solving).
```

A predicate such as "author_of" has an "arity;" that is, it may be applied to a fixed number of arguments. "author_of" is used here with an arity of two. Alternatively, "author_of" might have been used with a third argument to record the year of publication, such as:

```
author_of (Kowalski, Logic for Problem Solving, 1979).
```

Prolog derives its power from the facility to express rules. A rule looks like the following:

```
author_of (Kowalski, Logic for Problem Solving, 1979) :- wrote (Kowalski, Logic for Problem Solving), published (Logic for Problem Solving, 1979).
```

This rule can be read as "Kowalski is the author of *Logic for Problem Solving* in 1979 if Kowalski wrote *Logic for Problem Solving*, and *Logic for Problem Solving* was published in 1979." The symbol :- is used in Prolog to signify *if*. The predicates listed to the right of the *if* are considered to be in conjunction; that is, the comma between those predicates is read as *and*.

Variables may be used in rules. In the syntax presented here, variables begin with the _ character. For example:

```
author_of (Kowalski, _X, 1979) :- wrote (Kowalski, _X), published (_X, 1979).
```

Or, a more general rule of this structure would be

```
author_of (_Author, _Title, _Year) :- wrote (_Author, _Title), published (_Title, _Year).
```

Facts and rules, collectively called clauses, may be combined into a Prolog program. Such a program constitutes a knowledge base. A program might be

wrote (Kowalski, *Logic for Problem Solving*).

wrote (Clocksin, *Programming in Prolog*).

wrote (Mellish, *Programming in Prolog*).

wrote (Clark, *Prolog: A Language for Expert Systems*).

wrote (McCabe, *Prolog: A Language for Expert Systems*).

wrote (Clark, *Logic Programming*).

wrote (Tarnlund, *Logic Programming*).

published (*Logic for Problem Solving*, 1979).

published (*Programming in Prolog*, 1981).

published (*Prolog: A Language for Expert Systems*, 1980).

published (*Logic Programming*, 1983).

author of (*_Author*, *_Title*, *_Year*):-
wrote (*_Author*, *_Title*),
published (*_Title*, *_Year*).

co_authors_of (*_Author1*, *_Author2*,
_Title):-
wrote (*_Author1*, *_Title*),
wrote (*_Author2*, *_Title*),
not equal (*_Author1*, *_Author2*).

This program may be queried. A query is specified much like a clause except that it is punctuated with a question mark (?). The entry of a query against a program will cause the Prolog interpreter to attempt to find whether the query is true or, in the case of a query with variables, to find those values for the variables that can make the query clause true.

A query that might be entered with the above program is

published (*Programming in Prolog*, 1981)?

Prolog would respond "true." But for the query

published (*Programming in Prolog*, 1976)?

Prolog would respond "false." To the query expressed as:

published (*Programming in Prolog*,
_Year)?

Prolog would respond

_Year = 1981.

Prolog is a "relational" language, which means that it will search for all variable values that can make the query true. For example, a reply to the query:

wrote (*_Author*, *Programming in Prolog*)?

would be

Author = Clocksin.

Author = Mellish.

To the query:

co_authors_of (Clark, *_Author*, *_Title*)?

Prolog would respond

_Author = McCabe, *_Title* =
*Prolog: A Language
for Expert Systems*.

_Author = Tarnlund, *_Title*
= *Logic Program-
ming*.

A query that involves a rule such as that for "co_authors_of" requires that Prolog do more than a simple table search to find the response. The answer must be "deduced" by following a chain of rules and facts. This facility, known as "deductive retrieval," is powerful because relationships can be built up and searched as the combination of other relationships without being explicitly listed in the database.

Prolog implementations generally make available predefined, imperative predicates for input and output. An example might be "print" to cause output. The "co_authors_of" clause above might be reexpressed as:

co_authors_of (*_Author1*, *_Author2*,
_Title):-
wrote (*_Author1*, *_Title*),
print ("looking at", *_Author1*),
wrote (*_Author2*, *_Title*),
print ("candidate", *_Author1*),
not equal (*_Author1*, *_Author2*).

In this case execution of the query:

co_authors_of (_Author, Mellish,
_Title)?

would result in:

looking at Kowalski
looking at Clocksin
candidate Clocksin

_Author = Clocksin, _Title = *Programming in Prolog*.

looking at Mellish
candidate Mellish
looking at Clark
looking at McCabe
looking at Clark
looking at Tarnlund

This query track reveals plainly the backtracking nature of Prolog search.

Support for keyword index term storage and retrieval can be developed similarly. An index might be stored as a relation:

index_term (*Programming in Prolog*,
programming).
index_term (*Programming in Prolog*,
prolog).
index_term (*Logic for Problem Solving*,
logic).
index_term (*Logic for Problem Solving*,
prolog).
index_term (*Logic for Problem Solving*,
logic programming).

A Boolean query to find titles indexed for "programming" AND for "prolog" would be formed as:

index_term (_Title, programming),
index_term (_Title, prolog)?

A request to retrieve titles that are indexed for "logic" OR "prolog" would be expressed as two Prolog queries:

index_term (_Title, logic)?
index_term (_Title, prolog)?

A better way to express the second query here, so that citations duplicating the results of the first query would not be retrieved, is

index_term (_Title, prolog), not index_term (_Title, logic)?

The "index_term" relation may be combined with the relations concerning the citation information which were defined above. This would accommodate such queries as:

Index_term (_Title, prolog),
published (_Title, _Year), greater
(_Year, 1980)?

This query would return titles published after 1980.

This is a brief introduction to Prolog and how it might be used for conventional information retrieval. Additionally, the repertoire of the Prolog programming language described here is sufficient to carry out our knowledge-based approach to information retrieval.

KNOWLEDGE-BASED INFORMATION RETRIEVAL

Michie used the term *knowledge refinement* for the process of converting knowledge into logic for computer processing, proposing logic as an appropriate alternative form for instruction manuals. Knowledge refinement implies the conversion of the content of the material into logic. This logic then becomes "of" the source, rather than "about" the source.

To apply knowledge refinement to information retrieval, it is necessary to convert the knowledge contained in the archived documents into logic. For example, below are some rules drawn from Kowalski's preface:

technology (theorem proving) :-
objective (problem solving).
technology (logic programming) :-
objective (problem solving).
technology (logic programming) :-
objective (expressing programs).
technology (logic programming) :-
objective (expressing specifications).
procedure (resolution):- technology
(theorem proving).

These rules are an expression in logic of the knowledge that: (a) theorem proving is a technology that may be applied to problem solving; (b) logic programming is a technology that may be applied to problem solving, program expression, or specification expression; (c) resolution is a procedure that is involved in theorem-proving technology.

The logic resulting from documents can be combined into one knowledge base.

Rules derived from Clocksin and Mellish are added to the rules derived from Kowalski:

```

technology (theorem proving) :-
  objective (problem solving).
technology (logic programming) :-
  objective (problem solving).
technology (logic programming) :-
  objective (expressing programs).
technology (logic programming) :-
  objective (expressing specifications).
technology (Prolog) :- technology
  (logic programming).
procedure (resolution):- technology
  (Prolog).
procedure (resolution):- technology
  (theorem proving).

```

Queries can be directed against the knowledge in this knowledge base. Deductions can be carried out and actual answers to the query produced. For example, the query:

```
objective (problem solving), technol-
ogical (_What) ?
```

would result in the reply:

```

_What = theorem proving.
_What = logic programming.
_What = Prolog.

```

This query posed the problem of listing the technology or technologies that may be applied to problem solving. The reply lists theorem proving, logic programming, and Prolog as relevant technologies.

The knowledge base could be coded so as to print a trail of the deductive process. An example of this is

```

technology (theorem proving) :-
  print ("theorem proving is used in
  problem solving"),
  objective (problem solving).
technology (logic programming) :-
  print ("logic programming is used
  in problem solving"),
  objective (problem solving).
technology (logic programming) :-
  print ("logic programming is used
  in expressing programs"),
  objective (expressing programs).
technology (logic programming) :-

```

```

  print ("logic programming is used
  in expressing specifications"),
  objective (expressing specifica-
  tions).

```

```

technology (Prolog) :-
  print ("Prolog is used in logic pro-
  gramming"),
  technology (logic programming).
procedure (resolution):-
  print ("resolution is used in Pro-
  log"),
  technology (Prolog).
procedure (resolution) :-
  print ("resolution is used in theo-
  rem proving"),
  technology (theorem proving).

```

Now, querying for the uses of the resolution procedure:

```
procedure (resolution)?
```

would result in the reply:

```

resolution is used in Prolog
Prolog is used in logic programming
logic programming is used in prob-
lem solving
logic programming is used in express-
ing programs
logic programming is used in express-
ing specifications
resolution is used in theorem proving
theorem proving is used in problem
solving

```

In information retrieval, however, citations are the objective. In knowledge-based information retrieval, then, the trail of sources for the logic involved in the deduction is a by-product of primary importance. The following version of our knowledge base is further developed so as to yield a citation list in response to a query:

```

technology (theorem proving) :-
  print ("theorem proving is used in
  problem solving"),
  citation (1),
  objective (problem solving).
technology (logic programming) :-
  print ("logic programming is used
  in problem solving"),
  citation (1),
  objective (problem solving).
technology (logic programming) :-
  print ("logic programming is used
  in expressing programs"),

```

citation (1),
 objective (expressing programs).
 technology (logic programming) :-
 print ("logic programming is used
 in expressing specifications"),
 citation (1),
 objective (expressing specifica-
 tions).
 technology (Prolog) :-
 print ("Prolog is used in logic pro-
 gramming"),
 citation (2),
 technology (logic programming).
 procedure (resolution):-
 print ("resolution is used in Pro-
 log"),
 citation (2),
 technology (Prolog).
 procedure (resolution) :-
 print ("resolution is used in theo-
 rem proving"),
 citation (1),
 technology (theorem proving).
 citation (1) :- print ("Kowalski
 (1979)").
 citation (2) :- print ("Clocksin and
 Mellish (1984)").

Now the presentation of the query:

objective (problem solving), technol-
 ogy (_What)?

would result in the reply:

_What = theorem proving.
 Kowalski (1979)
 _What = logic programming.
 Kowalski (1979)
 _What = Prolog.
 Clocksin and Mellish (1984)

Similarly, querying for the uses of the reso-
 lution procedure:

procedure (resolution)?

would result in the reply:

resolution is used in Prolog
 Clocksin and Mellish (1984)
 Prolog is used in logic programming
 Clocksin and Mellish (1984)
 logic programming is used in prob-
 lem solving
 Kowalski (1979)
 logic programming is used in express-
 ing programs

Kowalski (1979)
 logic programming is used in express-
 ing specifications
 Kowalski (1979)
 resolution is used in theorem proving
 Kowalski (1979)
 theorem proving is used in problem
 solving
 Kowalski (1979)

Further modifications of this last version
 of the logic program could incorporate the
 conventional citation-type knowledge
 about the documents. The Prolog for this
 information was developed earlier in this
 paper. The citation trail could then list
 more complete bibliographic information,
 and/or the queries could specify the form of
 the citation list.

DISCUSSION

The process of converting knowledge
 into a formal, logical form is certainly not
 new: it has been done since the time of the
 ancient Greeks; any introductory textbook
 on logic includes such exercises. What is
 new is the facility to process the representa-
 tion as logic in the computer. Early work in
 artificial intelligence recognized the poten-
 tial of theorem-proving programs for ap-
 plication in information retrieval (Cooper,
 1964, for example). More recent work in
 theorem proving has resulted in logic-
 programming systems that are a practical
 support for knowledge-based information
 retrieval. Prolog is currently the best devel-
 oped and most efficient logic-program-
 ming system. (Walker describes Prolog's
 utility and facility for knowledge-based
 construction and use. Clocksin and Mellish
 provide an introduction to the language it-
 self. Sergot and others report on efforts to
 capture a legal document in logic-
 programming form; Leith critiques this ef-
 fort.)

The knowledge-based retrieval de-
 scribed here is a fundamentally different
 process from conventional information re-
 trieval. The keyword-based indexing and
 retrieval of conventional systems is surface-
 oriented and passive compared to this re-
 trieval process based on knowledge refine-
 ment. This knowledge refinement entails a
 transformation of the actual knowledge in
 the abstracts into a "computer-reason-

able" form that results in subsequent automated reasoning. In contrast, keyword-based methods manipulate unstructured words and phrases only. Conventional retrieval is by simple matching or—in the case of expert system-aided retrieval (such as described in Waters or Kehoe)—by rule-extended matching on these keywords, with no attempt to deal directly with the actual meaning. Natural language information retrieval (such as described by Biswas and others) merely imposes a natural language interpreter as an interface for the conventional keyword-based and rule-extended systems.

In practice, knowledge refinement of abstracts into Prolog is not as esoteric or difficult as it may seem. Capturing every nuance of meaning, which is necessary for a legal expert system, is unnecessary here. Only a small number of basic relationships need be captured, and these can be expressed in a small number of facts and rules for each citation item included. In information retrieval the primary interest is in the trail of citation references that results from querying and not necessarily in the answer that is derived. In most cases, the answer from a knowledge base could be totally wrong and still yield a set of references superior to that attainable with keyword-

based methods.

Large-scale application of knowledge-based information retrieval would require the development of logic to accompany items included in bibliographic databases. This logic could result from automatic or manual refinement of abstracts. Alternatively, authors could supply a logic abstract in addition to the conventional prose abstract and keyword list. The vocabulary problems in synthesizing the variously produced logic abstracts into a homogeneous database would be solved by the control techniques of conventional systems.

Immediate application of knowledge-based information retrieval is a possibility for individual research projects. The development of a Prolog knowledge base encompassing a particular aspect of about 100 sources is well within the capability afforded by microcomputer Prolog implementations. Such a knowledge base could be established initially by manual coding of the content of abstracts. As work progressed, the original facts and rules could be augmented with knowledge found in the body of the documents as well as with discovered rules that relate multiple documents and unify the knowledge base. Vocabulary for a single project of a single researcher could be managed informally.

BIBLIOGRAPHY

- Biswas, Gautam, and others. "Knowledge-Assisted Document Retrieval: 1. the Natural-Language Interface." *Journal of the American Society for Information Science* 38, no.2:83-96 (1987).
- Clocksin, W. F., and C. S. Mellish. *Programming in Prolog*. 2d ed. (New York: Springer-Verlag, 1984).
- Cooper, W. S. "Fact Retrieval and Deductive Question Answering Information Retrieval Systems." *Journal of ACM* 11, no.2:117-37 (1964).
- Kehoe, Cynthia A. "Interfaces and Expert Systems for Online Retrieval." *Online Review* 9, no.6:489-505 (1985).
- Kowalski, R. *Logic for Problem Solving* (New York: North Holland, 1979).
- Leith, P. "Fundamental Errors in Legal Logic Programming." *The Computer Journal* 29, no.6 (1986).
- Michie, D. "A Prototype Knowledge Refinery." In *Intelligent Systems*, p.57-69. J. E. Hayes and D. Michie, ed. (Chichester: Ellis Horwood, 1984).
- Sergot, Mark J., and others. "The British Nationality Act as a Logic Program." *Communications of the ACM* 29:370-86 (1986).
- Walker, Adrian. "Knowledge Systems: Principles and Practice." *IBM Journal of Research and Development* 30, no. 1:2-13 (1986).
- Waters, Samuel T. "Answerman, the Expert Information Specialists: An Expert System for Retrieval of Information from Library Reference Books." *Information Technology and Libraries* 5, no.3:204-12 (Sept. 1986). ■■

Library Automation with Workstations: Using Apple Macintoshes in a Special Library

Edward J. Valauskas

The Merriam Center Library introduced its staff to Apple Macintoshes as library workstations over an eighteen-month period, as part of a large-scale automation project. Each unit was brought in to handle a specific chore such as serials management, monographic acquisitions, or online access. Staff response to these computers was enthusiastic, easing the transition from keeping records with paper files to computerized database management. The success of these workstations as vehicles for automation is largely due to their ease of use, interactive interface, and interconnectivity.

With the rise in the use of personal computers in libraries over the past decade, there have been increasing opportunities to test their feasibility in a variety of work environments. Workstations, such as those offered by OCLC, have played a restricted role in libraries, largely because of their dependence on mainframes for computing power. With the rise in the computational strength of personal computers, true multi-tasking workstations are becoming a reality for libraries (for a comparison of OCLC workstations to the current IBM and Apple personal workstations, see appendix A). Of course, there are only a small number of libraries that would take the risk of using a group of interconnected computers to function as an automation base. Large public and academic libraries, with heavy investments in mainframe-based bibliographic files, would find library workstations an expensive and perhaps infeasible option. Small private, special, and public libraries would be hard-pressed to defend the choice

of a workstation as an automation vehicle if access is available to large bibliographic databases through dedicated and inexpensive terminals. Libraries in that netherworld between large and small collections, serving specialized yet vocal clients, are in the best position to consider personal computers and workstations. It was in this spirit that the Merriam Center Library embarked on an ambitious project to automate its operational functions solely around personal computers as workstations, which are literally positioned at every staff member's desk. The Apple Macintosh was chosen as the primary tool for this project.

THE WORKSTATION EXPERIMENT

The library of the Charles E. Merriam Center for Public Administration (commonly known as the Merriam Center Library or MCL) serves the staffs of the American Planning Association, American Public Works Association, International

Association of Assessing Officers, and researchers from the academic community and public sector in the Chicago area. The fifty-year-old library's collection contains some 50,000 monographs, 100,000 reports, and more than 2,000 serials, focusing on local and regional governmental issues. The staff of seven FTEs, including two professionals, operate the library.¹

BACKGROUND

All operational activities were based on paper files and manual records until 1981, when MCL developed a program in conjunction with the University of Chicago's Computation Center to produce MCL's bibliographic monthly, *Recent Publications on Governmental Problems*, and generate a book catalog. This project resulted in a complex program that sorts and formats bibliographic data, producing hard copy for publication.

MCL's staff has experienced continual problems with the program. With little documentation, staff turnover, and a lack of in-house programming expertise, there was an appreciable unfamiliarity with the idiosyncrasies of the program. Bugs in the program, communication errors between the library and the university mainframe where the program was mounted, and staff impatience meant lost files, inaccessible bibliographic records, and expenditures for unanticipated computer time to reconstruct records. Among some staff members there was a remarkable level of hostility to the program, the mainframe environment specifically, and automation in general. This ill will toward technology was only matched by the frustration of other staff members tediously maintaining massive paper files. It was into this atmosphere that personal computers were introduced to revamp the library's day-to-day operations.

COMPUTING REQUIREMENTS

What kind of computers were thought to be appropriate to this sort of environment? First of all, they had to be easy to use and highly reliable. Funds were not available for formal training or frequent repairs. Secondly, software and hardware had to be well documented and digestible for even the most technophobic worker. In tandem

with this requirement, off-the-shelf software would need to function as the library's primary operational vehicle; again, monies were not available for hiring programmers to create library-specific, customized software. Third, the library desired personal computers that could function as interconnecting workstations sharing peripherals and software. It was impossible to buy multiple pieces of hardware and copies of software. Each personal computer would be purchased singly; the success of each unit determined where the next computer would be added. It was decided that the Apple Macintosh, especially its interactive features, met many of the hardware requirements, and that its collection of software answered most of the library's initial needs. It was also agreed that the potential of the Macintosh for technological growth into a full-fledged workstation was quite high.

IMPLEMENTATION

The first unit was set to manage a portion of the library's serials collection, specifically some 800 titles that are routed to 125 researchers in the Merriam Center.² Each year, the staff prepared a union list of all titles for each researcher as well as a specific list of items routed to each individual. Sandwiched between their other duties, collecting this information manually from the paper files normally took several staff members a minimum of two months to accomplish. With the assistance of the serials technical assistant, information was transferred from the paper file onto a form describing fields that was set up in a template created on a Macintosh with the Microsoft (MS) *File* program. The entire transfer and production of route lists were accomplished in less than two weeks.

The initial master file for serials was set up by creating in Microsoft *File* a separate indexed field for each defined category in the original paper file, which meant creating fourteen separate fields (title, ISSN, publisher, publisher address, zip code, frequency of publication, main subject heading, subscription expiration, price, retention, route list, check-in, notes, and source), each accessible through MS *File*'s Find and Sort functions. All fields were for-

matted as text fields except subscription expiration (a date field) and price (formatted as a number field).

The flexibility of *MS File* allowed the introduction of several unique fields such as source and retention. The source field provides information on the source of a given title. Only 17 percent of the library's routed titles (some 800 in all) are received through paid subscription. Most are received so that their contents can be indexed in the library's bibliographic monthly *Recent Publications on Governmental Problems*. The retention field informs the staff and clients how long the library retains a given title; since space is limited, few periodicals are retained indefinitely.

STAFF INTERACTION

Flushed with the success of the route list program, the first Macintosh was set up as a serials workstation, operated entirely by a paraprofessional. His enthusiasm was remarkable; within a matter of a month, he had generated a number of subfiles from the master serials file, managing a variety of tasks formerly within the province of the paper file. His eagerness was also contagious. Other staff members watched the creation and growth of his automated files and began to discuss ways in which the program or the Macintosh could be applied to their tasks. Staff meetings became more frequent and were enriched by detailed descriptions of potential applications for future Macintoshes. Schedules were created so that time for experiments could be shared on the Macintosh while utilizing it and its programs to handle other chores. The staff wrote articles describing their experiences creating templates and files. Staff members were spending more time at work, both into the evenings and on weekends, to justify their needs for workstations.

CONFIGURATION OF THE WORKSTATIONS

Over an eighteen-month period, workstations were eventually acquired to meet the demands of all staff members. Units were added only when all personnel agreed to the necessity for a new unit in a given location, usually because of problems caused by sharing an existing unit for several func-

tions. Currently MCL uses Macintoshes as workstations for the following tasks (see figure 1):

1. to manage over 2,000 serials;
2. to handle the acquisition of monographs and pamphlets;
3. to prepare monthly issues of MCL's *Recent Publications on Governmental Problems* and access bibliographic files stored on a University of Chicago mainframe;
4. to format and edit texts of bibliographies, issued bimonthly by MCL for the Council of Planning Librarians;
5. to keep circulation records for all monographs, pamphlets, and serials;
6. to access bibliographic databases, such as Dialog for clients and OCLC for cataloging; and
7. to handle day-to-day correspondence and billing.

These chores are handled with the following equipment:

- two 512K enhanced Macintosh computers, all equipped with 800K external drives;
- two Macintosh Plus computers, both with 800K external drives and one with an attached 20M Apple hard disk;
- two Macintosh SE computers with 20M internal hard disks, one of which is outfitted with a PC card and a 5.25" drive to read and translate manuscripts submitted as MS-DOS files;
- an ImageWriter II with an Apple Talk board;
- a LaserWriter Plus; and
- three modems.

Four of the Macintoshes are located in the library's administrative and technical offices, along with the printers and several modems. All of these workstations are linked together with AppleTalk, permitting all equipment to share peripherals. Two Macintosh units are located in the library's reading room, along with a modem, where they are used for circulation and reference. A small number of software programs are used; they include Apple's *MacWrite*, *MacPaint*, *MacTerminal*, *MacDraw*, and *MacProject*; Microsoft's *File* and *Word*; and several telecommunications packages, including *Red Ryder*, *Kermit*, and *MicroPhone*.

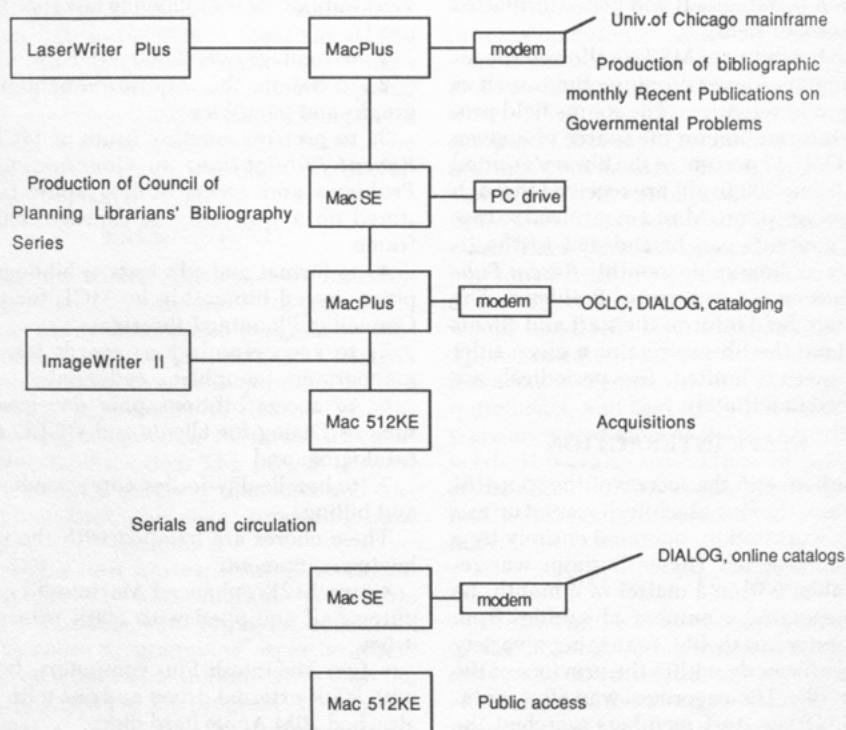


Fig. 1. *Macintosh Configuration in the Merriam Center Library*

FUTURE STRATEGIES

Staff enthusiasm for the workstations and their potential remains quite high. Over the next eighteen months, the library will need to add a file-sharing system and a large, on-site storage device to handle the burgeoning files and subfiles that have been created. The library's patrons have also asked for access, and the staff is currently investigating the means for setting up an electronic bulletin board that would permit twenty-four-hour access to some of the library's files and serve as a depository for reference inquiries. There is also some demand from the staff for upgrading all equipment to the minimum level of the Macintosh SE and for acquiring several Macintosh IIs. Given the staff's responsiveness to these computers, it will be quite interesting to see how they respond to even more powerful, interconnected units.

CONCLUSION

Given the initial conditions of the experiment in the beginning of 1986, it was difficult to imagine how the MCL staff would respond to a new way of doing things with a technology in which they had very little faith. By involving the staff in the entire process, making it highly dependent on their response to the technology occupying their desktops, automation was not treated simply as an object that was replacing old routines or reinventing them in a new format. This style of automation personally involved staff members by allowing each person the freedom to create individualized files based on common software and, in turn, a unique, interactive style with the workstation. Hence there was a certain degree of effort and pride in creating each workstation and demonstrating the utility of its product.

Perhaps only a small library with limited resources could try to demonstrate the potential of workstations with an initially skeptical staff. Nevertheless, given the computational power of some of the recently announced workstations, there may be an opportunity for other, larger libraries to experiment with staff involvement in at least some aspects of automation. It is dan-

gerous to underestimate the value of human resources and their potential contributions to the application of new technology. Fortunately, the Merriam Center Library was able, in a short period, to take advantage of a technology and discover the wealth of previously hidden resources in its staff.

REFERENCES AND NOTES

1. See also *The Macintosh Library: Uses of the Apple Macintosh in a Special Library*, Edward J. Valauskas ed. (Cupertino, Calif.: Apple Library Users Group, 1987), p.1-3.
2. See also Edward J. Valauskas, "Periodicals Management with the Macintosh in a Special Library," *Apple Library Users Group Newsletter* 4:25-27 (Oct. 1986). ■■

APPENDIX A: A COMPARISON OF CURRENT MICROCOMPUTER-BASED (PERSONAL) WORKSTATIONS AVAILABLE TO LIBRARIES.

In July 1987 OCLC announced the features of its newest workstation, the M310.¹ This unit is based on the Intel 80286 processor and is available in two configurations (labeled A and B on the accompanying table).² The workstation is anticipated to satisfy OCLC's clients well into the 1990s and be replaced by then-developed hardware and software based on the Intel 80386 processor.³ This design was based on at least a year's worth of research and the examination of microcomputers from eight vendors. Among some of the guiding characteristics, OCLC demanded that the units consist of widely available components that are easily expandable, highly reliable, well documented, and highly tolerant of a wide range of environments.⁴ These features were thought to be embodied in IBM PC-compatible, Intel 80286 processor-based computers.⁵ Table 1 compares some basic features of the OCLC work-

stations with others derived from personal computers.⁶

The superiority of the Apple and IBM workstations is clearly obvious in all categories. Unfortunately, OCLC's involvement in developing its workstation coincided with the development of the others, so that these models were probably not an option for OCLC. Nevertheless, some evidence suggests that OCLC consciously decided not to use the more advanced Intel 80386, which is some 250 percent faster than the 80286. This decision seems to have been based on a perceived lack of software for full exploitation of the 80386 in the library market. The argument depends on the development of operating systems that take full advantage of the 80386, so that appropriate, library-specific software can be created.⁷ It has been suggested that the M310 will function as a workstation into the 1990s. Predictions regarding

Table 1. Comparison of Current Personal Workstations Available to Libraries

	OCLC M310 A	OCLC M310 B	IBM PS/2 Mod80	Macintosh II
Operating System	MS-DOS	MS-DOS	OS/2	Mac/OS
Processor	Intel 80286	Intel 80286	Intel 80386	Motorola 68020
Internal Memory Std/Max	256K/640K	640K/1M	1M or 2M/16M	1M/1.5G
Hard Disk	10M to 40M	20M to 40M	44M to 115 M	20M to 80M
Internal Drive Slots	two floppy	one floppy, one hard	two floppy, two hard	two floppy, one hard

typical workstations at that time include use of 64-bit processors, a minimum RAM of 3M to 6M, and 100M of storage.⁸ The M310 would hardly compete technically or economically with such

devices, especially with prices for more sophisticated computers at around \$6,000 per unit or less.

APPENDIX REFERENCES AND NOTES

1. Described in "Future Workstations and Terminals," *OCLC Communications & Access: In Development*, OCLC Pub. no.66 (Dublin, Ohio: OCLC, Inc., July 1987), p.1.
2. *Ibid.*; See also OCLC, Inc., *OCLC's New Family of Workstations: Status Report-February 1987* (Dublin, Ohio: OCLC, Inc., 1987), p.5.
3. *Ibid.*, *OCLC's New Family of Workstations*, p.4-5.
4. *Ibid.*, p.3.
5. *Ibid.*, p.3-4.
6. Data on M310 workstations from "Future Workstations and Terminals," p.1-2 and *OCLC's New Family of Workstations*, p.5. Data on IBM Personal System/2 Model 80 and Apple Macintosh II from Alan Paller, "Clash Provokes Burst of Change," *Computerworld* 21:S3 (July 27, 1987).
7. "Despite the constant and rapid technological changes in personal computer hardware, the 80286 based machines are expected to be a good choice for OCLC and its users for the next 3-5 years. This is not a crystal ball prediction, but rather a recognition of the evolutionary nature of personal computers and their [sic] use. . . . hardware continues to outpace the software (operating system) [and] is ahead of the capabilities of the applications. Once the operating system catches up with the capabilities of the 80286 processor, then OCLC can begin to develop more sophisticated applications to meet user needs. Beyond the 80286, the 80386 processor may be useful for some multi-user file server type systems that require the capacity and performance that it can deliver. However, the 80386 is too powerful at this time to be considered for the basic workstation required for OCLC applications." (OCLC, Inc., *OCLC's New Family of Workstations*, p.4-5).
8. Paller, S12; Reed McManus and others, "Special Report: The Future of Computing," *PC World* 5:260-73 (May 1987); Frederic G. Withington, "Technology Forecast," in *Office Workstations in the Home* (Washington, D.C.: National Academy Pr., 1985), p.105-8. ■■

Twenty Years Ago in JOLA

The *Journal of Library Automation* publishes original contributions in all fields of research and development in library automation, including interlibrary communications; in research in information science directly related to library activities; and in the history and teaching of these subjects.

Scope statement, *JOLA* (Mar. 1968).

Reports and Working Papers

Back to the Concept: Perspectives on Series Authorities

Mary Dabney Wilson

Editor's note: The following paper was presented at the meeting of the LITA/RTSD CCS Interest Group on Authority Control in the Online Environment during the ALA Annual Conference in San Francisco, July 1987. The response from the Library of Congress, with a request for comments, follows this paper.

INTRODUCTION

Like the ugly stepsister whose foot is too large, series authority control seems to present more problems than will comfortably or easily fit into the glass slipper of most automated authority control systems. In order to get back to the concept, I plan to discuss the development of our thinking on series both as an element of the bibliographic description as well as an access point. Some of the problems occurring in series authority control derive from the fact that with series we frequently ask an element of description to do double duty and perform also as an access point. Because most catalogs existing today contain series headings that have been constructed according to different cataloging codes, it is important to stop and consider how we have come to our present rules for series description and series access. What we have done in the past with series has produced a

Mary Dabney Wilson is Head, Bibliographic Control Department, University of Texas at Arlington.

decided impact on today's catalogs and authority control systems.

I do not plan to discuss the usefulness of series as an access point. Opinions on this vary, but at least in the sciences and frequently elsewhere, the series is used to retrieve, albeit not as often as other access points. My assumption for this discussion is that series can serve as useful retrieval tools in a catalog, or, in AACR2 terms, that some series "provide a useful collocation." Therefore, I will be concentrating on the means by which this collocation has historically been carried out and some of the specific problems these practices have brought about in automated series authority control.

BACKGROUND

Authority control refers to the principle of uniform heading, that is, the principle that dictates that only one form of heading be used for each unique entity. The reason we are concerned with uniformity and uniqueness of entry forms derives from the purposes that catalogs serve. Cutter's Objects of 1876, reiterated by the Paris Principles of 1961, remain the rationale for authority control today. Helen Schmierer has perhaps best summarized these two principles:

First, the library catalog should enable a user to ascertain if the library has a particular item. In other words, the catalog should be able to respond to "known item" searches.

Second, the library catalog should show what items the library has that share a common characteristic. By this we mean the catalog should be able to respond to "category" searches.¹

Authority control exists to fulfill the second purpose, that is, to respond to "category" searches. Without authority control,

it is not possible to be certain that you have displayed *in one place* all materials sharing a common characteristic. Authority control, then, is concerned with those access points which can be shared by multiple bibliographic entities, and series by definition fall within that category.

THE PAST

The rules governing the recording of series statements have for the most part been consistent. Each code from 1908 to the present requires that if a bibliographic entity is published as part of a series, that fact is to be recorded. How it is to be recorded has varied considerably. Similarly, whether or not an access point is to be made, and if so, in what form, has also varied. Jean Decker made clear the distinction between the recording of the series and the creating of access by saying: "A series statement is not an access point, it is merely description. The terms 'series heading,' 'series entry,' or 'series added entry' refer to the series as an access point."² The terms *series note* and *series area* refer to the description.

The rules that preceded AACR2 generally could be said to have mixed issues of description and access. By that I mean that decisions as to which access points were to be made affected decisions as to how the item would be described. To illustrate, let us take AACR1 rule 134A. This rule had nothing to do with series but was the rule for recording author statements. In that rule were eight cases where the author statement could be omitted from the catalog entry; each case related to the choice, and sometimes the form of, name used in the main entry heading. For example, the author statement could be omitted entirely from the bibliographic description if it matched the form of name used in the main entry heading, or it could be omitted when the form of name in the heading for a corporate body was the same as that in the author statement except for the fact that the body was entered under place. This is what I mean by mixing issues of description with issues of access.

In all of the major U.S. codes before AACR2, issues of series access preceded issues of series description. The cataloger for the most part would first decide whether

access would be provided for the series and then, based on that decision, record series data in the series note. In *Catalog Rules: Author and Title Entries* (1908), the basic principle was to enter a series under title unless it was universally known under the name of the editor or publisher. For the series note, most of the instructions concerned how to record the location of the series statement within the publication being cataloged. For instance, you would see phrases such as "Half-title: . . ." or "Added t.p.: . . ." preceding the series. If the series entry were other than title (i.e., under author, editor, publisher, or society) that name was to be given before the series title in the series note.

The next two major U.S. codes, *Rules for Descriptive Cataloging in the Library of Congress* and *ALA Cataloging Rules for Author and Title Entries* (both 1949), brought the mixing of series description and series access to its fullest form. Rule 3:16 in *Rules for Descriptive Cataloging* contained the instructions for recording series statements. Following it was a listing of special situations. Rule 3:16B required the cataloger to decide first whether there would be a series access point made before the decision could be made as to which of two or more forms would be recorded in the series description. There were special instructions for three other situations which cause problems today. First, where the series title was combined with the title of the monograph, the cataloger was instructed to separate the series and give it in the regular position.

From morality to religion, being the Gifford lectures delivered at the University of St. Andrews, 1938.

To be recorded as:

Title: From morality to religion
Series: Gifford lectures, 1938.

Second were the instructions for a series of works by one author, where if the author appeared preceding the series note, the "appropriate possessive pronoun" was substituted for the author's name.

(*Its* Publications in research and records)

Third, the same rule further specified that if a corporate body was considered to

be the author of the series and not of the particular part being cataloged, it was included in the series note in entry form followed by the title.

(Kentucky. [State Geologist] Bulletin no. 20)

Rule 3:25 gave instructions on making the series entry or access point, and here it was first codified that if the series were to be entered as it appeared in the series note, only the word "Series" was given in the tracing. Of course, any variation between the series note and the series entry was to be shown explicitly.

What the *Rules for Descriptive Cataloging* did was to give general instructions for the recording of series, followed by three major exceptions: the separation of titles included with the title of the monograph, the possessive pronoun form, and the corporate name in entry form followed by the title.

The *ALA Cataloging Rules for Author and Title Entries* (1949) of course gave no instruction for the recording of the series statement but did contain the rules for their entry. Rule 5F said to enter a series under its title with two exceptions. The first, like the 1908 rules, was to enter under an editor or publisher a series which is familiarly known by the name of the editor or publisher, and second, enter under the name of a society, institution or other corporate body, a series issued by it without distinctive title.

With *AACRI* in 1967, matters regarding series description and series access did not improve. The provisions for separating series titles from titles of monographs, for the use of the possessive pronoun, and for including the name of the author in the series description preceding the title of the series were all the same.

The muddy demarcation between access and description began to clear with the work on the International Standard Bibliographic Descriptions. The first preliminary edition of *ISBD (M)* in 1971 and the first standard edition in 1974 were simple and direct for series description: the series title was to be given as it was found in the publication. There was no exception for the substitution of the possessive pronoun, nor for corporate bodies as authors. However, in the *ISBD* as it manifested itself in chapter 6

of *AACRI*, two of the three exceptions had reappeared. The rules for the Gifford lecture situation and for the possessive pronoun were back. There were no more author/title series, but for generic titles the author's name was given following the title—a precursor to *AACR2*'s paralleling of series area with title and statement of responsibility area.

The *ISBDs* began to compel catalogers to divorce description from access. With *AACR2* the separation is almost complete. For Areas 1-6 (title and statement of responsibility thru series) there are no more situations where description issues are related to access. However, there are two access points which we still relate to descriptions, those being the title and the series. These of course, are not dictated by *AACR2*. *AACR2* does not attempt to say how access points are to be made or indicated on a cataloging record. *AACR2* only instructs that the cataloger make certain access points. Catalogers have relied on previous rules primarily aimed at the card format to record what access points have been made and in what form they have been made. For the two access points related to descriptions, we still are thinking in terms of the "roman numeral" tracings "Title" and "Series" codified in 1949 and used when the title or series is to be traced as it appears in the bibliographic description. These conventions are not required by the *MARC* format. The *MARC* format is hospitable to either implicit or explicit title and series access. For series you may give implicit series access in the 440 field or explicit access in an 8xx field.

To demonstrate some of the problems that occur when a description field is used for access, let us take an example of a series title traced implicitly: "The Columbia history of urban life." When humans arrange cards in a file, it is not a problem to interfile series which appear in one case as "The Columbia history of urban life" and in another case as simply "Columbia history of urban life." In a machine file, because of the filing indicator available in the 440, it also should not be a problem to *arrange* these two records properly on a series display. The problem comes when you have linked authority control. In the *MARC Authorities Format*,

the entry form for most series is recorded in the 130 field. While the Authorities Format contains an indicator for nonfiling characters, the Library of Congress has decided to omit initial articles in all uniform titles (LCRI 25.3A and 25.4A). This can create problems in a catalog with linked authority control.

Most programs for linking bibliographic records with authority records can recognize that the link should be made when loading a new record, but in those systems where field text is stored only in the authority file, as it is in many systems, when the record is reconstructed for display, you get the version of the series that was on the authority record, not necessarily what was on the piece.

In systems where the field data are stored in both the authority file and the bibliographic records, you may not have the problem just mentioned, but in both types of systems, if you find it necessary to change the data in the authority record, when that change is communicated globally to the bibliographic records, you may find you have changed data in a description field to be in a form appropriate only for access. For instance, you may have a series authority record with links to bibliographic records and find that the Library of Congress has established this particular series with a qualifier. If you change your first authority record to agree with the Library of Congress version, and if that change is communicated globally to the attached bibliographic records, then you will be substituting data in a 440 "description field" for data which is only appropriate to an 830 "access field." There are labor intensive methods of getting around this problem in some systems, and there are systems which are smart enough to recognize and properly deal with changes of this sort, but the important point to make is that these types of problems occur only in automated series authority control, not with other types of authority control. Why? Because we ask a field of description to function also as an access point.

The problems just mentioned can occur in a purely AACR2 catalog. But the combination of series description with series access under previous cataloging codes create

even more serious problems for automated authority control. Many of the series previously entered under corporate bodies cannot pass AACR2 rule 21.1B2 and are therefore entered under title proper. Those with common or generic titles now require the addition of a qualifier. Automated authority control systems can have difficulty processing these changes. Some simply cannot handle changes across tag groups, that is from a name tag (410 for author/title series) to a title tag (440 or 830). For those systems that can, the results are not always desirable. When a straight substitution of text or "flip" occurs, you may get a series uniform title construction in a 4xx field.

If the 410 name/title problem is difficult to resolve, even more difficult for the machine to process is the possessive pronoun construction. Consider the situation of the *Instituut voor Toegepaste Sociologie te Nijmegen*. The authority record for this corporate body has two troublesome references. If your source database is OCLC, you may have already found in your catalog some remarkable series access points. These access points were created because of standard processing that OCLC does when loading revised MARC records. Each name heading is passed by the Name Authority File, and when a 410 subfield 'a' with *Its* occurs, it is automatically switched to *Instituut voor Toegepaste. . .* Although OCLC is aware of this problem, they are unable to identify the affected records and have asked to have them reported. If the second indicator in the 410 is properly used, it is certainly possible to program a correct substitution for the data in the 1xx field of the bibliographic record, but then you are faced with the change across tag groups mentioned earlier.

It is easy to be critical of our past practices, but I feel certain that the reasons for what today may appear as arbitrary manipulations of descriptive data were in fact based on a valid principle: conservation of resources. If the series note could double as an access point, then time and space were saved. This was the only virtue of the "Its" technique. Such practices might be justified in a manual catalog where "cataloging shorthand" was interpreted and manipulated by human beings and where shorten-

ing the length of records was a significant consideration, but they are counterproductive in an online environment.

We have seen some of the problems created when series description also serves as access. From past practices we must contend with the "Its technique" and name/title constructions. From current practice, we have the omission of the initial article in established series and the trouble encountered when an existing series requires a change such as the addition of a qualifier. What is the solution?

The solution is to completely divorce series description from series access. Indicate explicitly all authority-controlled access points in their entry form. Leave the description to serve its designated function, to describe. Whereas the rules for access point choice and form can be expected to change in the future, a properly formulated description of a bibliographic entity which is complete should never have to be modified. It is the beauty of the online catalog that changes in access point form can be handled easily, but only if they are separate from the description.

There will be those who have card format catalogs who will worry about the lengthening of the cataloging record. I'm quite certain I remember when ISBD was adopted that many were voicing those same concerns. Remembering AACR 134, they would question the provisions for giving the author statement as it appeared in the work in all cases. "Why restate the author, particularly lengthy corporate names with hierarchies, when those names already appear in the main entry?" We have all coped with that decision and have for the most part, I believe, seen the validity of separating description from access. Can we take this final step and do the same with series? If we can, the glass slipper of linked authority control may actually fit for series as it already does for other headings. ■■

REFERENCES

1. Helen F. Schmierer, "The Relationship of Authority Control to the Library Catalog," *Illinois Libraries* 62:599-603 (Sept. 1980).
2. Jean S. Decker, "AACR2 and Series," *Cataloging & Classification Quarterly* 3, no.2/3:59-63 (Winter 1982/Spring 1983).

Library of Congress Comment and Call for Response

Judith G. Fenly

The Library of Congress (LC) is considering changes in order to resolve the issue, raised in Mary Dabney Wilson's paper, of using the 440 field for both description and access. The proposal is directed to the problems created by the aforementioned dual role in a system where the authority file is linked to the bibliographic file.

LC is willing to consider discontinuing the use of the 440 field altogether and begin using the 490 field with the first indicator set to 1 in lieu of the 440. LC would also then propose to change the definition of indicator value 1 from "Series traced differently" to "Series traced." Series tracings (830) would be created for all 490 series with indicator value 1, regardless of the fact that in some cases the 830 would duplicate the 490. LC's proposed change would cover only prospective cataloging, leaving a mixed practice regarding series traced the same in the files. There would be no systematic update of the LC database.

This proposal was announced by Lucia J. Rather, Director for Cataloging, LC, at a meeting of the LITA/RTSD CCS Interest Group on Authority Control in the Online Environment during the ALA Midwinter Meeting in San Antonio. Rather asked for a straw vote regarding the LC proposal: about half of the audience was immediately in favor of the proposal, another half abstained, and a few voters were immediately opposed to the proposal. Rather announced that LC would be continuing its discussion of this proposal through spring and invited letters of comment to Sally H. McCallum, Chief, Network Development and MARC Standards Office, Library of Congress, Washington, DC 20540. ■■

Judith G. Fenly is Automated Operations Coordinator, Library of Congress, Washington, D.C., and Chair of the LITA/RTSD Authority Control in the Online Environment Interest Group.

We've got you covered.

EBSCO has all your serials needs covered! We are committed to maintaining the largest database of serials titles in the subscription agency business and are the one source for almost any serial published anywhere in the world. Our title file currently contains information on over 45,000 publishers. And, because we are as committed to keeping your customer service and serials management needs covered as we are to increasing our database of serials titles, we operate 23 regional processing offices throughout the world. At EBSCO, in addition to a qualified sales representative,

your account will be serviced by a knowledgeable customer service representative and an entire regional office staff having a strong commitment to provide you with sound customer service.

Isn't that what you expect from a professional subscription agency?

EBSCO
SUBSCRIPTION SERVICES

P.O. Box 1943
Birmingham, Alabama 35201
(205) 991-1182

News and Announcements

Hugh Atkinson Memorial Award

Nominations and applications are sought by May 15 for the Hugh C. Atkinson Memorial Award, newly established to honor Atkinson's life and accomplishments. The award will recognize outstanding achievements by academic librarians, working in automation or management, who have improved library service, development, or research.

To be given for the first time at the 1988 ALA Annual Conference in New Orleans, this annual award consists of an unrestricted \$2,000 cash prize and a plaque.

Individuals may nominate themselves or be nominated by others. To be eligible for the award the nominee must be a librarian employed in a university, college, or community college library during the year prior to application for the award and must have at least five years of professional experience in an academic library. The nominee must have demonstrated achievement (including risk taking) related to automation or management that has contributed significantly to improvements in the areas of library automation, management, development, and/or research.

Those wishing to nominate someone (including themselves) for the award should write a letter outlining the ways that the candidate meets the above criteria. Letters should be accompanied by a current version of the candidate's vita and should be sent to Hugh Atkinson Memorial Award, ACRL/ALA, 50 E. Huron St., Chicago, IL 60611. *Deadline for nominations is May 15, 1988.*

This new award is funded by an endowment—which now stands at \$60,000—created by individual and vendor contributions given in memory of Hugh C. Atkinson. Additional funds are sought to bring the endowment to at least \$100,000. Send your tax-deductible contributions to

Hugh Atkinson Memorial Award at the address given above.

Hugh C. Atkinson was born on November 27, 1933, and died October 24, 1986. He was a graduate of St. Benedict's College and of the Graduate Library School of the University of Chicago and worked in the libraries of the University of Chicago, Pennsylvania Military College, State University of New York at Buffalo, Ohio State University, and University of Illinois at Urbana-Champaign (UIUC). He established a reputation as one of the major innovators in modern librarianship, especially in the last two institutions.

In the award proposal, Michael Gorman, a longtime friend and colleague at UIUC, stated

His ideas on library automation, on cooperation, and on the organization of libraries were a decade or more ahead of their time. In the UIUC library, one of the largest libraries in the world, he oversaw innovations in each of these areas. He has been described as the 'father' of the Illinois LCS network (the most developed statewide network of its kind in North America) and was certainly responsible for establishing an ethos of cooperation in the state of Illinois which has been of great benefit to the citizens of that state. Within the UIUC library he was responsible for the creation of a major online catalog and for the reorganization of the administrative structure of that library, which has rapidly become the paradigm for many other major research libraries. Those who knew him will remember his vision, his kindness, and his individuality.

The new award is jointly sponsored by the Association of College and Research Libraries (ACRL); Library Administration and Management Association (LAMA); Library and Information Technology Association (LITA); and Resources and Technical Services Division (RTSD), four divisions of the American Library Association.

For further information about the

award, contact Michael Gorman, Chair, Hugh C. Atkinson Memorial Award 1988, University of Illinois, 246A Library, 140 W. Gregory, Urbana, IL 61801; (217) 333-0318 or Mary Ellen K. Davis, ACRL Program Officer, ACRL/ALA, 50 E. Huron St., Chicago, IL 60611; (312) 944-6780. ■■

Dan Cochran Receives 1987 LITA/CLSI Scholarship Award

The Library and Information Technology Association (LITA), a division of the American Library Association, has awarded Dan Cochran the \$1,500 LITA/CLSI scholarship.

Cochran is enrolled in the library science graduate program at Kent State University. He made a significant contribution to the CD-ROM catalog project at Ritter Library of Baldwin-Wallace College in Berea, Ohio, working closely with Patrick Scanlan, library director, in its development. Ritter will be the first library in northeastern Ohio to implement a CD-ROM catalog. Cochran also worked on training and technical communications at Predicasts, a producer of online databases.

Explaining his career interests, Cochran said, "I view libraries as the ideal socio-technical catalysts to serve information needs across cultures, disciplines, and interests. And professional information specialists/librarians are the mediators of these unique services and opportunities."

The scholarship "is a cash award of \$1,500 made to a beginning student on the master's level in an ALA-accredited program in library and information science with emphasis on library automation."

The scholarship requirements include academic excellence, leadership, and evidence of a commitment to a career in library automation and information technology as well as prior career experience in those areas.

The scholarship is supported by a contribution from CLSI, Inc., Newtonville, Massachusetts, and administered by the Education Committee of LITA. ■■

LITA Program Assistant Appointed

Mary K. Simon joined the Library and Information Technology Association staff

as program assistant on November 16, 1987. Simon previously served for six years as director of information for the 23,000-member American Public Works Association (APWA). Her responsibilities included managing the association's library and information service; monitoring national, state, and local legislation pertinent to the public works field; serving as staff liaison to various association task forces; and conducting research for interest groups within the association, particularly the Research Foundation.

In addition, she served as the executive secretary/director for one of APWA's seven institutes for professional development—the Institute for Buildings and Grounds (IBG). In this capacity she directed all of the institute's programs and oversaw program and conference planning, including the sessions held at APWA's national conference. Simon represented IBG's interests and fostered ongoing relationships with other intergovernmental agencies that resulted in joint sponsorship of major research projects. She also undertook studies, conducted surveys, and prepared publications for dissemination to institute and association members and the public.

Simon received a master's degree in library and informational science from the University of Missouri-Columbia. She holds certification in the Chicago Public Relations Society and is active in several volunteer groups, including the Junior League of Evanston and the League of Women Voters. ■■

Aries Systems and Faxon Form Alliance to Distribute Medline Database on CD-ROM

Aries Systems Corporation of North Andover, Massachusetts, and the Faxon Company, Inc., of Westwood, Massachusetts, have formed a partnership and will offer subscriptions to the Medline Knowledge Finder CD-ROM database. This is a powerful and flexible bibliographic retrieval system that simplifies searching databases of citations to the biomedical literature.

Designed for both end-user and professional searching, the Knowledge Finder database contains full citations of medical journal articles, many with abstracts, with

up to five years of citations on a single optical disk. The service runs on Macintosh personal computers, and its easy-to-use search functions provide a variety of options, including natural-language search statements.

Knowledge Finder runs on Apple Macintosh Plus, SE, and II computers and is compatible with most CD-ROM disc players that support the Small Computer System Interface (SCSI) protocol. All Knowledge Finder CD-ROM discs use the "High Sierra" data directory structure.

Subscription plans include both annual subscriptions with multiple updates and onetime quarterly subscriptions. Prices range from \$595 to \$1,995 and include the Knowledge Finder software, software updates, documentation, and telephone support.

Aries Systems was founded in 1986 to develop innovative information retrieval and dissemination products, with special emphases in the fields of biomedicine and science. Medline Knowledge Finder is the first of these products. Aries works cooperatively with a variety of organizations in the information industry, extending the general information storage and retrieval architectures used in Knowledge Finder. ■■

Compact Disc of Education Materials in OCLC Database

OCLC has announced the availability of the Education Materials in Libraries (EMIL) compact disc database for use with OCLC's Search CD450 System. Compiled from the OCLC online union catalog, the EMIL database contains more than 450,000 bibliographic records pertaining to education.

The EMIL database is a comprehensive collection of education-related bibliographic records for materials printed during the twentieth century and more than 17,000 bibliographic records for materials printed prior to 1900. It also includes records for manuscripts, machine-readable data files, software, AV materials, musical scores, maps, games, flash cards, slides, sound recordings, filmstrips, and more.

The EMIL database complements previously published discs in the Search CD450 Education Series: ERIC, Current Index to

Journals in Education, and Resources in Education.

Each Search CD450 database uses the same software, allowing patrons and staff to access bibliographic information for many subject areas through one familiar system.

Order forms and pricing information for the Search CD450 System and its reference databases are available from participating OCLC network offices or from its Electronic Publishing and Information Delivery Division. Use these toll-free telephone numbers:

National: 1-800-848-5878
Ohio: 1-800-848-8286

■■

Book on Inmagic

Inmagic Inc., developer of textbase software for micro- and minicomputers, has produced a new version of "Biblio Guide: Using Inmagic in Libraries," a book of models for setting up Inmagic databases for online catalogs, serials, loans, and acquisitions. The guide gives Inmagic users a quick start with their library automation projects, and the package includes a disk containing all of the database models and more than ninety report designs illustrated in the guide. These examples can be used as they are or modified to suit particular needs. Numerous tips on library automation are also included. Elizabeth B. Eddison is editor of Biblio Guide.

Since 1980 Inmagic software has been installed in several thousand locations in more than forty countries worldwide. It is used by information managers in a variety of organizations, such as Fortune 500 corporations, management consulting firms, law offices, hospitals, government offices, as well as personnel, real estate, and design firms and special collections in school, university, and public libraries.

Inmagic products are distributed through specialty dealers, value-added resellers, and consultants, as well as from its Cambridge, Massachusetts, headquarters. The price of Biblio Guide is \$145, and the Inmagic microcomputer version sells for \$975. Multiple copy discounts, corporate licenses, and a network version are available. ■■

LITA Newsletter Available by Subscription

LITA Newsletter, a quarterly publication of the Library and Information Technology Association (a division of ALA), is now available by subscription. Subscriptions are \$15 per year in the United States, Canada, Spain, and other PUAS countries; \$25 in other foreign countries.

The newsletter complements LITA's journal, *Information Technology and Libraries*, with a mixture of news, reports, and opinions and provides coverage of conference programs and discussions within LITA, the only professional association devoted exclusively to the use of technology in libraries.

In addition to reports on the plans and activities of the association, *LITA Newsletter* includes columns on library technology such as authority control, expert systems, technical standards, online catalogs, and the MARC formats. The column on technical standards has served as a forum on standards issues since the newsletter's inception in 1980.

The first issue available by subscription is Winter 1988 (V.9, no.1). Orders may be sent to Subscriptions Department, American Library Association, 50 E. Huron St., Chicago, IL 60611-2729. ■■

Montgomery County Selects Eyring

Agnes M. Griffin, director of the Montgomery County Department of Public Libraries in Rockville, Maryland, has announced that the county has signed a \$1.8 million contract with Eyring Library Systems, Salt Lake City, for a new integrated library automation system.

Montgomery County, population 665,000, is located in the Washington, D.C., suburban area. The library system is one of the most active in the area, claims a high readership of 12.6 per capita, and currently serves 72 percent of the households. The new system will serve the county's twenty-three libraries (four regional, eleven community, eight branch) with a combined collection of more than two million items ranging from compact discs to best-sellers and government documents.

The system is designed to allow for the growth of the current libraries as well as the planned addition of four new libraries. After seven years, the combined collection is expected to contain approximately 3.5 million items.

In the first phase, the Eyring system will organize the circulation, bibliographic maintenance, cataloging, and management information for the entire county library organization. There will be 199 terminals for staff use.

In the second phase, through the computerized catalog, any of the country's estimated 480,000 patrons will be able to search for any book or other item in the collection of any of the county libraries.

In addition, anyone using the computerized catalog will be able to request community information on topics such as educational opportunities, club and organization activities, the County Council agenda, etc. Both the catalog and community information will be accessible twenty-four hours a day on dial-up lines available to anyone with a computer or terminal equipped with a telephone modem.

Montgomery County library officials anticipate more than seven million computerized catalog inquiries by staff during the first year of operation and a circulation of approximately nine million items. The system is designed to allow for expected growth through 1995 without becoming overloaded (by which time circulation of fourteen million items per year is expected, along with at least twenty-five million catalog inquiries).

The Eyring system uses software developed by the Colorado Alliance of Research Libraries, a consortium of university and public libraries, and is hosted on a Tandem TXP computer system. Part of the contract to deliver the system calls for it to be available 98 percent of the time. ■■

Faxon's FI\$CAL Software

The Faxon Company has announced the release of FI\$CAL, a micro-based software package that performs transfer of selected Faxon invoice data into PC applications for financial planning and analysis.

Designed to run on IBM PCs and PC-

compatibles, this new software transfers user-selected data from Faxon invoices into files readable by Lotus 1-2-3, Symphony, dBASE II, or dBASE III. Within those programs, the invoice data can be manipulated in many ways, depending on the requirements of the individual client. For example, titles and total serial expenditures can be sorted by various departments over designated time periods for purposes of comparison, reporting, and serial budgeting and planning. Any user who is familiar with Lotus or dBASE commands can use the imported files to meet a variety of specific needs.

Mary Ellen Clapper, manager, Library/Vendor Interfaces, said of the service: "FI\$CAL will help the increasing number of library professionals interested in customization, enhanced local control, and creative, flexible manipulation of serial invoice data. The automatic data transfer

saves time and eliminates manual keying error."

Clients can arrange to receive their current invoice information of PC diskette as an annual, quarterly, or monthly service. FI\$CAL requires an IBM PC or PC-compatible with 256K memory and two disk drives. ■■

SOLINET Signs Agreement with Blackwell for Authority Control

The Southeastern Library Network (SOLINET), incorporated in 1973, has announced the signing of an agreement with Blackwell North America, Inc., under which SOLINET will broker Blackwell's authority control services.

Blackwell services provide authority control for name, title, series, and subject headings. ■■

Now Available... Biological Abstracts/RRM® Collective Indexes, 1980-1984 on Microfiche!



Access five years of **Biological Abstracts/RRM (Reports, Reviews, Meetings)** (**BA/RRM**) references *quickly and easily* with the new **BA/RRM Collective Indexes** on microfiche!

Just look at what the **BA/RRM Collective Indexes** have to offer you!

- convenient and rapid access to references to over 700,000 research reports from Volumes 18-27 of **BA/RRM**
- access to research reports, reviews, U.S. patents and books in biology and biomedicine
- and, best of all, *unique* coverage of life science meetings, symposia and papers!

Don't waste time searching volumes of **BA/RRM page by page**—order the **BA/RRM Collective Indexes, 1980-1984**, on microfiche TODAY!

Contact: BIOSIS, Customer Services, 2100 Arch Street, Philadelphia, PA 19103-1399, USA. Or call toll free 1-800-523-4806 (USA except PA); (215) 587-4800 worldwide; Telex 831739; Fax (215) 587-2016.



BIOSIS®

Effective Technology, Excellent Service: Putting the Pieces Together



LITA National Conference • October 2-6, 1988 • Boston, MA

The Library and Information Technology Association (LITA) convenes its second national conference in Boston. The conference agenda covers all areas of information technology, with particular emphasis on library applications.

- Technology Seminars will provide substantive overviews of six topics including expert systems and artificial intelligence, optical technologies, microcomputers, integrated library information systems, facsimile and optical scanning technology, telecommunications
- Forty Technical Sessions on such topics as authority work in an online environment, administration of technology and implications for traditional services, preservation of library materials and disaster planning, optical publishing, relations between libraries and computer centers, the electronic library
- New Product Reviews by selected exhibitors providing in-depth demonstrations of newly introduced or enhanced products and services
- Showcases of real-life uses of new technologies presented by library and information science practitioners sharing their most recent developments and projects
- Exhibits—over 100 exhibitors and 150 exhibit booths displaying the latest in library technology
- Open House Visits to see automated library systems in many of the leading libraries in the Boston area
- Post-Conference Workshops on data conversion, the Apple Macintosh as a library workstation, design and use of online authority control systems, hacking IBM PC system hardware, preparing for serials automation, library systems facilities management

Choose the pieces you need to bring together technology and service — Come to Boston in October for LITA's 2nd National Conference.

**For more information: LITA National Conference '88
American Library Association
50 East Huron Street
Chicago, IL 60611-2729
(312) 944-6780**

Recent Publications

Book Reviews

Dewey, Patrick R. *101 Software Packages to Use in Your Library: Descriptions, Evaluations, and Practical Advice*. Chicago and London: American Library Assn., 1987. 160p. paper, \$17.95 (ISBN 0-8389-0455-6).

Walton, Robert A., and Nancy Taylor. *Directory of Microcomputer Software for Libraries*. Phoenix, Ariz.: Oryx, 1986. 564p. paper, \$37 (ISBN 0-89774-342-3).

Significantly different results can be achieved from the same premise, as evidenced by two new microcomputer software guides.

Walton and Taylor take a strict definition of library software for their *Directory of Microcomputer Software for Libraries*. Facts on 249 software packages specifically designed for library functions are presented in a clear, tabular format. A supplemental annotated listing contains 74 additional software titles that are no longer available, part of a larger package, or lacking full information.

The *Directory* includes an amazing amount of software, from the small in-house projects of individual institutions to software packages marketed by the largest library-market vendors. Too frequently, though, the minimal single-paragraph abstract does not provide a full understanding of the program's capabilities. Otherwise, the level of detail and clear presentation is impressive.

Two pages are allotted to each program in the *Directory*. The information on each is broken down into standard categories: abstract; vendor; software requirements (operating system, language); and hardware requirements. The categorization and uniform-page format facilitate comparison but often result in white space be-

tween the categories of succinctly explained software.

The alphabetical (by title) arrangement of the information complicates comparison, however. Though a comprehensive subject index is provided, along with a needless title index and vendor index, grouping the programs by function would have been a more convenient arrangement. A software citation index, presented in title order, lists citations from library and computer periodicals for many of the programs. Because the authors have ferreted out many obscure programs from non-profit developers that have escaped media attention, many of the programs lack any citations. References date from 1983 through early 1986.

The work also includes an excellent opening chapter containing thirty-one "criteria for selecting and evaluating microcomputer software."

Patrick Dewey's *101 Software Packages To Use In Your Library* is well described by its subtitle: *Descriptions, Evaluations, and Practical Advice*. Arranged by type of software, Dewey opens each chapter with a general explanation of how the programs in the genre function, the services and uses they can provide, and some selection guidelines. Tabular format descriptive reviews complete each chapter. Supplemental materials include a glossary of microcomputer terms and sections on good work habits and basic computer care.

The major difference between the books is that Dewey only includes software that he recommends, based on firsthand examination, while the aim of the *Directory* is to be comprehensive and not judgmental. A serious flaw in *101 Software Packages* is that there is no indication of whether an

omitted product is considered flawed or was only unknown or unavailable to the author. The lack of such a distinction limits the credibility of the work.

Though Dewey's book is smaller than Walton and Taylor's, and contains fewer programs in each genre, Dewey attempts to cover the entire range of software useful in a library, not just library software. Chapters on communications packages, integrated software, newsletter/desktop publishing, utilities and word processing are included. While these chapters provide a good introduction to the topic with capsule reviews, information on such programs is widely and regularly available in computer and general interest magazines, and the features and prices of the packages change often. The result is only about half the programs in Dewey book are purely library software.

The later publication date belongs to Dewey's work, but it is no more up-to-date, the citations dating from 1982 to 1985. And despite his inclusion of general and business-market software, Dewey's hardware orientation is behind the times, listing more applications for the Apple II than the IBM-PC line, and describing CP/M as "the most widely used operating system available for micros." Much of his word processing chapter is based on the results of his 1984 survey of library word processing use, which not surprisingly demonstrates a heavier use of Apple equipment than IBM-PCs.

101 Software Packages provides a valuable service to the novices of library automation, and microcomputing in general. Dewey presents clear, basic explanations of computer functions and digested software recommendations that will encourage those with unused Apples to join the revolution. Experienced microcomputing librarians may find the book limited and dated.

Librarians experienced in microcomputing and seeking new software packages will find *Directory of Microcomputer Software for Libraries* of more use for its large collection of strictly library-oriented software and clear distillation of basic facts.—*Ronald A. Gagnon, North Shore Community College Learning Resource Center, Beverly, Massachusetts.* ■■

European Conference on Library Automation, 11-12 September 1986, Harrogate, England. LA Conference Proceedings Series in Library Automation, no.5. London: Library Assn., (dist. by ALA Publishing Services), 1987. 84p. paper, \$15 (ISBN 0-85365-508-1). "Proceedings of the conference organized and sponsored jointly by the Library Association and CLSI."

This is the fifth publication in a little-known series for North Americans published by the Library Association. The proceedings and indeed this series are designed to assist information specialists to "understand and exploit information technology." The focus of the series is decidedly Eurocentric with particular emphasis on technological developments in the United Kingdom. Three years have elapsed since the last publication in this series, yet this conference and its proceedings are by far the most distinguished to date because of the scope and future historical value to librarians.

First, some background notes are needed to explain the series. The title of the earliest proceedings published in the series is at first glance somewhat misleading. *Information Technology in 1982* contains the published papers, as the subtitle goes on to explain, of the annual meeting of the STATUS Users Group, held at the University of Loughborough in late 1982. STATUS is information management system software developed in the midseventies, and it was the basis of a bibliographic database used mostly by technical libraries in Britain. The book, although somewhat parochial, is a good overview of what STATUS libraries were doing during Information Technology Year in 1982. The second proceedings to be published were also held that year. *Information Technology on Screen: New Approaches in Videodata, Teletext and Cable* has a title almost too long for its modest sixty-three pages. The papers on tele-software, cable, Laser Vision, teletext, and view data tended to be somewhat technical. The third offering in the series is *U.K. Library Database System and Union Catalogues*, the proceedings of a seminar by the Cataloguing and Indexing Group held in January 1983. The focus is naturally Brit-

ish, but in addition to the primary discussion of the U.K. Library Database System (UKLDS), other topics such as network development, interlibrary lending, and union catalogs were discussed. The fourth publication in the series, also published in 1983, is *Databases for Books: Their Uses for Selling, Acquiring and Cataloguing*. These scant fifty-eight pages of the proceedings of the MARC Users Group are of primary interest to catalogers and acquisitions librarians.

After a three-year hiatus comes perhaps the best offering in this series. The fifth in the LA Conference Proceedings Series in Library Automation is the *European Conference on Library Automation*. Held in late 1986, it simply attempts to outline developments in library automation in Europe with special focus on the United Kingdom, France, Netherlands, Scandinavia, and Southern Europe. The issue of European cooperation is covered in two final chapters. There are huge intentional gaps in national coverage in this conference, and because there are we really don't see the big picture. These are, after all, the proceedings of a conference, not a monograph. Several of the papers were poor, and many of them contained the general conference chatter one might expect. Yet in spite of these occasional blemishes, this book is of interest to American librarians and information specialists.

The chapter "Automation in France" by Pierre Le Loarer is excellent and gives one an appreciation of the pace in a highly centralized, bureaucratic environment. Kari Marklund's chapter on "Library Automation in Scandinavia" shows the diversity of development especially when there are many in-house developed systems available, most of which are unknown to Americans. Clearly in Denmark, Finland, Sweden, and Norway the role of the National Library is very strong. This centralization is a theme repeated by many of the authors as is the divergent development between public and academic libraries. John Eyre's contribution on "Library Automation in Southern Europe" is also interesting because the efforts are so different from the rest of Europe. Spain, Italy, Greece, Portugal, and Yugoslavia are the countries most

frequently mentioned in this clustering of Mediterranean nations. Eyre tends to generalize, but only to explain the slower pace of development, hardware problems, reinvention of the wheel (writing library software from scratch), traditional attitudes toward librarianship as a profession, and funding. The second of the two articles on European cooperation, by Christian Lupovici, is quite substantial. Lupovici discusses LIB-2, one of three major studies being done for the European Commission. Each of the twelve countries will provide factual data on the state of library technology in their respective nations. Each survey has four distinct parts: computerized catalogs, computer applications in library management, computer-based user services, and an interlibrary loan component. Two short-term goals of the study, besides addressing the "disparities" between the northern and southern countries, would be to explore areas of cooperation in interlibrary loan and access to foreign catalogs now made difficult because so much library software and hardware are different. From a European perspective, technical problems are the least of their worries. Other dilemmas that beg resolution are creating a single ILL network with common rules, dealing with multilingualism, and multiple cataloguing formats: INTERMARC (used in French-speaking nations), MAB (used in West Germany), and UNIMARC.

Not surprisingly, these proceedings mention in passing various technical aspects of librarianship: circulation subsystems, union catalogs, library software and hardware (microcomputers, minis, and mainframes), OPAC's, CD-ROM, and bibliographic utilities, to name a few. Although individual national treatment is spotty, some strong contributions salvage this book. Those interested in the international aspects of librarianship and the development of technology will find that these proceedings nicely complement information found in IFLA and UNESCO publications. This book is therefore recommended for large academic libraries.—Tom Smith, *Library of Congress, Washington, D.C.* ■■

Mandel, Carol A. *Multiple Thesauri in On-*

line Library Bibliographic Systems: A Report Prepared for Library of Congress Processing Services. Washington, D.C.: Cataloging Distribution Service, LC, 1987. 94p. paper, \$15.

Applying computer support to multiple controlled vocabularies in bibliographic systems accessed by librarians and library patrons is the subject of a report written by Carol A. Mandel, director of the Technical Services Group for Columbia University Libraries, under contract for the Processing Services Department of the Library of Congress. The report's objectives are as follows: (1) to survey the state of the art in computer support of multiple controlled vocabularies, including such support at the Library of Congress, (2) to describe the current multiple controlled vocabulary environment at the Library of Congress, and (3) to provide the Library of Congress with various approaches for improving its current computer support of multiple controlled vocabularies.

The report consists of five chapters and an executive summary. Chapter 1 compares the multiple controlled vocabulary environment of commercial database searching with that of library catalogs and concludes that the situation is more problematic for libraries because their systems are intended for direct use by patrons. Few online library systems have been designed to support one controlled vocabulary; none supports *more than one* controlled vocabulary. Thus, the burden of searching in a multiple controlled vocabulary environment falls upon library patrons.

Chapters 2, 3, and 4 describe the state of the art in computer support of multiple controlled vocabularies. Such support is divided into three broad functional areas: (1) thesaurus management, (2) subject authority control, and (3) subject searching.

Eleven appropriate features for a thesaurus management system are listed in chapter 2. These features were generated from a review of existing thesaurus management systems; however, no existing system supports all eleven features. The features are suitable for managing one or more thesauri as long as each thesaurus is managed separately. When integration among several

thesauri is desired, additional computer support is necessary. Mandel summarizes five approaches to thesaurus integration from a Unesco report prepared by F.W. Lancaster and Linda C. Smith: (1) mapping, (2) intermediate lexicon/switching language, (3) integrated vocabulary, (4) microthesauri, and (5) macrovocabularies. She enumerates additional features that a thesaurus management system should have to support an integrated approach to thesaurus maintenance.

In chapter 3, Mandel describes subject authority control capabilities of seven online bibliographic systems. Each system's subject authority file, links and references, and retrieval and display features are described; numerous figures are included. Mandel concludes, however, that while the state-of-the-art in subject authority control is well developed, no system "makes subject searching through multiple authority files accessible to the library patron."

In chapter 4, Mandel asserts that online catalogs have always involved manipulation of multiple vocabularies, i.e., patrons' terms and the catalog's controlled vocabulary. The ability of a system to match terms entered by patrons with the catalog's controlled vocabulary is affected by the design of the system's subject searching features. In a multiple controlled vocabulary environment, retrieval is also affected by the system's approach to providing library patrons with access to databases indexed by different controlled vocabularies. Four approaches are described: (1) segregated files, (2) mixed vocabularies, (3) integrated vocabularies, and (4) front-end navigation.

In chapter 5, Mandel describes six controlled vocabularies used at LC: (1) *Library of Congress Subject Headings (LCSH)*; (2) *Legislative Indexing Vocabulary (LIV)*; (3) *Thesaurus of Graphic Materials (TGM)*; (4) *Subject Headings for Children's Literature*; (5) *National Union Catalog of Manuscript Collections (NUCMC)*; and (6) *Handbook of Latin American Studies (HLAS)*. The author describes how LC's present systems provide computer support for various aspects of thesaurus management, subject authority control, and subject searching, and sug-

gests what directions LC can take to enhance existing systems or design new ones.

Although the report was prepared for LC and emphasizes its computer systems and multiple vocabularies, it is essential reading for libraries whose card and/or on-line bibliographic files contain more than one controlled vocabulary. Specialized thesauri are common for certain subjects (medicine, art); formats (microforms, pictures, 35 mm slides); referral files (local clubs, events); and audiences (children).

Libraries have focused on the construction of bibliographic and authority files for monographs, serials, and sound recordings, and other materials whose subject matter is represented by *LCSH*. Once such files are constructed, libraries' attention will shift to other files and collections whose subject matter is represented by specialized thesauri. If the thesauri are created and maintained by the library itself, computer support for thesaurus management will be necessary. If the thesauri are created by an external agency, computer support for authority control and subject searching will be desirable functions. Furthermore, libraries will have to resolve many of the same questions now facing the Library of Congress. Can separate parallel thesaurus management systems be employed for each thesaurus or is some degree of integration necessary? Should subject searching on different thesauri be integrated? What approach to multiple-vocabulary searching is easiest for library patrons? When choosing an integrated library system or a replacement system, libraries will have to assess the merits of such systems regarding computer support for multiple vocabularies. Mandel's report is highly recommended for library staff, educators, or online library bibliographic system designers interested in online catalogs, subject searching, subject authority control, and thesaurus construction and maintenance. The report is well organized and clearly written. Thus, readers can easily distinguish between material specific to LC and material that applies to libraries in general.—*Karen Markey, School of Information and Library Studies, University of Michigan, Ann Arbor.* ■■

Software Reviews

Catalog Carder. Right-On Programs, 1737-36 Veteran's Highway, Central Islip, NY 11722. Hardware requirements: IBM-PC and compatibles and Apple Series. Price: \$95.

Librarian's Helper. Scarecrow Press, 52 Liberty St./P.O. Box 4167, Metuchen, NJ 08840. Hardware requirements: For PC-DOS, MS-DOS, or CP/M-86, 256K RAM and at least one disk drive. For CP/M (including Apple with CP/M card), 64K RAM and two disk drives. Price: IBM or compatible, \$225; Apple with CP/M card included, \$325.

Librarian's Helper and *Catalog Carder* are among several types of software being promoted to speed the process and eliminate cataloging drudgery in schools and small libraries. Both claim to conform to exact AACR2 standards. Producers also promise that the menu-driven programs are easy to use by people with minimum knowledge of cataloging.

Though I applied the two systems on numerous books, I was also anxious to see if the programs would be useful in our Special Collections Library at the University of Arizona. We do not have to catalog our books but we do process hundreds of pamphlets, broadsides, manuscripts, and the University archives. In addition, we often need to create replacement cards for our books and other items. Until now, we have had to go through the onerous process of typing and xeroxing the card, then adding numerous tracings. Obviously this consumes an enormous amount of time, especially since we produce two sets of cards (one is interfiled in our Main Library catalog which strictly follows AACR2 convention).

Helper couldn't quite do it all, but it came darn close. Given a few adjustments, the program could be an invaluable addition to our library software. *Carder* turns out to be a poor cousin indeed and is no bargain at \$95.

Carder is copy-protected but can be copied to a hard disk. Right-On Programs does guarantee a quick replacement copy for \$5 if the program is damaged. It also needs a boot from DOS but once done, this prod-

uct's on-screen prompts guide the user through a simple program of eight fixed fields. These include: catalog number; author or editor; illustrator; title; publication information; copyright date; book description; note field; and allowance for four subject headings. The computer beeps at the end of each field and this could be a nice feature. Unfortunately, it more often than not beeped before the entry was complete. This probably was due to the fact that the program was designed to use 16.5 type and limit each cataloged item to one card.

As a result, these overly short fields do not allow adequate space for longer entries. For example, a hundred 16.5 pt. characters are allotted for title. My sample entry (figure 1) simply truncated. In addition, 44 spaces are not enough to include publication information. Nor could I squeeze my sample subjects into 30 spaces. Tracings allow four subjects and whether you want them or not, there are always two added entries which appear as follows: "I. Author II. Title."

The program neither accepts title main entry nor prints out with hanging indentation. I tried it and got "I. Title I. Title." Forget trying to follow convention and include an author's birthdate; the statement of responsibility comes out like this: "/ by Lucy Elizabeth, 1953 Harding."

Carder promises "user friendliness" and strict adherence to AACR2 rules. As it turns out, the user does need to be familiar with AACR2 because very little punctuation is included; there is none at all in the physical description area. If you leave out a field—publication, for example—you get punctuation anyway (figure 1). Additionally, if

you want capitalized subject headings added to cards, you must type them in caps in the tracings area. One needs to experiment to find out what is and what is not included here.

According to documentation, data for ten card sets can be entered at a sitting. Designers admit that once a set is printed, all data is deleted and cannot be reprinted without being typed in again. Worse still, as I found to my remorse, all my data were lost when I used the "align printer" function. Moreover, in order to print cards the user must go through a cumbersome process of entering the call number for each entry and printing before calling up the next one. If a number has been incorrectly entered there is simply no way to call the card to print. By the way, three lines are allowed for call number; too few for some.

Conversely, it was pure delight to experiment with *Librarian's Helper*, which passed most tests quite nicely. *Helper* is a sophisticated program accompanied by excellent documentation from Jennifer Pritchett and Fred Hill. It also supplies cross references to Akers' *Simple Library Cataloging*. It did not take long to get started, enter data, then sit back and watch the printer spin out dozens of AACR2-perfect cards containing as many as ten subject headings that automatically capitalize when needed, ten added entries (quite a few more tracings than *Carder* and others I have seen), four generous note fields, and as many "continued on next card" as needed. For title added entry the system allows users to type the title as it should appear (figure 1). The multi-pitch capability (10-16.5 characters per inch) is a helpful feature and

Petrology and tectonic setting of the Livingston Hills ...	
E9791	
1978	Harding, Lucy Elizabeth, 1953-
213	Petrology and tectonic setting of the Livingston Hills formation, Yuma County, Arizona / by Lucy Elizabeth Harding, 1978. viii, 57 leaves : ill. (some col.) maps (1 fold in pocket) ; 28 cm.
	Thesis (M.S. - Geosciences)--University of Arizona. Includes bibliography.
	1. Petrology--Arizona--Livingston Hills. 2. Petrology--Arizona--Ploosa Mountain. 3. Rocks, Sedimentary. 4. Geology--Arizona--Yuma County. I. Title: Petrology and tectonic setting of the Livingston Hills ...

Petrology and tectonic setting of the Livingston Hills Forax	
E9791	Harding, Lucy Elizabeth, 1953
1978	Petrology and tectonic setting of the Livingston Hills Formation, Yuma County, Arizona / by Lucy Elizabeth, 1953 Harding. -- , 1978. viii, 57 leaves : ill. (some col.) maps (1 fold in pocket) ; 28 cm.
	Thesis (M.S. - Geosciences)--University of Arizona. Includes bibliography
	1. Petrology--Arizona--Livingston. 2. Petrology--Arizona--Ploosa No. 3. Rocks, Sedimentary. 4. Geology--Arizona--Yuma County. I. Author II. Title

Fig. 1. Librarian's Helper and Catalog Carder

I used it with cards requiring lengthy notes and subjects. This is where *Helper* outshines *Carder* by using variable-length fields within a fixed record length of 3,000 characters. This feature is probably more than adequate for most users. OCLC's limit, for example, is 5,000 characters.

For those of us who conform slavishly to AACR2, there are a couple of problems that the designers should be able to adjust. First, the call number should repeat on titles that require more than one card. (There are a generous five lines for call number and a sixth for volume, if desired.) The second problem also appears on "continued" cards. In this case, the period is dropped after an author's name (figure 2).

There are thirty-two fields, and the prompt for seldom-used fields can be toggled on or off as desired. Some of these include: formatting title main entry cards with hanging indentations; a series note; LC and ISBN numbers; parallel and uniform titles; and subsequent statements of responsibility. This latter field could use a slightly more space for longer titles, however.

The literature promises that "the program assumes no knowledge of computers and minimal knowledge of cataloging."

This is only partly true. Users must learn what to do with the prompts, especially if they have typed in errors, because the editing process can be cumbersome at first. It requires the cataloger to return to the Main Menu for each prompt in the area(s) where a mistake was entered—better to have a good typist. Most punctuation is supplied by the program, however, and those who do data entry can work from a User Input Form, which the designers have provided in the documentation.

Helper has other useful features. Unlike *Carder*, it can save data which can be recalled easily by record number. Moreover, it will sort this data by author, title, and/or call number and print out bibliographic lists if required. I set up separate diskettes for the different formats so I could get one printout for broadsides, or one for pamphlets, etc. Multiple sets of labels are automatically created with sequential numbers for multi-volume sets during data entry. Additionally, the user can use the import/export feature of a database manager to move data in and out of the program format and the designers have provided a description of the data structure to assist this process.

Az 326	Fred Harvey Company. Papers, 1896-1945 / /by Fred Harvey Company. 17 boxes and 6 v.
	Approx. 2000 photographs of hotels, eating houses, and news stands operated by Fred Harvey. Also includes photos of railroad stations, Indian exhibits at the San Francisco, Calif., Panama-Pacific International Exposition, 1915, and the San Diego, Calif., Panama-California Exposition (1915-1916), and personal pictures of Southwestern Indians, scenes in Mexico, and individuals.
	Correspondence consists of letters of commendation for services rendered to patrons of the dining cars, eating establishments, and (continued on next card)

Az 326	Fred Harvey Company. Papers, 1896-1945. 17 boxes and 6 v.
	Approx. 2000 photographs of hotels, eating houses, and news stands operated by Fred Harvey. Also includes photos of railroad stations, Indian exhibits at the San Francisco, Calif., Panama-Pacific International Exposition, 1915, and the San Diego, Calif., Panama-California Exposition (1915-1916), and personal pictures of Southwestern Indians, scenes in Mexico, and individuals.
	Correspondence consists of letters of commendation for services rendered to patrons of the dining cars, eating establishments, and (continued on next card)

	(Card 2)
Fred Harvey Company	
hotels operated by the company.	
There are two guest registers (1922-1930) for Herall's Rest and Herall Creek Camp, both at the Grand Canyon; ledger book (1906-1933) for various facilities and operations at the Grand Canyon; an album of blueprint plans for news stands; and additional miscellaneous items.	
Inventory with the collection.	
1. Title: Fred Harvey Company Papers.	

	(Card 2)
hotels operated by the company.	
There are two guest registers (1922-1930) for Herall's Rest and Herall Creek Camp, both at the Grand Canyon; ledger book (1906-1933) for various facilities and operations at the Grand Canyon; an album of blueprint plans for news stands; and additional miscellaneous items.	
Inventory with the collection.	
1. Title: Fred Harvey Company Papers.	

Fig. 2. Librarian's Helper

1987, Amsterdam [etc.]: Elsevier, 1987. 437p. \$69.50 (ISBN 0-444-70302-0). "Published on behalf of the American Society for Information Science."

Automated Systems for Access to Multilingual and Multiscript Library Materials—Problems and Solutions. Ed. by Christine Bossmeyer and Stephen W. Massil. IFLA-Publications 38. Munich [etc.]: Saur, 1987. 225p. \$34 (ISBN 3-598-21768-4). "Papers from the Pre-Conference held at Nihon Daigaku Kaikan Tokyo, Japan, August 21–22, 1986." "Edited for the Section on Library Services to Multicultural Populations and the Section on Information Technology."

Boaz, Martha. *Librarian/Library Educator: An Autobiography and Planning for the Future.* Metuchen, N.J. and London: Scarecrow, 1987. 314p. \$27.50 (ISBN 0-8108-1988-0).

Boettcher, Cheryl. *The Kimono Imagined.* University of Illinois Graduate School of Library and Information Science, Occasional Papers, no. 180. Champaign: Univ. of Illinois, 1987. 38p. paper, \$3.50 prepaid (ISSN 0276-1769).

Boss, Richard W. *Information Technologies and Space Planning for Libraries and Information Centers.* Boston: Hall, 1987. 121p. \$36.50 (ISBN 0-8161-1859-0); paper, \$28.50 (ISBN 0-8161-1870-0).

CD-ROMs in Print, 1987. Comp. by Nancy Melin Nelson. Westport, Conn. and London: Meckler, 1987. 102p. paper, \$29.95 (ISBN 0-88736-179-X).

Combining Libraries: The Canadian and Australian Experience. Ed. by L.J. Amey. Dalhousie University School of Library and Information Studies, no. 2. Metuchen, N.J. and London:

Scarecrow, 1987. 433p. \$39.50 (ISBN 0-8108-2049-8).

The Library Microcomputer Environment: Management Issues. Ed. by Sheila S. Intner and Jane Anne Hannigan. Phoenix and New York: Oryx, 1988. 258p. paper, \$27.50 (ISBN 0-89774-229-X).

Nonbook Media: Collection Management and User Services. Ed. by John W. Ellison and Patricia Ann Cody. Chicago and London: American Library Assn., 1987. 388p. paper, \$35 (ISBN 0-8389-0479-3).

Online Database Search Services Directory: A Reference and Referral Guide to More than 1,700 Libraries, Information Firms and Other Sources Providing Computerized Information Retrieval and Associated Services Using Publicly Available Online Databases, 2d ed. Ed. by Doris Morris Maxfield. Detroit: Gale, 1988. 1,268p. \$155 (ISBN 0-8103-2114-9).

Plumbe, Wilfred J. *Tropical Librarianship.* Metuchen, N.J. and London: Scarecrow, 1987. 318p. \$29.50 (ISBN 0-8108-2057-9).

Saffady, William. *Integrated Software Packages for Microcomputers.* Westport, Conn. and London: Meckler, 1987. 238p. \$29.95 (ISBN 0-88736-172-2).

A Track to Unknown Water: Proceedings of the Second Pacific Rim Conference on Children's Literature. Ed. by Stella Lees. Metuchen, N.J., and London: Scarecrow, 1987. 406p. \$32.50 (ISBN 0-8108-1199-5).

Words on Tape: An International Guide to the Audio Cassette Market 1987/1988. Westport, Conn. and London: Meckler, 1988. 427p. \$29.95 (ISBN 0-88736-222-2). ■ ■

Now available by subscription!

LITA newsletter

- Contemporary columns on technology in libraries, including authority control, expert systems, technical standards, online catalogs, the MARC formats and more....
- Focussed, timely coverage of conference programs and discussions within LITA, the only professional association devoted exclusively to the use of technology in libraries....
- Current reports on the plans and activities of the association....
- All in an informed, informal mixture of news, reports and opinions.

Now available to non-members by subscription.

Library and Information Technology Association
a division of the American Library Association

Order from: Subscriptions Department OR PHONE:
American Library Association 312-944-6780
50 East Huron Street
Chicago, IL 60611-2729

Please enter a one-year subscription to the LITA Newsletter for me and send the next four issues for \$15.00.

MAIL TO:

Name

Address

City/State/Zip

BILL TO:

Name

Address

City/State/Zip

Payment is enclosed Please bill me.

Allow 8 weeks for first issue. Add \$10 outside the U.S.

Letters

To the Editor:

Since the publication of "A Technical Evaluation of the Linked Systems Project Protocols in the Name Authority Distribution Application" in the December 1987 issue, OCLC has made further progress with implementation of the LSP protocols. Unfortunately, insufficient experience has been gained to enable technical evaluation of these additional implementations at this time; however, I thought a current update of status might be of interest to your readers.

Distribution of LC Name Authority records from LC to OCLC has continued since late January 1987, without major problems. OCLC has successfully completed testing of the contribution process with the library, and name authority records have been successfully contributed to LC by OCLC via the OCLC contribution facility. OCLC has also completed development of the necessary facilities to support inter-site searching of the OCLC database by LC, and the library has successfully searched the OCLC database on several occasions. Final adjustments to the OCLC user interface for record creation and editing are in process. As soon as these changes are integrated into the OCLC system, user documentation can be completed. OCLC is currently reindexing its name authority file to incorporate certain normalization conventions into OCLC's searching capability and upon completion of this process, near the first of the year, will begin training Indiana University staff as the first OCLC LSP user.—*Larry L. Learn, Director, Office of Telecommunications Planning, OCLC, Dublin, Ohio.* ■■

To the Editor:

Ben Chitty's article "Indexing for the Online Catalog" in the December 1987 issue is a thoughtful and well-developed analysis. Unfortunately, however, it con-

tains several errors of fact with regard to the TOMUS system distributed by Carlyle Systems.

1. On page 302, both the text and table 4 indicate that TOMUS uses keyword indexes only. In fact, TOMUS provides both heading and keyword indexes for any field, and most TOMUS libraries have heading indexes as well as keyword indexes.

2. The same text and table say that browsing is not available on TOMUS; in fact, browsing is available.

3. The table indicates that subfields are not indexed in TOMUS; in fact, they are indexed in virtually all TOMUS systems.

4. The table places a "no" opposite "indicators"; in fact, TOMUS recognizes a variety of indicators in various fields. These may include (on request) those mentioned in footnote 3 (source indicators in 6xx fields).

The table should thus have a "no" only opposite "permuted" and a "yes" in the rest of the column under TOMUS.

In the interest of accuracy, we would appreciate your printing this correction.—*Stephen R. Salmon, Chairman, Carlyle Systems Inc., Emeryville, California.* ■■

To the Editor:

As a long-time observer and participant in the development of library automation, I found A. B. Chitty's comparisons of heading, keyword, and permuted approaches to online catalog indexing ("Indexing for the Online Catalog" *ITAL* Dec 87) quite interesting and provocative. His analysis did not take into account one alternative, however.

In the ALOHA system, marketed by Advanced Libraries & Information, Inc. (ALII), headings are typically keyword-indexed. (Like a number of integrated library systems, index building is table-driven). Keyword searches which match more than one heading (single hits go directly to a detailed display) produce a

dynamically-created list of headings and cross-references (See and See Also), which can be browsed. In Chitty's terms, this perhaps would be described as a hybrid keyword/heading approach; however, it provides the contextual functionality which he attributes solely to permuted indexes. In one sense, it could be described as a permuted approach, except that in ALOHA the permuted display is created on demand, or post-coordinated, rather than created ahead of time, or pre-coordinated.

Chitty's assertion of the functionality of presenting keywords in context when there are multiple hits is, I believe, quite percep-

tive and accurate. My argument lies only in pointing out a variant that his analysis did not consider.

Finally, I would most emphatically agree with the thrust of Chitty's argument that more attention needs to be paid to the underlying principles of indexing when analyzing online catalog needs. Many online catalogs are rather like the proverbial dog who plays the piano: they are certainly online, but they really don't play very well.—
Douglas C. Livsey, Vice-President for Sales and Marketing, Advanced Libraries & Information, Inc., Salt Lake City, Utah. ■■

INDEX TO ADVERTISERS

ALA	pages 90, 100, 104
Biological Abstracts	page 89
Blackwell North America	2d cover
Data Trek	page 6
EBSCO	page 84
Midwest	4th cover
OCLC	3d cover
H. W. Wilson	page 5

INSTRUCTIONS TO AUTHORS

Information Technology and Libraries welcomes manuscripts related to all aspects of library and information technology. Some specific topics of interest are mentioned on the masthead page. Feature articles, communications, letters to the editor, and news items are all considered for inclusion in the journal. Feature articles are refereed; other items generally are not. All material is edited as necessary for clarity and length.

Manuscripts must be typewritten and the original submitted with one duplicate. Do not use onion skin. All text must be double-spaced, *including footnotes and references*. Manuscripts should conform to *The Chicago Manual of Style*, 13th ed., rev. (Chicago: Univ. of Chicago Pr., 1982). Illustrations should be prepared carefully as camera-ready copy, neatly drawn in a professional manner on separate sheets of paper. Manuscript pages, bibliographic references, tables, and figures should all be numbered consecutively.

Feature Articles consist of original research, state-of-the-art reviews, or comprehensive and in-depth analyses. An abstract of one hundred words or less should accompany the article on a separate sheet. Headings should be used to identify major sections. Authors are encouraged to relate their work to other research in the field and to the larger context of economic, organizational, or management issues surrounding the development, implementation, and use of particular technologies.

Communications consist of brief research reports, technical findings, and application notes. An abstract need not be included.

Letters to the Editor may offer corrections, clarifications, and additions to previously published material, or may be independent expressions of opinion or fact related to current matters of concern in the interest area of the journal. A letter commenting on an article in the journal is shared with the author, and a response from the author may appear with the letter.

News and Announcement items may announce publications, conferences, meetings, products, services, or other items of note.

Book Reviews are assigned by the book review editor. Readers wishing to review books for the journal are invited to contact the book review editor, indicating their special areas of interest and expertise.

Names and addresses of the journal editors may be found in paragraph three on the masthead page. In all correspondence please include your own name, institutional affiliation, mailing address, and phone number.



FIELD-TESTED

Policies and Procedures

College Library Information
Packets (CLIP Notes)

Designed to collect basic data and sample documents which college and small university libraries can use to establish or refine specific services or operations.

- ✓ Managing Student Workers in
College Libraries CLIP Note #7

■ *Compiled by Michael D. Kathman
and Jane McGurn Kathman*

Contains job descriptions, application and interview forms, orientation and training information, quizzes and tests, and evaluation documents used with student workers in college and small university libraries.

\$17.00pbk.; ACRL member \$14.00 182p. 0-8389-7097-4 1986

- ✓ Special Collections in College
Libraries CLIP Note #6

■ *Compiled by Christine Erdmann*

Sample documents from college libraries on projects, publicity, financial support, archives, preservation, and use policies.

\$18.00pbk.; ACRL member \$15.00 95p. 0-8389-7004-4 1986

- ✓ Mission Statements for College
Libraries CLIP Note #5

■ *Compiled by Larry Hardesty, Jamie Hastreiter,
and David Henderson*

Provides guidance to college libraries seeking to develop or refine their statement of purpose. Contains philosophical statements outlining the purposes of selected college libraries.

\$20.00pbk.; ACRL member \$15.00 107p. 0-8389-6944-5 1985

- ✓ Online Bibliographic Database
Searching in College Libraries
CLIP Note #4

■ *Compiled by David Carlson and P. Grady Morein*

"...making our college and university students cognizant of online services is a must. To those charged with making such services available on the nation's campuses this inexpensive guide is likewise a must." (*Online*)

\$19.00pbk.; ACRL member \$15.00 132p. 0-8389-6624-1 1983

Association of College and Research Libraries
a division of the American Library Association
c/o ALA Publishing Services, Order Department
50 East Huron Street · Chicago, Illinois 60611-2795

We stand behind every SC350 system

A serials control system should offer more than technical capabilities. That's why OCLC supports every SC350 installation with a team of specialists, ensuring that you will get the most out of your flexible serials control system.

SC350 has the features you want in a microcomputer-based serials control system. Multi-user capability, so staff members at different workstations can share one database • Easy check-in • Efficient methods for filing and tracking claims for missing issues • Ability to generate routing slips • Notification of what needs to be bound and what's at the bindery • Fund accounting and management reports to help you keep track of how your money is being spent.

Plus the ability to link to the Online Union Catalog and the LS/2000 local library system.

But it's the people who really make the difference.

To discover why more than 100 libraries have already chosen SC350, call us toll-free.

(800) 848-5878
Ext. 6403

(800) 848-8286 (Ohio)
Ext. 6403

**OC
LC** ||| **Local
Systems**



MATSS—The Perfect Software System For Automated Acquisitions

As a long time reliable source for the supply of books and cataloging services, Midwest Library Service has developed MATSS (Midwest Automated Technical Services System)—a fully integrated software acquisitions system for libraries.

The software package consists of the program diskettes and systems manual, which is easy to teach and easier to learn.

Features of MATSS are as follows:

- Open order database.
- Automatic checking for order duplication.
- Supports downloaded MARC records from major bibliographic utilities and CD-ROM databases.
- Orders are produced in printed formats or a BISAC variable format for electronic transmission to vendors.
- Orders can be transmitted toll-free to Midwest Library Service.
- Vendor and fund accounting.
- Orders transmitted to Midwest Library Service can be received automatically with little, if any, keying.
- Automatic claim/cancellation process.
- Management reporting.
- All functions are menu-driven.
- Catalog card and label production.
- On-site installation and training.



Midwest Library Service

11443 St. Charles Rock Road
Bridgeton, MO 63044, USA

Call toll-free 1-800-325-8833

Missouri librarians call toll-free 1-800-392-5024

Canadian librarians call toll-free 1-800-527-1659