Volume 6
Number 2

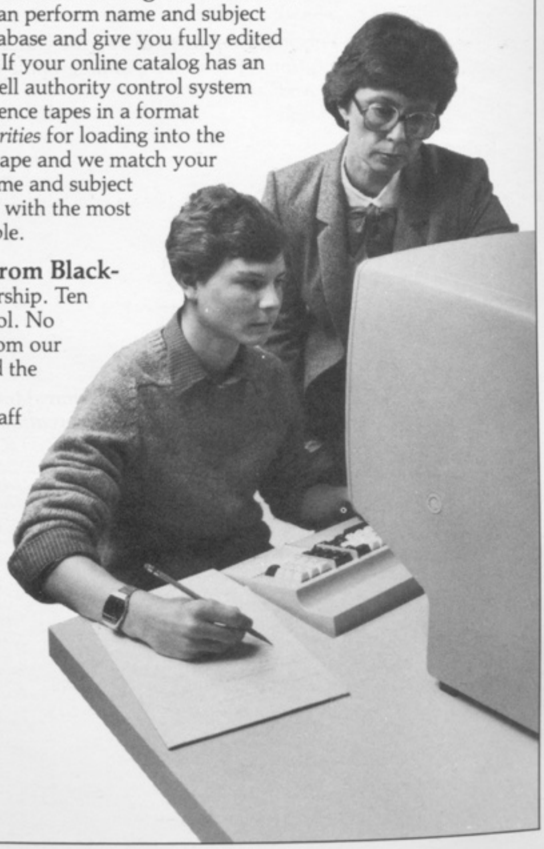
ISSN 0730-9295

Information Technology and Libraries

June 1987

CONTENTS

ITLBDC 6(2) 81–168 (1987); Inf Technol Libr



Information Technology and Libraries

Volume 6, Number 2: June 1987

CONTENTS

ITLBDC 6(2) 81–168 (1987); Inf Technol Libr

Information Technology and Libraries

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 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 80 of the March 1987 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, Arizona 85287. Copies of materials 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. Advertising arrangements should be made with William Z. Schenck, University of Oregon Library, Eugene, OR 97403.

Information Technology and Libraries is a perquisite 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. 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, 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, Otdyelnyi 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. ♾

The Linked Systems Protocol and the Future of Bibliographic Networks and Systems

Michael K. Buckland and Clifford A. Lynch

In this paper, we will describe the Linked Systems Protocol and examine the implications of its use for library automation and in particular for the future of bibliography, bibliographic systems, and bibliographic networks.

As background, we should first distinguish between the Linked Systems *Protocol* and the Linked Systems *Project*.

- The Linked Systems *Protocol* is a draft national standard for bibliographic information retrieval being developed by Committee D of the National Information Standards Organization (NISO), formerly American National Standards Institute (ANSI) Committee Z39, Subcommittee D. The current formal working title for the Linked Systems Protocol is *Z39.50: American National Standard—Information Retrieval Service Definition and Protocol Specifications for Library Applications.* The protocol is often referred to as Z39.50 to avoid confusion with the Linked Systems Project, since both the protocol and the project share the acronym LSP.
- The Linked Systems *Project* is a development project undertaken initially by the Library of Congress, the Research Libraries Information Network, and the Western Library Network, with funding from the Council on Library Resources. The participants are seeking to link their computer systems to exchange MARC authority records [as part of the Name Authority Cooperative Organization (NACO) project] and MARC bibliographic records.

The distinction between the protocol and the project is important. The project participants are contributing to the development of the protocol through their participation in the standards committee. As a practical matter, project participants defined an initial working draft of the Linked Systems Protocol, along with various other actual standards and working versions of standards that were not yet standardized, in order to establish a full "working" suite of networking protocols. They are now proceeding with the considerable undertaking of developing a workable implementation. In the meantime, the protocol has continued to evolve through successive drafts under the aegis of NISO Committee D.[1]

This paper does not discuss the project, which has been described elsewhere.[2,3] Nor does the paper explore technical issues relating to the protocol, except as they affect its use in applications. (See reference 4 for a detailed technical discussion; however, one should remember that the protocol is still a

---

Michael K. Buckland is Assistant Vice-President for Library Plans and Policies and Clifford A. Lynch is Assistant Director for Computing and Telecommunications, Division of Library Automation, University of California System.

This is a revised and expanded version of a presentation given at the LITA Linked Systems Project Program at the American Library Association meeting in New York, June 29, 1986.

The authors would like to thank Mary Jean Moore for her editorial assistance.

draft and somewhat volatile: it has undergone two major revisions since Lynch's article was written.)

PROTOCOL STANDARDS AND THE ISO OPEN SYSTEMS INTERCONNECTION REFERENCE MODEL

The Linked Systems Protocol is only one element defined within a sweeping and complex networking environment called the Open Systems Interconnection (OSI) protocol suite. The OSI protocol suite, in turn, is a concrete manifestation of an abstract conceptual model for computer networking. For the reader to understand fully the potential role of the Linked Systems Protocol, we must first review this context.

The OSI reference model is a conceptual framework developed by the International Standards Organization (ISO) for considering the various elements involved when two computers communicate with each other. These elements range from the electrical interface used to connect wires to the application programs that are specific to the particular task at hand. (The reference model does not define the communicating entities themselves, but only the interaction between them.) This conceptual model divides these activities into seven layers; roughly speaking, layers one through four are concerned with networking and telecommunications, and layers five through seven with computing and computer applications. (For an introduction to the OSI reference model, see reference 5; for a discussion of different protocol reference models, including the OSI model, see reference 6; and for a discussion of networking standards, see reference 7.)

The OSI reference model is itself a standard but is useful mainly as an intellectual tool that provides standardized terminology and concepts for discussing protocol definitions and actual systems. After this reference model was defined and accepted, ISO undertook to define actual protocols for each layer of the model. For virtually all layers, several alternative protocols were needed. At the lower layers, alternative protocols were needed to suit different kinds of network media and technologies; at the higher layers, different application protocols were needed for different applications (e.g., file transfer, banking transactions, library automation, etc.). For the middle layers—layers four (transport) and five (session)—only one protocol, or at most a few, will be defined at each layer, at least for the time being.

As definitional work progressed, it became clear that there were numerous oversights and missing details in the OSI model. Thus, fine structure was defined within individual layers—for example, three sublayers were defined for the network layer (layer three), and the application layer (layer seven) began to incorporate management and common application functions. In addition, missing components, such as network management and security, were defined.

Progress on this ambitious agenda of tasks has varied widely: some parts are defined and are accepted national or international standards, other parts are at various stages within the standards adoption process, and some are still in the definitional phase. Because so much of the OSI protocol suite is a unity and because the suite is not yet complete (in the sense that it does not today define everything necessary for a fully functional, operational internetwork), vendors have to date implemented relatively little of the hardware and software needed to support the protocol suite.

It is important here to clarify what is meant when people say that existing networks or network protocols conform to the OSI reference model. Since the reference model is only a definitional and analytic tool that can be used in talking about a network, the only thing that actually conforms to the model is a protocol *definition*. Any given software or hardware implementation that claims to be an OSI Level "X" Protocol within the OSI protocol suite can, at most, conform to a particular protocol definition, which, in turn, may conform to the OSI reference model. Individual software and hardware implementations, therefore, do not conform to the model. Further, different OSI protocols (even at the same layer) are not necessarily compatible. Conformance to the OSI reference model thus is not a guarantee that different systems can work together.

THE LINKED SYSTEMS PROTOCOL (Z39.50)

The Linked Systems Protocol can be viewed as having two components: one explicit, one implicit. The *explicit* component is an application layer (layer seven) protocol, as defined by the OSI reference model discussed above. The *implicit* component exists in that the draft standard assumes that the protocol is operating within an environment defined by the OSI reference model and by the current work in adjoining protocols within the protocol suite (the presentation layer and the common application services layer). This is consistent with the standards committee's charge and with the current state of protocol standardization work. (All work in devising national standards for networking protocols is being carried out within the OSI reference model environment.) In addition, meaningful conformance to the standard can only be considered in the environment of the OSI protocol suite.

The standard proper is an application layer protocol, which *could* be used, with some refinement and augmentation, with protocol suites other than the OSI suite. Since existing telecommunications networks necessarily predate the completion of the OSI protocol suite, it will probably be several years—perhaps a decade or more—before most operational networks adopt the OSI protocol suite. In fact, because replacement of existing protocol suites is so disruptive and expensive, there will undoubtedly be efforts to develop translating gateways and other facilities to permit at least limited interoperability between the OSI protocol suite and other protocol suites, such as IBM's Systems Network Architecture (SNA) and the Department of Defense's Transmission Control Protocol/Internet Protocol (TCP/IP). If these efforts are successful, multiple protocol suites may coexist for quite a long time. Further, because most existing networks were built from the bottom up, they tend to be most advanced and robust in their analogs of the lower layers of the OSI model (layers one through four), which provide network function, and poorest in functionality at the application layers. Thus, there is great interest, from the DARPA (Department of Defense Advanced Research Projects Agency) internet community and the National Science Foundation networking projects, in superimposing higher-level OSI protocols on existing lower-level non-OSI protocols (in existing networks) to obtain functionality quickly and, simultaneously, set the stage for a smooth, gradual transition from existing network protocols to the OSI suite in the future.

The Linked Systems Protocol has little real significance, except in conjunction with an operational telecommunications network. It is the *combination* of the Linked Systems Protocol and telecommunications networks that enables the benefits of each to be used. Reliable, affordable, high-bandwidth telecommunications networks will have to be developed in order to reap the benefits of the protocol simply because the purpose of the protocol is to use telecommunications to link bibliographical systems. Implementations of the protocol will have to be developed for those who want to benefit from existing and future networks.

Therefore, implementation of the Linked Systems Protocol on top of extant computer networks (outside of the OSI protocol suite) is not only possible but *necessary* if the Linked Systems Protocol is to be widely available in the near future.

FUNCTIONALITY

The objectives behind the Linked Systems Project were well focused and specific. The project concentrated on two tasks in bibliographic retrieval:

- the retrieval and transfer of catalog records and
- the retrieval and transfer of catalog authority records.

However, both librarians and library users need additional tasks performed—e.g.,

- retrieval and transfer of records from indexing and abstracting services,
- retrieval and transfer of records from circulation systems, and
- retrieval and transfer of records from acquisitions systems.

Library catalogs are but one special type of bibliographic database. Users of libraries search many different bibliographic databases, such as ERIC, Medline, *Chemical Abstracts*, and so on. It is not the user's fault that different systems have different command languages and contain records in different formats. The differences present problems, and the Linked Systems Protocol is the most plausible approach to overcoming these differences because it provides a means of decoupling databases and user interfaces.

In addition, library users are interested not only in whether a library owns a book but also in whether it is currently available. Both users and librarians often want to know whether their own or another library has a given book on order. Thus, although the Linked Systems Protocol has its roots in the user's need for catalog information, users require a broader range of accessible information, including bibliographies (comprehensive databases of what material exists that meets the given criteria, independent of location) and circulation and acquisitions information (time- and geography-specific information about what physically exists and where it is located).

What is needed is a family of protocols or an extended protocol that can handle retrieval of such a broad range of bibliographic information. Fortunately, the Linked Systems Protocol is evolving in this direction. There is clearly much work to be done, especially since standardized formats have not yet been developed for the *communication* of many different types of records. The protocol is concerned only with the retrieval and movement of records from one machine to another; it does not address the issue of a standard record format, such as MARC, to describe the structure and content of the records being retrieved and moved. Development, refinement, and wide use of MARC or other standard formats in areas such as abstracting and indexing are badly needed.

Beyond all that, reference librarians will recognize the need for protocols that can handle information retrieval from, for example, numerical databases and textual databases. Connections need to be made to developments such as the Association of American Publishers' Electronic Manuscript Project,[8] which defines a communications format for documents in machine-readable form.

IMPLICATIONS OF THE PROTOCOL

We base our speculations concerning the implications of the Linked Systems Protocol on the following assumptions:

- that the protocol will quickly evolve to cover a broad range of bibliographic queries;
- that the development of the protocol will be accompanied by the development of telecommunications networks; and
- that the protocol and the telecommunications networks can be made compatible.

Implications for the Individual Librarian

Most of the discussion of the Linked Systems Protocol has been in the context of the Linked Systems Project and has concentrated on the benefits of the protocol to the major national bibliographic agencies. That perspective can be complemented by examining the benefits to the individual librarian.

A plausible scenario for cataloging in the future is that the cataloger's work station will use the protocol to hunt for a particular record. If the cataloger were to enter a search key for a given book, the work station might check: (1) in the local library catalog; (2) on a locally mounted optical digital disk; (3) in the local or regional union catalog (if there is one); (4) on the bibliographic utility of choice; and (5) on the bibliographic utility of second choice. It would then check through other remote files in some predetermined order until a record was found or the attempt was abandoned, to be replaced by original cataloging or deferral of the search to some later date.

One would hope that the forwarding of the search through successive sources would be automatic, rather than being a series of searches individually initiated by the cataloger. Further, it should be clear that there would be ample opportunity to design software that would cleverly choose the order in which databases are searched, based on a wide variety of criteria. (Basically, the system could employ heuristic algorithms that use information such as the type of material being searched, statistics about the contents of various databases, and the costs of accessing various databases to maximize the likelihood of locating a satisfactory copy of the record in question at minimum cost.)

This scenario is in keeping with the current trend of bibliographic utilities becoming *sources* of data, with *processing* of the data handled in a distributed manner, locally, rather than at the utility.

Implications for the "National Database"

A consequence of the scenario described above is a new, empirical definition of the currently vague concept of the "national database." In practice, this database would become the union of databases conveniently accessible through the Linked Systems Protocol and reachable through whatever network or interconnected networks (internet) a given cataloger uses. Clearly, this makes the development of internetworks linking as many resources as possible a major priority for the library community in the immediate future.[9]

Historically, the concentration of cataloging in the bibliographic utilities meant that location symbols and original cataloging were added to the databases as a byproduct of cataloging. This is decreasingly the case with the distributed, local processing of catalog records that may or may not have been derived from the utilities. (They may, for example, have been derived from optical digital disks of records.) As a result, there is much interest in tape-loading the results of the local creation and processing of catalog records into the utilities. This practice is questionable, however. It can be argued that if the utilities believed in the Linked Systems Protocol, they should use it, not tape-loading, to extend the range of accessible titles. Recourse to the Linked Systems Protocol, at least with selected very large bibliographic databases, would be consistent with the pragmatic definition of the national database as one that is conveniently accessible through the internet via the Linked Systems Protocol.

Implications for the Library User

In considering the implications of the Linked Systems Protocol, we should take the radical step of ending where we should have begun: with the library user, the supposed beneficiary of all the technological apparatus under development. It is the individual scholar who needs to conduct one search after another, who uses bibliographies at all levels (bibliographies of articles as well as catalogs of books), whose work is likely to be stultified if limited to the collections of his or her local library. It is this library patron, concerned with exploring more of the bibliographic universe, who will be the most empowered by the Linked Systems Protocol.

The Linked Systems Protocol can do two things for such an individual library user:

- It can tie together the disparate databases that multiple library automation systems have created at many libraries and allow a patron to establish whether a given book is held by the local library or is on order, and, if held, whether it is out on loan, on reserve, being bound, or (supposedly) on the shelf. It will link together the diverse subsystems involved in comprehensive library automation without the necessity for monolithic integrated systems.[10] This implies the existence of a local network.
- It can give the scholar access to the entire bibliographic universe. Thus, it furthers the tendency that exists today in the scholarly use of libraries to identify first what material exists and only then determine where the material is and how to lay hands on it. This implies the existence of a national internet linked to the user's local network.

The ability to determine the existence of material without reference to its location will become ever more important. Today, some large academic libraries will find a book or article for the user, either in the library's own collection or through interlibrary loan, freeing the user from the need to pay much attention to physical location. Tomorrow, delivery of much material may be accomplished through electronic imaging technologies, rendering the physical location of the material almost irrelevant.[11] In such an environment, the traditional method of keeping separate catalogs for different libraries and separating book and serial titles becomes increasingly unhelpful, whereas a broad bibliographic approach that unites the catalogs of all major research libraries and links diverse types of bibliographies, such as bibliographies of articles and bibliographies (or catalogs) of books, becomes ever more useful.

All libraries stand to gain a great deal from the implementation of the Linked Systems Protocol, but it is interesting to note which libraries will benefit most. Traditionally, smaller libraries have a greater interest in large libraries than vice versa. One might posit an inverse law of linked systems benefits: the smaller the library, the greater the benefit. Further, small libraries are more numerous than large ones: hence, there are more beneficiaries.

Just as there are many more small libraries than large ones, so also there are many more library users than small libraries. The most exciting implication of the Linked Systems Protocol is that the greatest benefit will be derived when the protocol—and, of course, supporting networking capabilities—are a standard feature of the scholar's work station.

REFERENCES

1. Current drafts of Z39.50 are available through Ray Denenberg, current chair of NISO Committee D, at the Library of Congress.
2. S. H. McCallum, "Linked Systems Project, Part 1: Authorities Implementation," *Library Hi-Tech* 10:61–68 (1985).
3. R. Denenberg, "Linked Systems Project, Part 2: Standard Network Connection," *Library Hi-Tech* 10:71–79 (1985).
4. C. A. Lynch, "Linked Systems Protocol—A Practical Perspective," in *Telecommunications Networks: Issues and Trends*, M. E. Jacob, ed. (White Plains, NY: Knowledge Industries Pr., 1986), p.67–91.
5. A. Tanenbaum, *Computer Networks* (Englewood Cliffs, N.J.: Prentice-Hall, 1981).
6. M. A. Padlipsky, *The Elements of Networking Style and Other Essays and Animadversions on the Art of Intercomputer Networking* (Englewood Cliffs, N.J.: Prentice-Hall, 1985).
7. C. A. Lynch, "Protocols in Perspective," *ASIS Bulletin* 11:9–11 (Aug./Sept. 1985).
8. *Electronic Manuscript Project: Standard for Electronic Manuscript Preparation and Markup* (Washington, D.C.: Association of American Publishing, 1986).
9. C. A. Lynch and E. B. Brownrigg, "The Telecommunications Landscape: 1986," *Library Journal* 111, no.16:40–46 (Oct. 1986).
10. ________, "Beyond the Integrated Library System Concept: Bibliographic Networking at the University of California," in *Proceedings of the Second National Conference on Integrated Online Library Systems,* D. C. Genaway, ed. (Canfield, Ohio: Genaway and Associates, 1984), p.243–52.
11. ________, "Library Applications of Electronic Imaging," *Information Technology and Libraries* 5:100–105 (June 1986). ■■

Desktop Typesetting and the *LITA Newsletter*

Walt Crawford

The LITA Newsletter *recently changed production techniques, moving from typesetting and traditional galley pasteup to production of page proofs on an inexpensive desktop laser printer. The change was intended to improve the timeliness of the newsletter and allow more pages for expanded coverage without expanding the budget.*

In this article the editor of LITA Newsletter *discusses the background for the decision, the current production techniques, and some benefits and drawbacks of desktop typesetting.*

The *LITA Newsletter* had minor problems. Production took ten weeks from the time that people sent in reports until the issues were mailed out. Space was limited, with three issues each year limited to eight pages and the postconference issue limited to twelve pages. With LITA's burgeoning interest groups and greater conference activity, some conference reports were delayed an extra three months, and there was little room for new topics. LITA had budget problems, and since the *Newsletter* brings in no new revenue, more money for more pages would be difficult to justify.

Contemporary technology offered a solution to the problem: desktop typesetting, or desktop publishing without graphics. This new technology makes it possible to produce almost twice as many pages, cut the production cycle by almost half, and cut the production budget significantly. This article discusses the background for the change in production, some of the benefits and drawbacks of desktop typesetting, the way the change proceeded, and specific aspects of the new production technique.

STATE OF THE NEWSLETTER

The *LITA Newsletter* began when the Information Sciences and Automation Division (ISAD) of ALA became the Library and Information Technology Association in 1980. Pat Barkalow served as the first editor, producing eight pages an issue, three times a year. Working with ALA's Central Production Unit (CPU), Pat initiated the maroon *LITA Newsletter* banner, the two-column format, and an approach to LITA coverage that relied heavily on columns from each section of LITA. The regular column "Standard Fare" also began under Pat's editorship, at the instigation of the Technical Standards for Library Automation Committee (TESLA). Carol Parkhurst became the second editor in summer 1982. During her editorship, the *Newsletter* began quarterly publication, postconference issues grew to twelve pages, and new typefaces were adopted for the banner, headlines, and body text. These two editors developed a strong newsletter, with good coverage of LITA activities. The newsletter was professionally produced with attractive typography.

I was appointed as the third *LITA Newsletter* editor, beginning with the 1985 postconference issue. At the same time, the LITA reorganization took full effect. The

Walt Crawford is Principal Analyst for Special Projects at the Research Libraries Group and editor of the *LITA Newsletter*.

sections were dissolved, eliminating section editors and a steady supply of section coverage. The result was not a shortage of copy but an excess—some 1985 conference coverage was held over to the following issue. The same thing happened with the post-Midwinter issue—almost three pages of the Summer 1986 issue consisted of delayed reports from Midwinter.

LITA had abandoned its sections but increased its activities. During the 1984 Annual Conference in Dallas, LITA had nineteen functional meetings (board, executive committees, program planning, etc.), four programs, and eighteen meetings devoted to specific areas of interest, including seven discussion groups. During the 1986 Annual Conference in New York, the simplified LITA required only eleven functional meetings. There were only three programs, but twenty-five meetings were devoted to specific areas of interest, including ten interest groups. Since then, even more interest groups have formed. LITA is becoming involved in more areas, requiring more coverage by the *LITA Newsletter*.

More coverage means more pages. Better coverage would mean a shorter delay between reporting and publication. I also wanted to encourage new topical columns along the lines of "Standard Fare" and "MARC Notes"; such columns would also need space.

It was clear that expansion might be difficult if it involved a larger budget. LITA is still recovering from the effects of a dues increase, and the new ALA accounting methods help to make LITA's budget tight. The best solution to the problems of the *Newsletter* would reduce the budget or at least not increase it. Other events made one particular alternative attractive. I had started to use an IBM PC/XT at RLG, along with a word-processing program (The FinalWord) with unusually strong printer control capabilities. RLG had purchased three Hewlett-Packard LaserJet laser printers for various purposes, offering some direct experience with low-cost laser printing. Finally, a special discount program made it possible for me to acquire a single HP LaserJet at a little more than half-price. After seeing some other uses of the LaserJet and doing some investigation, I came to the conclusion that it would be possible to use the PC/XT, The FinalWord, and the LaserJet to prepare page proofs for the *LITA Newsletter*. It would allow faster production and, by eliminating typesetting charges, save enough money to amortize the printer and increase the post-Midwinter issue to twelve pages. In January 1986, I drew up a proposal to implement desktop preparation for *LITA Newsletter* page proofs.

DESKTOP PUBLISHING AND DESKTOP TYPESETTING

The term *desktop publishing* typically refers to the use of a microcomputer and laser printer to generate finished pages that combine text and graphics and that typically use multiple text fonts. Contemporary usage also usually means that pages can be designed and viewed on the microcomputer screen just as they will appear when printed—WYSIWYG layout—What You See Is What You Get. The word *publishing* is a misnomer; desktop publishing systems actually support desktop page production, and page production is only one aspect of publishing.

WYSIWYG layout and the need to combine text and graphics make desktop publishing complex and moderately expensive. A full desktop publishing system costs at least \$8,000: \$4,500 for an Apple LaserWriter, at least \$1,000 for word processing, image-building, and page-layout software, and typically more than \$2,500 for the hard-disk computer and high-resolution display needed to work effectively. A scanner, to import photographs and other images to the computer, will add another \$2,000, bringing the total to \$10,000 or more.

Taste, Talent, and Typography

Desktop publishing as defined above appears to be very "high tech," using the computer display and software to replace scissors, paste pot, T-square, and other tools of traditional publication layout. While the computer does replace those tools, it doesn't do the work itself. The user must still determine typefaces and fonts, see that the copy is properly edited, see that the pages are filled and balanced, and do all the other production work. The computer will assure that headlines are "pasted up" in a straight line but won't do anything to make sure that a page is legible, attractive, or coherently arranged.

The computer can replace the typesetter, but not the typographer. Choosing the appropriate faces and fonts still requires taste and talent. If anything, most desktop publishing systems make layout and typography more difficult, since it is not usually possible to view the full page at once (without eliminating detail), and type displayed on the screen will not exactly match type as printed.

When skilled typographers and layout artists use desktop publishing systems, the results are faster production at lower cost. When desktop publishers lack an eye for appropriate combinations of type and must work with parts of a page, the results can include days of wasted time and some peculiar results.

Desktop Typesetting: A Simpler Alternative

Full desktop publishing did not appear to be a useful solution for the *LITA Newsletter*. The *Newsletter* has never used graphics extensively and, in recent issues, used few photographs; the need was for more text, not more graphics. I could contribute basic hardware in the form of an IBM PC/XT—but, at least in early 1986, the best desktop publishing software required an Apple Macintosh. I had read enough accounts of desktop publishing to know that I lacked the patience for the extensive experimentation needed to make page layout programs produce attractive results.

The *LITA Newsletter* required a system that would produce effective pages based on what the editor wanted—not require the editor to keep manipulating onscreen images for hours on end. I proposed a system that did not require graphics, required only a small set of type fonts and faces, and did not require that the pages appear onscreen as they would in print. An apt term for the system I wanted to use would be *desktop typesetting*—although the software and hardware support page layout as well.

The proposal called for a Hewlett-Packard LaserJet, the most widely sold desktop laser printer on the market and, at the time of the proposal, the cheapest laser printer that supported proportional typefaces. It also called for the *F* font cartridge, containing well-designed versions of typefaces similar to Times Roman and Helvetica: the former for body text and small headlines, the latter for larger headlines. With a special discount the combination of LaserJet and font cartridge cost just over $2,000—making the experiment quite inexpensive, since typesetting was costing more than $1,500 a year.

Naturally, desktop typesetting also requires a computer and software. For the initial experiment, RLG provided both. I use an IBM PC/XT at work, and RLG uses The FinalWord II as its common microcomputer text formatting system. The combination of The FinalWord II, LaserJet and *F* cartridge provides a powerful and inexpensive desktop typesetting system, but one that requires several trial runs to produce the best results.

The Hewlett-Packard LaserJet

Hewlet-Packard has sold more than 300,000 LaserJets since the model was first introduced. The LaserJet was initially intended to be a replacement for a daisy-wheel printer and works well in that capacity. At full retail, it costs less than twice as much as a fast daisy-wheel printer; at up to eight pages a minute (or, for single-spaced pages, 400 characters a second or more), it can print six to eight times as fast as the fastest daisy-wheel printers; and, with the noise level of a desktop copier, it is much less offensive than a daisy-wheel printer. It can't produce carbon copies, but it can produce multiple fonts on the same page.

The LaserJet is based on a laser print engine made by Canon in Japan. The same print engine is used for the Apple LaserWriter and quite a few other desktop laser printers and is based on Canon's desktop copier. It is intended for medium duty—200 or 300 pages a day at most. As a replacement for several daisy-wheel printers, that's plenty, but the unit is not designed for continuous printing several hours a day, day in and day out.

Hewlett-Packard designed several font

cartridges for the LaserJet and soon found themselves in the typesetting business, particularly with the *F* cartridge. The *F* cartridge contains a Times Roman–like face in 10 point medium, bold, and italic; a similar face in 8 point medium; and a Helvetica-like face in 14.4 point bold, suitable for medium-sized headlines. (The point is the standard measure of type size; there are 72 points to an inch.) All the faces are proportional and well designed. The printer has a resolution of 300 dots per inch (90,000 dots per square inch), and the output quality is better than typewritten, though not as good as typeset. Where dot matrix printers strive for "near letter quality" output, the LaserJet and other laser printers achieve "near typeset quality."

The LaserJet was not designed for intensive graphics. It can only print a small portion of a page in high-resolution graphics. It was also not designed for complex combinations of fonts. Hewlett-Packard later introduced the LaserJet Plus, allowing larger areas of high-resolution graphics and making it possible to have fonts defined in software as well as on cartridges. Neither printer comes close to the sophistication of the Apple LaserWriter and LaserWriter Plus, both of which support full-page graphics at high resolution, and both of which support a wide range of typefaces in any point size desired.

The LaserJet's prime advantages are price, market share, and the Hewlett-Packard name. The LaserJet costs about half as much as the LaserWriter, and HP sells several times as many LaserJets as Apple sells LaserWriters. Because HP has sold 80% or more of all desktop laser printers and because most of those printers are connected to PC-compatible computers, many PC-DOS software packages support the LaserJet. Microsoft Word, PC-Write 2.6, WordPerfect and The FinalWord II all offer strong support for the LaserJet, as do many other software packages.

The FinalWord II

The FinalWord II provides a wide range of controls over the appearance of printed output, but such controls are represented by text within a document. When a writer takes advantage of FinalWord's advanced capabilities, the system is almost the antithesis of WYSIWYG—what you see looks very little like what you'll get. Full use of FinalWord means preparing output styles that define what output should look like. I defined an overall output style called *LN* (for *LITA Newsletter*), specifying top and bottom margins, left and right margins and "binding edge" (wider left margin on odd-numbered pages, wider right margin on even-numbered pages), and other characteristics of *LITA Newsletter* pages. That output style also includes descriptions for the spacing of headlines, boxed headlines and items, and special handling for "bulleted" text, where each paragraph is preceded by a small black box or bullet.

FinalWord output styles take effect when a document is printed, as do characteristics of the printer. Text is entered and edited in any convenient form, with special text to indicate special needs such as headlines and bullets. Figure 1 shows the beginning of a column in its FinalWord form; figure 2 shows the same column as it would appear in the *LITA Newsletter*.

The FinalWord II offers another advantage for the *LITA Newsletter*, where some of the contributions arrive in machine-readable form (over ALANET, BITNET, or RLIN mail): it uses pure ASCII text, with no imbedded control characters or other special characters. If a contribution is in machine-readable form, it is only necessary to edit the copy and add headlines and other features.

Desktop Typesetting: An Inexpensive Solution

At today's prices, a person can put together a complete desktop typesetting system for roughly half the price of a good desktop publishing system. The LaserJet has a normal discount price of around $2,400, and PC-DOS computers with hard disks are available for $1,200 to $1,500. Adding $450 or so for the font cartridge and word-processing system brings the total to a little more than $4,000. The FinalWord doesn't try to represent the final page in detail, so the computer doesn't need a high-resolution graphics display.

Four thousand dollars will pay for 100 to 400 pages of typesetting; in the case of the *LITA Newsletter*, it would pay for roughly 115 pages. There are extra costs for pages of

```
@Box [
@Ca [Standard Fare]
@Cb [Pierre Badin LaTes, III
Column Editor]
]
@Lb [NISO (National Information Standards Organization)]
   NISO Standards Committee FF has been newly formed to develop standards for the
description of computer software. The standard will establish what eye-legible data should
be printed on the physical medium carrying the computer software (e.g., a floppy disk), on
the the physical packaging for the medium (e.g., a floppy disk envelope), and on the package
in which the medium is stored (e.g., a box or plastic bag).
   A number of draft standards will be available in late 1986 or early 1987:
@Itemize[
SC G (Common Command Language for Online Interactive Retrieval Systems) is reviewing
the extensive comments that were received on the first draft, which was distributed this
spring. The committee will produce a second draft in early 1987.

A revised draft of Z39.18-1974 (Guidelines for Format and Production of Scientific and
Techni@| cal Reports) was circulated for comment in March and will be balloted in
September by NISO voting members.

SC AA (Interlibrary Loan Data Elements) will circulate a draft standard for review and
com@| ment, also late in 1986.
]
```

Fig. 1. Text for LITA Newsletter.

LaserJet output, but they are very small. The LaserJet uses a cartridge combining toner (ink) and those elements of the printing mechanism subject to fast wear; that cartridge retails for $90 and should last for 2,000 to 3,000 pages, a per-page cost of three or four cents. Any paper (or other flat surface) that will work in a plain-paper copier will work in a LaserJet; good quality twenty-pound copier paper rarely costs as much as one cent a page. Thus, materials should cost no more than five cents a page. I estimated that it would take ten trial runs to get an issue of the *LITA Newsletter* to look right. For a twelve-page issue, that might use up as much as $6 worth of supplies—and take about five minutes to format and print each trial run.

EXPECTED BENEFITS AND DRAWBACKS OF DESKTOP TYPESETTING

Desktop typesetting isn't a universal solution; for a given situation, the technique will offer some benefits and some drawbacks. I anticipated the following benefits for the *LITA Newsletter:*

• Enough savings from eliminating typesetting costs to allow two or three twelve-page issues each year, while amortizing the equipment over five years.

• Single keying, with copy being entered only once. If copy was received in machine-readable form, no editorial keying would be required. In any case, the second keying by the typesetter would be eliminated. Additional keying takes time and introduces new errors.

• Greater editorial control, since layout would be done as part of editing and editorial revisions could be made through the last day of page production. If an issue ran a quarter-page too long, the editor could trim words, sentences, and paragraphs, rather than eliminating whole articles or sections of articles. Detailed editorial revision is difficult working over a distance of 2,000 miles; when minor changes require additional typesetting charges and production delays, minor changes aren't reasonable.

• Faster production, by eliminating typesetting, pasteup, and the mailing of galleys and page proofs. The old production schedule required seventy days from the copy deadline to the date issues were mailed; of that time, the editor had one week to assemble and edit the issue. I esti-

Standard Fare

Pierre Badin LaTes, III
Column Editor

NISO (National Information Standards Organization)

NISO Standards Committee FF has been newly formed to develop standards for the description of computer software. The standard will establish what eye-legible data should be printed on the physical medium carrying the computer software (e.g., a floppy disk), on the physical packaging for the medium (e.g., a floppy disk envelope), and on the package in which the medium is stored (e.g., a box or plastic bag).

A number of draft standards will be available in late 1986 or early 1987:

- SC G (Common Command Language for Online Interactive Retrieval Systems) is reviewing the extensive comments that were received on the first draft, which was distributed this spring. The committee will produce a second draft in early 1987.
- A revised draft of Z39.18-1974 (Guidelines for Format and Production of Scientific and Technical Reports) was circulated for comment in March and will be balloted in September by NISO voting members.
- SC AA (Interlibrary Loan Data Elements) will circulate a draft standard for review and comment, also late in 1986.

Fig. 2. Formatted Version of Figure 1.

mated that desktop typesetting could reduce that to forty-nine days, allowing three weeks for editorial work and page preparation.

• Narrower demands on the editor's time, since there would be no need to proofread galleys and page proofs sent from Chicago. Under the old schedule, the editor could not reasonably plan vacation or business trips during a three- or four-week period after copy was sent to CPU. While the new technique would require more actual work on the editor's part, it would be concentrated into a short period; once page proofs were mailed, nothing more would be required.

There would definitely be drawbacks to desktop typesetting. Those drawbacks included the following:

• Heavier workload for the editor, since all copyediting and hyphenation would have to be done as part of editing. The FinalWord II has no automatic hyphenation routines. "Soft hyphens"—places where lines can be broken within words—must be added manually after inspecting initial galleys. The increased need for copyediting and fact checking led me to propose adding an associate editor at the same institution to share the load.

• Loss of the fact checking and other editorial work done by CPU. That was another reason to add an experienced associate editor. Since the first issue was produced, LITA's executive director has also asked to review the galleys prior to issue production, a checking step that previously took place during CPU's work.

• Loss of print quality, since 300 dots per inch is far below typesetting resolution of 1,200 to 2,400 dots per inch. I knew that 300 dots per inch produced characters with little or no graininess and no "dottiness," but I also knew that the LaserJet printing I had seen was *near* typeset quality. I could certainly tell the difference between LaserJet output and typesetting but felt that the loss of quality would be acceptable for a newsletter.

• Difficulty of handling photos, since they would have to be added as crude halftones produced by photocopying them with a white-dot overlay. An electronic scanner might produce somewhat better quality but would double the hardware cost and require more sophisticated software.

• Loss of special features such as boxed items, shaded items, and varying type sizes for headlines and body type. Horizontal rules are available on the LaserJet (and can be done with varying thickness, with some difficulty); boxes could be created by entering vertical lines after printing, but I wanted to avoid manual layout work as a potential source of problems. Shading could also be added after printing but would probably not work as well. Varying type sizes for headlines appeared to be less of a problem; while the type repertoire of the *F* cartridge is limited, it appeared suitable for this purpose.

• Loss of type density, potentially requiring more pages to print the same text. The *LITA Newsletter* was set using 10 point Garamond on 11 point leading (the total space allowed for each line). Garamond is a relatively narrow type. My initial experiments used the *B* cartridge, Times Roman, set 10 point on 10.9-point leading, and showed that the characters were wide enough to require 10 to 15 percent more space for the same text. Fortunately, the increased number of pages far outweighed the loss of type density.

The drawbacks were obvious. The benefits were also obvious and outweighed the drawbacks. Perhaps the clinching argument was that it makes sense for LITA to use contemporary technology to produce its newsletter. I ordered a LaserJet in May 1986 and received it in June, leaving six weeks to become familiar with the system and experiment with techniques before issue 26 was due.

IMPLEMENTING DESK TYPESETTING

Between May 1986 and August 12, 1986, when camera-ready pages for issue 26 were mailed to CPU, I got more complete information on the *LITA Newsletter* budget and the effects of the experiment. That period also allowed time to experiment extensively with various aspects of desktop typesetting, resulting in some changes in technique.

The first requirement was to appoint an associate editor, one who could offer informed advice along the way and provide editorial support promptly during the actual production cycle. Kathleen Bales, manager, Systems Analysis and Design at RLG, agreed to serve as associate editor. She was already active in LITA, had previous editorial experience, and offered all the other talents needed for the job.

The second requirement was to get hard numbers for the budget. How much would an eight-page, twelve-page, or sixteen-page issue cost without typesetting? Were there other budget implications that I'd overlooked? That process involved some pleasant surprises:

• While a twelve-page issue costs $1,252 for printing as compared to $924 for the initial eight pages (including the second color on the front page), a sixteen-page issue costs $1,277.80—about $25 more than a twelve-page issue. In other words, the major choice is really between eight-page issues and sixteen-page issues, with twelve pages making sense only if there is not enough copy for sixteen pages. Postage and handling are the same for all three issue sizes. Enough will be saved from typesetting to pay for the LaserJet (over five years) and to increase all four issues to sixteen pages if that is appropriate.

• I had ignored charges for CPU's time, partly because I had never seen a detailed budget for the previous year. The revised estimate from CPU, eliminating copyediting and layout, cuts their charges in half, eliminating more than $1,900. Thus, the overall budget was actually reduced significantly.

Typography and Layout

The 10 to 15% extra space required for wider letters bothered me, although the surprising development that total pages could be increased by 78% rather than 22% made the space requirements less troublesome. The *F* cartridge uses narrower letters than the *B* cartridge, reducing the extra space requirements to as little as 5%. Thus, the effective increase in available space could be as much as 70% or more. In practice, the numbers are even better. The banner and mailing strip take up half a page, so that the actual increase was from thirty-four available pages to sixty-two available pages, or a net effective increase of almost 78%.

The FinalWord II defines the standard *F* body text as 10 points set 10.9 (using 0.9 points of leading). The smaller body text, 8 points, was set 9; the headline text, 14.4 points, was set 15. One point of leading is typical for typeset material, but the results seemed a bit crowded and thick on LaserJet output. That makes sense; LaserJet fonts are somewhat thicker than typeset equivalents.

Most other word-processing packages set body text at 12 points, using a normal six lines per inch. Two points of white space, while only one thirty-sixth of an inch, is a bit too much for typeset material and is one of the typical characteristics of typewritten or computer-printed material. The other characteristic, excessive white space between words in fully-justified text, can be controlled by careful hyphenation.

After some experimentation, I arrived at intermediate leading for the *LITA Newsletter*. Body text is set 11.4; small text (used in Board of Directors minutes) is set 9.5; Helvetica headlines are set 16.5. After making those changes and adjusting other aspects of the output format, the resulting pages have just one line less in a full column than the typeset *LITA Newsletter*, a loss of 1.6%, resulting in much more legible pages. Between slightly wider letters and slightly more leading, the new body typeface requires about 7% more space for a given set of text; thus, thirty-four pages of the old format will require a little more than thirty-six pages in the new format. The real difference is more complicated but works to reduce that increased requirement.[1] Figures 3 and 4 show two sections of a typeset *LITA Newsletter* issue; figures 5 and 6 show the same text prepared using the LaserJet and The FinalWord II.

Producing Issue 26

The new techniques clearly cut the production cycle by at least three weeks. For the first year, it seemed sensible to use two of those three weeks to allow interest group and committee chairs two more weeks to submit conference reports, using the other week to make conference reports more immediate. The first desktop-typeset issue had an unusually late deadline for conference reports. That was probably a mistake. Most reports came during the conference or shortly thereafter, and the last few reports trickled in very late. For future years, the schedule improvement will be used to make the postconference issue appear much sooner after the conference.

The production schedule allowed three weeks for editorial work. Even though issue 26 was the first LaserJet-produced issue and was the longest issue in the history of the *LITA Newsletter*, three weeks was more than enough time—the issue was ready a full week ahead of deadline. Depending on the editor, a realistic production cycle may be less than six weeks.

Even with sixteen pages, there was more copy than space. Contributors to the *LITA Newsletter* are submitting reports that include the substance of a meeting. With more pages available, it make sense to include that substance and certainly serves

Authority Control in the Online Environment

Barbara B. Tillett
Topical Editor

This column is intended to provide news on activities, products, publications, etc., on authority control in the online environment. If you have items you would like included, please contact the column editor, Barbara B. Tillett, Scripps Institution of Oceanography, Library C-075C, University of California–San Diego, La Jolla, CA 92093.

ALA Midwinter News

The LITA/RTSD CCS Interest Group on Authority Control in the Online Environment met several times during the ALA Midwinter Meeting in Chicago. A regular part of the interest group's business meeting in an update on authority control news, which includes publications, products, LC update (NACO, LCSH tapes, LSP), studies and research, and plans for future meetings. The following are highlights from the January 1986 meetings.

Two publications of interest are Robert Burger's *Authority Work,* published by Libraries Unlimited, 1985, and the 1984 IFLA publication, *Guidelines for Authority and Reference Entries.*

A new product to produce authority and reference cards from the online national resource authority file via OCLC, *Lance's Authority System,* is available from Xavier University Libraries at $5 for the product disk and $5 for the manual.

Fig. 3. Typeset Material, Example 1.

the profession better. The result of that service is that the first sixteen-page issue still forced three pages of copy over to the next issue.

The original estimate of ten trial runs was fairly accurate. A typical article requires two passes for copyediting and three passes for hyphenation. The full newsletter requires from three to six passes to balance pages, avoid short columns, and otherwise make the full issue look right.

The FinalWord II does an aggressive job of preventing widows and orphans (single lines stranded at the top or bottom of a page) and of assuring that headlines appear together with the first two lines of an article. The effect of all this intelligence is that FinalWord can leave some pages with quite a bit of white space at the bottom. Page six of issue 26 shows a mild example of the problem, a page that is two or three lines shorter than other pages. I needed three passes to avoid at least one page that was a full inch shorter than other pages.

The most pleasant surprise in issue 26 was legibility. I had anticipated some loss

LITA Board of Directors: Highlights of Meetings at the 1985 ALA Annual Conference

Editor's Note: Board motions appear as worded. Some issues covered in other newsletter columns and articles are not repeated here. Full minutes are available on request from the LITA Office, American Library Association, 50 E. Huron St., Chicago, IL 60611.

The LITA Board of Directors met three times during the 1985 Annual Conference in Chicago. Members are Nancy Eaton, Lois Kershner, Kenneth Dowlin, Brian Aveney, Pat Barkalow, Michael Gorman, Ernest Muro, Louella Wetherbee, Bonnie Juergens, Mary Ghikas, Charles Kritzler, Jo-Ann Michalak (ex officio), and Don Hammer (ex officio). All votes were unanimous.

Optical Information Systems Interest Group. Nancy Eaton has found considerable interest among the membership in optical information systems and is putting together an interest group. After brief discussion, it was moved by Dowlin, seconded by Kershner, and passed that:

The Optical Information Systems Interest Group be approved as an Interest Group of LITA. (Absent: Barkalow, Juergens, Kritzler, Muro.)

INDEX: EDUCATION 1986. Don Hammer asked the board to consider cosponsorship of INDEX: EDUCATION at INFOMART in 1986, suggesting a tentative agreement to be fleshed out after INDEX: EDUCATION 1985 takes place. If LITA is to take part, it should set up a program committee now and begin planning. Lois Kershner has asked the Education Committee to be the planning group for next year.

Members of the Board raised a number of concerns about timing and some other aspects of INDEX: EDUCATION and its possible effects on other LITA activities. LITA is cosponsoring the INFOMART conference on behalf of ALA at Bob Wedgeworth's request; if LITA decides not to continue, ALA may ask some other unit to serve as a cosponsor.

Fig. 4. Typeset Material, Example 2.

in print quality, but felt that the new version would still be fairly legible. When the issue finally appeared in print, my own reaction was the same as the consistent reaction of those who have commented: the new version is actually *more* legible than the old. While the Garamond is certainly more elegant, the slightly heavier font produced by the LaserJet, with slightly more white space between lines, is bolder and easier to read. What appeared to be a drawback has turned into a benefit.

IMPLICATIONS FOR THE FUTURE

The experiment in desktop typesetting is off to a successful start. With the help of

Authority Control in the Online Environment

Barbara B. Tillett
Topical Editor

This column is intended to provide news on activities, products, publications, etc. on authority control in the online environment. If you have items you would like included, please contact the column editor: Barbara B. Tillett, Scripps Institution of Oceanography, Library C-075C, University of California, San Diego, La Jolla, CA 92093

ALA Midwinter News

The LITA/RTSD CCS Interest Group on Authority Control in the Online Environment met several times during the ALA Midwinter Conference in Chicago. A regular part of the Interest Group's business meeting is an update on authority control news, which includes publications, products, LC update (NACO, LCSH tapes, LSP), studies and research, and plans for future meetings. The following are highlights from the January 1986 Midwinter meetings.

Two publications of interest are Robert Burger's *Authority Work*, published by Libraries Unlimited, 1985, and the 1984 IFLA publication, *Guidelines for Authority and Reference Entries*.

A new product to produce authority and reference cards from the online national resource authority file via OCLC, *Lance's Authority System*, is available from Xavier University Libraries at a cost of $5 for the product disk and $5 for the manual.

Fig. 5. LaserJet Version of Figure 3.

many LITA members, the *LITA Newsletter* should see a growing number of topical columns, longer and more substantive reporting, and generally broader coverage of the division. Conference and meeting reports should appear sooner, and preconference issues should be more up-to-date.

It may take a year to prove the experiment a success, but I'm willing to bet that it will be. Desktop typesetting seems to be a situation in which everybody gains. The editor gains full-time use of an enormously productive tool and gains greater control over a larger newsletter. LITA member-

LITA Board of Directors: Highlights of Meetings at the 1985 ALA Annual Conference

Editor's Note: Board motions appear as worded. Some issues covered in other newsletter columns and articles are not repeated here. Full minutes are available on request from the LITA Office, American Library Association, 50 E. Huron St., Chicago, IL 60611.

The LITA Board of Directors met three times during the 1985 Annual Conference in Chicago. Members are Nancy Eaton, Lois Kershner, Kenneth Dowlin, Brian Aveney, Pat Barkalow Michael Gorman, Ernest Muro, Louella Wetherbee, Bonnie Juergens, Mary Ghikas, Charles Kritzler, Jo-Ann Michalak (ex officio) and Donald Hammer (ex officio). All votes were unanimous.

Optical Information Systems Interest Group. **Nancy Eaton has found considerable interest among the membership in optical information systems and is putting together an interest group. After brief discussion, it was moved by Dowlin, seconded by Kershner, and passed that:**

> **The Optical Information Systems Interest Group be approved as an Interest Group of LITA. (Absent: Barkalow, Juergens, Kritzler, Muro.)**

INDEX: EDUCATION 1986. **Don Hammer asked the board to consider cosponsorship of INDEX:EDUCATION at INFOMART in 1986, suggesting a tentative agreement to be fleshed out after INDEX:EDUCATION 1985 takes place. If LITA is to take part, it should set up a program committee now and begin planning. Lois Kershner has asked the Education Committee to be the planning group for next year.**

Members of the Board raised a number of concerns about timing and some other aspects of INDEX:EDUCATION and its possible effects on other LITA activities. LITA is cosponsoring the INFOMART conference on behalf of ALA at Bob Wedgeworth's request; if LITA decides not to continue, ALA may ask some other unit to serve as a cosponsor.

Fig. 6. LaserJet Version of Figure 4.

ship gains more extensive and livelier coverage in a more timely fashion and a budget saving at the same time.

Problems may arise with the next editor. As CPU pointed out, this method of production requires that an editor have the tools and knowledge to make it work. Two years from now, the cost of tools may be irrelevant; five years from now, it almost certainly will be. In two years, the combined cost of PC-DOS compatible computer, laser printer, and software is likely to be little more than $3,000; at that price, a complete system could be amortized over an editor's first three-year term and still make good economic sense.

The larger problem is the interest and patience needed to make desktop typesetting work. The article recruiting candidates for editor in 1985 said that word-processing skills were nearly mandatory. That's reasonable, but is it reasonable to insist that an unpaid editor should be willing to hyphenate copy and do layout by iterative refinement? I regard it as an interesting challenge—in short, as fun—but some other well-qualified potential editors might regard it as a form of torture.

Two years is a fairly long time in the microcomputer field; five years is almost a lifetime. After the current desktop publishing mania dies down and people see what works and what doesn't, there will be a large (though not enormous) market for sensible tools to produce high-quality publications. The next editor should find programs that reduce the guesswork while retaining control. By some standards, The FinalWord II is archaic software; however, it still represents a level of power and control available in few other programs. More good choices should be available in a few years.

CONCLUSIONS

Desktop typesetting and desktop publishing solve some problems and cause others. To get the best results takes patience, persistence, and a willingness to experiment and compromise. For those willing to spend the time and energy, desktop publishing can improve the appearance of previously typed documents and reduce the cost of typeset publications.

Desktop typesetting makes sense for the *LITA Newsletter*. It also makes sense for libraries who need handsome documents without typesetting costs. It isn't a good universal replacement for typographers, typesetters, and the other craftspeople of the printing industry, but where it works, it works very well.[2]

REFERENCES AND NOTES

1. The new "Board Highlights" typeface is denser than the old one, and headlines are smaller in the new format.
2. A related article, "Common Sense and Desktop Publishing," in *Library Hi Tech* 5, no.1 (1987), will cover aspects of desktop publishing in more detail. ■■

Database Management Principles of the UCLA Library's Orion System

James Fayollat and Elizabeth Coles

ORION, the integrated online library system developed by the UCLA Library, incorporates a number of database management features that enhance system efficiency for record retrieval and display. The most important of these features is the physical storage of records in alphabetical order within logical files. Specific design features related to record storage and to retrieval are described, as are the design and structure of linked files that support authority control and provide additional paths for user access to the database.

LIBRARY AUTOMATION AT UCLA: THE ORION SYSTEM

The online information system began more than twenty years ago as a serials processing system developed in the UCLA Biomedical Library with National Library of Medicine grant funding.[1-4] In the early 1970s the system was expanded to include cataloging and monographic acquisitions functions.[5-6] In 1977 the UCLA Library administration decided to expand the system for use by all campus libraries and to enhance its capabilities in order to create a truly integrated online library information system.[7] In 1982, when public access was implemented, the technical processing system was renamed ORION.

ORION now controls all technical processing, fund accounting, and catalog maintenance activity for a large, academic, research library system consisting of twenty-one libraries on the UCLA campus. A linked authority control module dynamically maintains subject, name, and series access points to bibliographic records for both cataloged and in-process materials. An integrated circulation/inventory module was scheduled for implementation in spring 1987. Public access to all materials, whether on order, in processing, or cataloged, is provided to the campus community from terminals in campus libraries, departmental offices, or, on a dial-access basis, from home microcomputers. ORION also supports databases for several collections in other campus departments and off-campus institutions.

ORION operates on the Office of Academic Computing's IBM 3090 model 200 mainframe computer, under IBM's MVS/XA operating system and TSO terminal monitor. This computer also handles a large portion of campus academic computing activity; at present, ORION uses approximately 10 to 15 percent of the IBM 3090's capacity. Nearly 200 dedicated ORION terminals are located in campus libraries. ORION programs, which are written in the PL/I programming language, are developed and maintained by a staff of programmer/analysts in the library. These locally developed programs support a full range of information system functions from the database management level through record retrieval, display, and editing capabilities.

James Fayollat is Head, ORION Systems Office, and Elizabeth Coles is Senior Programmer/Analyst at the UCLA Library.

FOCUS ON DATABASE MANAGEMENT

The heart of ORION is its database management system, which incorporates a number of features that make it especially well suited for library applications. The most important of these features is the physical storage of records in alphabetical order, rather than in random or chronological order. The latter method is widely used and suitable for applications where efficient and inexpensive addition of records to the database is more important than fast, low-cost retrieval and display of those records, or where the desired order for display of records cannot be easily predicted. Random storage, however, imposes penalties for retrieval. Before records satisfying a query can be displayed, they must either be sorted or recourse made to an intermediate index maintained in the desired order. Inefficiencies in retrieval may become more noticeable as a database grows and it becomes more likely that a query will result in the retrieval of a large number of records.

If the ultimate aim of an automated library system is to serve the information needs of the library's users, including staff, a compelling case can be made for adopting data management schemes that emphasize retrieval efficiency. ORION's technique of storing records in alphabetical order by main entry at all times requires more computer resources to add new records than some other methods, but it dramatically speeds retrieval and lowers the cost of presenting users with useful search results. Even now, when the UCLA Library is engaged in large retrospective conversion projects, the ratio of searches to database additions clearly supports the need for database management techniques that allow for maximum efficiency in record retrieval. More than 12,000 new bibliographic records are currently added to the ORION database each week. By contrast, more than 115,000 keyword searches are processed weekly; each of these searches may potentially retrieve several or many records. When retrospective conversion is complete, the number of records added to the database will drop by 75 percent, and the ratio of searches to database additions will be even higher.

In this paper we will outline the principles underlying the design of the ORION database management system. We will also attempt to describe how those principles have been implemented to serve the needs of a large library system and how ORION's data management features facilitate the specific functions of an integrated library information system.

CONTEXT FOR SYSTEM DEVELOPMENT

Local development of an automated library system always occurs in a context of requirements and needs as well as perceptions of basic principles that must govern the system. Throughout the ORION development effort, each of the following principles and requirements has influenced the direction of system development and the way in which the various functions and features were implemented.

- *National standards require that the system be MARC-compatible.* ORION stores records in a slightly modified MARC format, and facilities have been developed for import and export of MARC data.
- *The ORION database must accommodate very large files because the library's collections comprise over 5.5 million volumes.* As of January 1987, ORION included 1.7 million bibliographic records and 2.2 million authority records. Over 600,000 cataloging records will be added to the database annually for the next several years as retrospective conversion projects are completed. The system must be able to accommodate this rapid growth without any degradation in performance or excessive maintenance costs.
- *The system must be flexible enough to serve the data storage and retrieval needs for a variety of library collections of widely varying sizes and data types.* ORION supports numerous files for several institutions and is capable of defining groups of files as data "collections" with common characteristics and rules for update and display of data.
- *Retrieval of records from the database must be fast, even for complex searches or those with a large number of results, and*

the records must display in alphabetical order. ORION's database management system helps to achieve this objective: response time is normally well under one second, even during peak-use periods.

- *The data in the system must be as current as possible at all times.* This requirement militates against batch-update procedures. In fact, ORION batch processing is limited to the following: loading data from weekly OCLC archive tapes and from floppy disks containing Library of Congress authority records downloaded from OCLC; file maintenance tasks; and production of such paper products as purchase orders, claim letters, fund reports, and on-demand lists. All record updates, by contrast, are done immediately online.
- *Each bibliographic entity in the database must be represented by a single master bibliographic record, or "unit record," for storing data efficiently and retrieving all copies under the same access points.* The unit record concept applies to each data "collection," so that master bibliographic records are shared within an institution's set of files but not with other collections.
- *Each campus library's technical processing information must be under its own control.* Each library's local data coexists in the same unit record. ORION stores this information with the bibliographic data in repeatable MARC-like fields, tagged in the range 900 through 999. Processing staff can edit only the 9XX fields for their library.
- *Multiple access points to bibliographic records, as well as multiple access methods, must be provided to meet the widely varying requirements of different users of bibliographic information.* ORION's inverted index files provide Boolean keyword searching and a flexible means of combining various access points in a single search. The linked "browse" files provide access via name or subject heading, title phrase, or call number. The alphabetical storage of records also provides an extremely efficient form of exact searching.
- *Users must be able to search an entire library collection or a subset.* ORION allows searches in individual logical files (such as the Serials file), in various default combinations of files, or in any combination of files specified by the user.
- *Authority file records must be linked to bibliographic records to facilitate maintenance of headings under authority control and to allow use of the authority files as tools for retrieving records.* Authority files must be in synchrony with the bibliographic files at all times.

The structures provided for storing, maintaining, and retrieving data can have an enormous impact on the success of a system. ORION's data management scheme has played an important role in satisfying the objectives listed here; the design features of the system are described in detail in the following sections.

DESIGN FEATURES—RECORD STORAGE

Logical Files and Physical Files

ORION's database management system maintains a distinction between logical and physical files. A logical file is a group of records with common characteristics that are kept together for purposes of retrieval, collection identification, or product generation. The UCLA database includes a number of logical files for bibliographic records. The Monograph Catalog, Monograph in-Process, and Serials files are the largest.

Logical files may consist of one or more physical files. This permits the size of logical files to be independent of physical constraints imposed by storage media and maintenance considerations. Currently, the maximum feasible size for a physical file in ORION is a full IBM 3380 single-density disk pack volume of more than 600 megabytes. This provides enough space to store more than 300,000 bibliographic records and their associated index files. However, it is considerably smaller than the largest logical grouping of records in the system; the Monograph Catalog file, for example, includes more than 1.3 million records and is growing rapidly. Therefore, physical files may be grouped together into a logical file that has a single alphabetical filing sequence. The Monograph Catalog is divided into five physical files, each of which contains records from one segment of the alphabet. The relationships of physical to logical files are controlled by a set of dynamically updatable system tables, and new files or groupings of files may be created at any time. The physical/logical file design allows for logical files of essentially unlimited size.

ORION physical files consist of sets of "blocks" of 15,400 characters,[8] each of which occupies one-third of a physical track on an IBM 3380 disk pack. ORION records, whether bibliographic or authority, are stored within these blocks and are managed by ORION programs at the application level. One or more index blocks store information about the first record in each data block. Within blocks, records are stored in a slightly modified MARC format.[9] ORION stores records in their entirety; no fields are replaced by references to other files. Processing fields that are "local" to the various branch libraries are also stored with the bibliographic fields. This promotes efficient retrieval and display, since all the data to be displayed is in one place and need not be reassembled.

Alphabetical Record Order within Logical Files

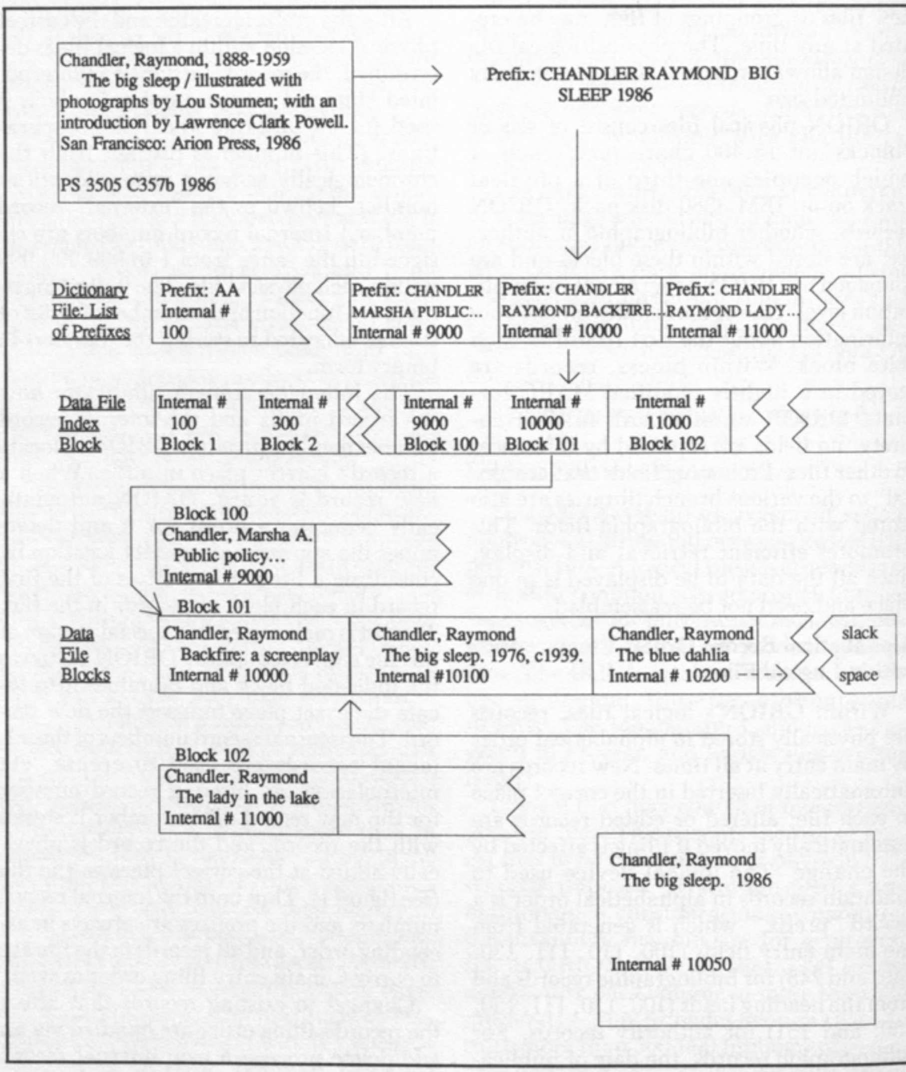
Within ORION's logical files, records are physically stored in alphabetical order by main entry at all times. New records are automatically inserted in the correct place in each file; altered or edited records are automatically moved if filing is affected by the change. The logical device used to maintain records in alphabetical order is a record "prefix," which is generated from the main entry fields (100, 110, 111, 130, 240, and 245) for bibliographic records and from the heading fields (100, 110, 111, 130, 150, and 151) for authority records. For bibliographic records, the date of publication is also included in the prefix in a translated form that allows different editions of a single work to file in reverse chronological order. Special provisions are made for filing numbers and other special data types according to American Library Association filing rules. The prefix, once generated, is physically appended to the front of the MARC record as a varying-length data field. Prefixes rarely exceed 100 characters in length; the maximum allowable length of 350 characters has proved sufficient to provide correct filing for all records encountered thus far.

After the prefix is created and the correct physical location within a logical file is determined, the record is assigned an interpolated "internal record number," which is used for all indexing and related operations. (This number is distinct from the chronologically assigned system identifier number, known as the "external" record number.) Internal record numbers are assigned in the range from 1 to 999,999,999 within each physical file. The 9-digit maximum is a function of the number of bytes of storage allocated to storing the numbers in binary form.

The following scenario illustrates how the record prefix and the internal record number combine to allow ORION to locate a record's correct place in a file. When a new record is added, ORION automatically generates a prefix for it and determines the approximate correct location by consulting a list of the prefixes of the first record in each block of records in the file. This list is maintained in a special section of the file's inverted index. ORION retrieves the indicated block and examines it to locate the exact place to insert the new record. The internal record numbers of the adjacent records are used to create, via interpolation, an internal record number for the new record. This number is stored with the record, and the record is physically added at the correct place in the file (see figure 1). Thus both the internal record numbers and the prefixes are always in ascending order, and all records in the file are in correct, main entry filing order as well.

Changes to existing records that affect the record's filing order are handled via an add/delete process: a new internal record number and prefix are assigned; ORION then adds a copy of the updated record to the file in the new location and deletes the original record from the old location. Changes that do not affect filing simply result in the replacement of the record in its original location.

The arrangement of records by internal record number makes it extremely simple to retrieve records in main entry order for display because it obviates the need to sort or merge them or to refer to an ordered intermediate index. It minimizes the number of read operations required to locate exist-



A prefix is generated from the main entry fields in the new record. The dictionary file stores the prefix of the first record in each block of the data file; comparing the new prefix with this list yields an approximate internal record number. This number is used in turn to locate the data block where the new record should go. An interpolated internal number is assigned, and the new record is physically inserted in alphabetical order. The dictionary file and linked browse files are updated concomitantly.

Fig. 1. Adding a Record.

ing records since their physical locations can be precisely determined without reference to pointers or overflow areas. Since the records that satisfy a search request need not be sorted, it is possible to begin displaying them without reading all of the records from disk first; only enough records to fill the screen need to be read. Finally, because the internal record numbers are used in the inverted index files and are maintained in sequential order there, many simplifying assumptions can be made in the performance of Boolean logic operations. This in turn reduces the computer cycles necessary for retrieving records in a Boolean keyword search.

Dynamic Space Management of Physical Files

ORION leaves a certain amount of empty or "slack" space in each block for the addition of new records. In the most active files, 15 to 20 percent of the space in each block is normally reserved for this purpose. Less active files require less slack space for efficient operation.

If a new record needs to be added, ORION inserts it at the correct place in the block and moves following records in the same block to make room. If a number of records are added to the same block, the block will eventually fill up. In this case, ORION automatically invokes routines that shift records from the ends of the filled block to adjacent blocks, much as cards in a card catalog might be shifted to make more room in a drawer that is completely full. If the adjacent blocks are also full, the system automatically shifts records from as many blocks as necessary in order to make room for the new record in its correct place in the alphabetical sequence. In practice, these routines are seldom invoked.[10]

File Maintenance

Like other database systems, ORION files require that occasional maintenance chores be performed to keep the system operating at peak efficiency. Current cataloging or acquisitions work tends to generate records that are widely scattered through the alphabetical sequence. ORION's data storage scheme, which distributes unused storage space throughout files, is ideally suited for this sort of input. However, retrospective conversion projects or the cataloging of a number of publications with the same entry may cause a large number of records to be added to a small portion of the database. When this happens, one or both of the following bottlenecks may develop.

First, as blocks in the affected area of the file fill up, the number of dynamic disk space shifts required to add new records in that area increases. Disk space shifts are monitored, and ORION produces a daily report that alerts systems staff to run a batch job to "space out" the records in the affected physical files, redistributing the slack space.

Second, the process of interpolation by which internal record numbers are assigned may lead to a situation where there is no available number between two existing records. To illustrate: if two adjacent records have internal numbers 1,000 and 3,000, a new record inserted between them may be assigned internal record number 2,000; a second new record may be assigned 1,500; etc. The interval between the numbers can only be divided in half a finite number of times before adjacent numbers are assigned. In fact, the interpolation algorithm does not merely divide the available interval in half, but attempts to assign a number to one side or the other of the midpoint. Nonetheless, internal record number blockages can occasionally happen. When they occur, an online command can be used to reassign the internal numbers in the affected range. Because the internal record numbers serve as the links from the inverted dictionary files to the bibliographic records, affected dictionary entries are adjusted at the same time so that the data files and their dictionaries remain synchronized. Preventive measures are taken by periodically running an overnight batch job that reassigns all of the internal numbers in a physical file and rebuilds the file's dictionary from scratch. Disk space may be redistributed at the same time.

In practice, file maintenance absorbs a relatively small part of system resources, even though ORION is currently growing at the rate of more than 12,000 new records each week. Monitoring the files and submitting and logging the batch file maintenance jobs for more than forty active ORION files require staff time of three to four hours a week. File maintenance in general consumes 2 to 3 percent of the machine resources used by the system as a whole. Because each physical file is a separate entity for maintenance purposes, maintenance can be accomplished in small segments and scheduled on a rotating basis. All batch maintenance jobs are processed after midnight, so users are essentially unaffected by the unavailability of files on which maintenance jobs are running.

DESIGN FEATURES—RECORD RETRIEVAL

Inverted File Structure

ORION uses inverted "dictionary" files to provide access for record retrieval. Numerous keywords and numbers in each record are indexed, including, for bibliographic records, all terms from the author, title, series, added name, and subject heading fields, as well as the permanently assigned ORION external record number, Library of Congress card number, ISBN or ISSN, OCLC control number, and other numbers—for example, invoice, GPO, or publisher's serial number.

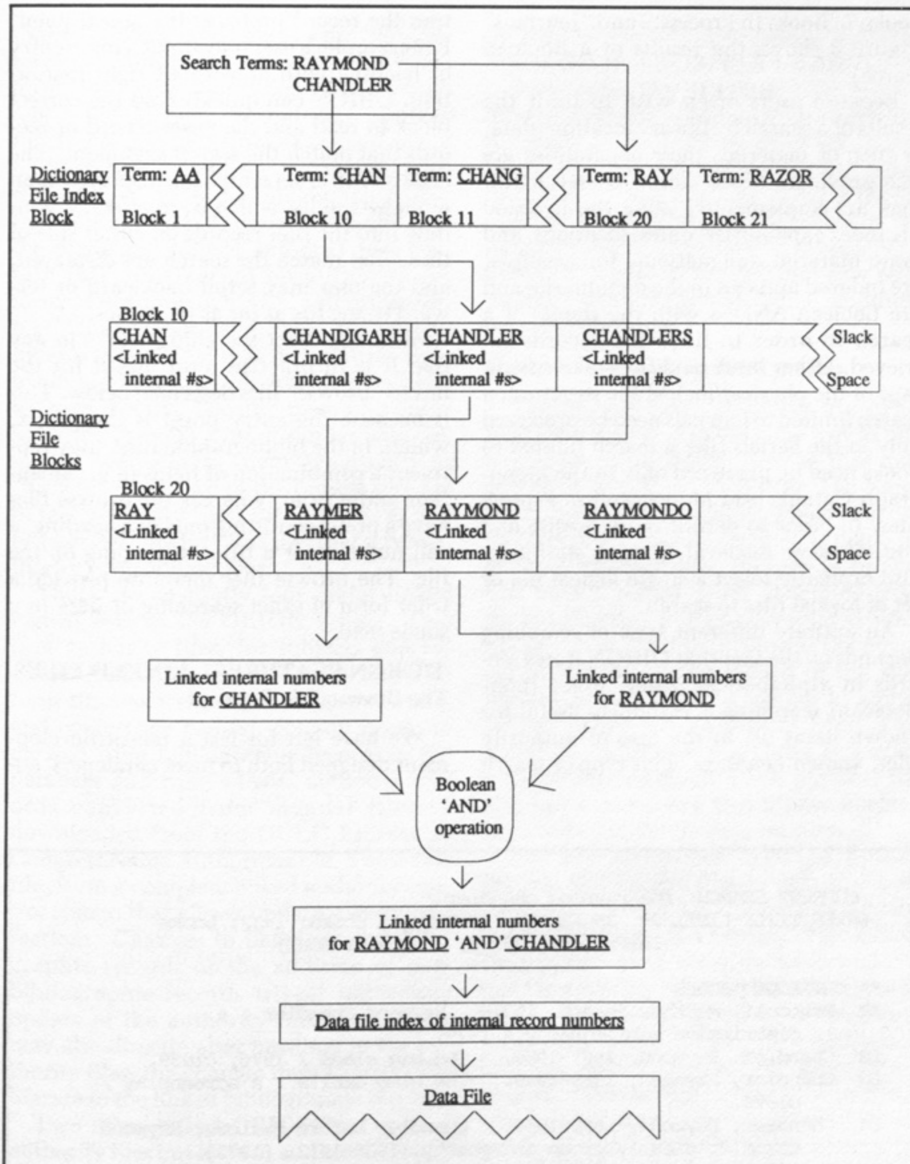
Each physical ORION file has an associated inverted dictionary file that includes all indexed terms from records in that file. The different types of indexes are distinguished by a 1- or 2-character index prefix that identifies the type of index and causes indexes of the same type to fall together within the inverted file. Defining and implementing a new index type requires only that a new index prefix be specified and that the programs be instructed which subfield, field, or fields to index under the new type.

Dictionaries, like bibliographic data files, are composed of 15,400 character-keyed blocks. Each dictionary contains one or more blocks that store an index indicating the first term in each of that dictionary's data blocks in order to facilitate rapid location of a term. Each data block in the dictionary contains one or more terms and the associated links to the bibliographic records. Terms may span several adjacent blocks, so that there is no theoretical limit to the number of links that each term may have. As in the data files, ORION leaves slack space in each dictionary block for the addition of new terms or new links to existing terms. When a term or link is inserted, data following in the same block is moved to the right to accommodate the new entry. If a block becomes full, ORION automatically shifts data to adjacent tracks to create more space, just as in data files. Also like the bibliographic data files, the dictionary files are subject to the need for maintenance. Disk space shifts are monitored, and the files can be spaced out periodically via a batch process.

The ORION internal record number, which determines a record's place in a physical file, serves as the link from entries in the dictionary back to the records in the data files. The links for each term are stored in numerical order. Therefore, the records that the links represent can be easily retrieved in alphabetical order. This scheme also tends to simplify Boolean operations on combinations of terms because the links are known to be in order at all times. The disadvantage of this method is that when internal record numbers change, as they may if numbers need to be reassigned to restore a reasonable interval between adjacent entries, all dictionary links to the affected records must also be changed. In practice, this is largely accomplished by means of preventive, overnight batch maintenance jobs that reconstruct the entire inverted file when a data file is completely renumbered. This kind of process consumes about half of the resources devoted to file maintenance.

Retrieval Capabilities

ORION's dictionaries provide full Boolean keyword access (see figure 2). Each linked record number in the dictionary has an 8-bit "mask" that shows how the term is used in the record. The mask allows searching by specific index types, such as names, titles, subjects, or series; or by defined combinations of types—for example, "find nt dickens" to search terms in either name or title fields. Index types may also be combined with logical operators—for example, "find name Commager AND find subject history." Logical AND, OR, and AND NOT operations, term truncation, and "wild card" characters are supported. It is possible to search more than one dictionary, and therefore more than one physical or logical file, at a time. When such a cross-file search is performed, each dictionary is searched in turn, and all Boolean operations are performed before going on to the next dictionary; results are cumulated dictionary by dictionary until the search is complete, whereupon the results may be displayed. ORION's public users normally search the Monograph Catalog, Monograph in-Process, and Serials files. These



The dictionary file index block directs the program to the dictionary blocks where the terms *Raymond* and *Chandler* reside. The linked internal numbers for each term are collected and a Boolean AND operation is performed. The result is a list of internal numbers satisfying the query. The internal numbers are then used to retrieve the records in the data file.

Fig. 2. Boolean Search Operation (Find Name Search, Boolean ANDing Raymond *and* Chandler).

three logical files represent eight physical files and include more than 1.7 million bibliographic records. A typical name/title search across these files takes less than one second. Because three logical files are being searched, results of these Boolean searches are displayed in three alphabetical sequences, which are labeled "Cataloged

Books," "Books In Process," and "Journals" (figure 3 shows the results of a Boolean search).

Because users often wish to limit the results of a search by library location, date, or type of material, these capabilities are also provided. Some "limit" search modifiers are implemented using the inverted file index capabilities: dates, locations, and some material designations, for example, are indexed as terms in the dictionaries and are Boolean ANDed with the results of a search in order to limit the records retrieved. Other limit modifiers take advantage of the physical/logical file structure: a search limited to journals need be processed only in the Serials file; a search limited to books need be processed only in the Monograph Catalog and Monograph in-Process files. In order to permit more precise and cost-effective retrieval, library staff may also explicitly select a single logical file or set of logical files to search.

An entirely different type of searching depends on the fact that ORION stores records in alphabetical order: exact (non-Boolean) searching is extremely useful for known items or, in the case of authority files, known headings. This type of search uses the record prefix as the access point. For example, a user may input a main entry or heading with or without right truncation. ORION can quickly find the correct block to read and the exact record or records that match the search argument. The results of an exact search are displayed as an alphabetical list which is, in effect, a window into the file: records on either side of those that match the search are displayed, and the user may scroll backward or forward in the file as far as he wishes.

Although exact searching works in any file, it is in practice most useful for the linked "browse" files described below. This is because the entry point is the prefix, which, in the bibliographic files, may represent a combination of fields (e.g., an author and a title), whereas the browse files have a prefix consisting only of a heading, a call number, or a title, depending on the file. The browse files therefore provide a truer form of exact searching of data in a single field.

DESIGN FEATURES—LINKED FILES

The Browse File Concept

We have left for last a recent development designed both to meet catalogers' au-

```
  CURRENT SEARCH: fna raymond chandler
-BRIEF TITLE LISTING-  35 RESULTS (Cataloged Books: 1-32; Books
                       In-Process: 33-35; Journals: NONE)

** CATALOGED BOOKS
D1  Bruccoli, Matthew Joseph, 1931-   Raymond Chandler : a
      descriptive bibliography / 1979.
D2  Chandler, Raymond, 1888-1959.   The big sleep / 1976, c1939.
D3  Chandler, Raymond, 1888-1959.   The blue dahlia : a screenplay /
      c1976.
D4  Chandler, Raymond, 1888-1959.   Chandler before Marlowe; Raymond
      Chandler's early prose and poetry, 1908-1912. [1973]
D5  Chandler, Raymond, 1888-1959.   [Correspondence. Selections]
      1981. Selected letters of Raymond Chandler.
D6  Chandler, Raymond, 1888-1959.   Farewell, my lovely / 1976, c1940.

*OPTIONS:  -TYPE  D1   (or D2,D3...) TO SEE ALL OF A SPECIFIC RECORD.
           -TYPE  UP   TO SCROLL UP THE LIST; PRESS ENTER TO SCROLL DOWN.
           -TYPE  D1+  (OR D2+, D3+...) TO BEGIN THIS LIST AT A SPECIFIC
                       NUMBER.
           -BEGIN A NEW SEARCH (E.G. FNT .....)
```

Fig. 3. Public Brief Title Display (Find Name Search, Boolean ANDing Raymond *and* Chandler).

thority control needs and to provide new retrieval options for all users. ORION's linked browse/authority control files represent a hybrid type, in that they combine functional and structural characteristics of bibliographic data files and of inverted dictionary files. Like the bibliographic files, they contain data of immediate usefulness: the name and subject authority files contain full information on authoritative forms of headings in MARC authorities format. On the other hand, these files serve as indexes in the sense that they contain links to bibliographic records and provide access to those records. Because these files allow browsing through a list of similar data elements, we refer to them generically as "browse" files. Two of the browse files are in fact authority files; the latter term will be reserved here for those files that serve authority control functions.

There are currently four major linked logical files in the ORION system. Two true authority files, for subjects and for names (the latter includes series and uniform titles as well as personal, corporate, and conference names), contain headings derived from bibliographic records in the database and from MARC authority records converted from manual files or downloaded from the OCLC Library of Congress Name Authorities file. These two files form a complete linked authority control system that allows updating in two directions. Changes to headings in bibliographic records or the addition of new bibliographic records trigger immediate update of the authority files. Catalogers may also directly alter headings in the authority files; the changes immediately proliferate to the linked bibliographic records.

Two other linked ORION files are not authority files but serve as additional access devices for the bibliographic files. These are the Title Browse file, which currently includes only serial titles, and the Call Number file. These files are not directly updatable by staff but automatically reflect changes in the bibliographic files. They allow browsing by title and by call number across the various logical files to which they are linked. The Call Number file allows simulated shelf browsing by users and provides the exact call number access necessary for the ORION circulation/inventory control module that is being developed.

Browse files may be linked to one or more bibliographic files. For example, the Name Authority file is linked to all of the UCLA Library's major bibliographic files, while the Title Browse file is linked only to the Serials file. The linkages are controlled by dynamically updatable system tables. New browse files may be created and new relationships defined through online modification of these tables.

Data Management in Browse Files

The browse files themselves are structured and managed in the same way as bibliographic files; logical browse files may consist of several physical files. The Name Authority file, for example, which contains more than 1.6 million records, is made up of three physical files. Records in browse files have a record prefix generated from the heading field and an interpolated internal record number that controls the physical placement of the record in the file. Browse records are thus always in alphabetical order within logical files. This structure facilitates searching the exact form of the headings. Each physical file also has a dictionary that allows Boolean keyword access to indexed terms.

The key difference between bibliographic and browse files in data management terms is that browse records contain links to bibliographic records. This allows them to serve as an access device as well as a tool for authority work. Each link contains the physical file name and the external record number of the linked bibliographic record. Internal record numbers are not used as links because they change over time, whereas the external numbers are permanently assigned. Using the external number as the link means that search routines must include looking up each external record number in the dictionary of the bibliographic file in order to find the corresponding internal record number and retrieve the record; however, it also means that the browse files need not be modified when internal numbers change.

The concept of maintaining alphabetical sequence at all times has been extended to the links within each browse record.

ORION maintains the links in alphabetical order across all logical files to which the browse file is linked. Thus, all bibliographic records retrieved via a browse search can be displayed in alphabetical order without sorting, and a single sequence, rather than a separate sequence for each linked logical file, can be presented. Ordering the links is facilitated by storing the first character in the prefix of the linked bibliographic record as part of the link in the browse record. Because the vast majority of browse records have only one or two links, one character is sufficient for correct sorting most of the time.[11] When it is not sufficient, ORION checks the bibliographic files themselves in order to determine where in the sequence of links a new one should be added. For the few records with a large number of links, this procedure—adding a new link—incurs a substantial amount of computer resource "overhead." This drawback is relatively minor, however, when weighed against the advantages of this scheme for record retrieval.

The implementation of dynamically linked browse files has added a new level of complexity to ORION's database management functions. A complex web of changes, all of which are accomplished immediately online, may result from a single change to an ORION record. An update to a bibliographic record, or the addition of a new one, will cause updates to the linked browse files. Updates to the authority files proliferate changes into the linked bibliographic records. If an authority file update changes a main entry field in a bibliographic record in such a way as to affect filing, the bibliographic record will move to a new place in the physical file or even to a new physical file. The addition of a new authority record with cross-references will cause the creation of appropriate cross-reference records, and a change to an authority record will also change the associated cross-reference records. If a browse record changes in such a way that it becomes identical to an existing record, the two records automatically merge and their links combine. In all cases, dictionary files are updated concomitantly so that they always correspond to the data in the browse and bibliographic files.

Many trade-offs had to be weighed carefully as this design evolved. The final choices were governed by the overall impact on computer resource and disk space usage, desired functionality, maintainability, and user expectations. For example, authority control via the authority files is fairly resource-intensive in terms of computer cycles and read-write operations because ORION stores headings both in bibliographic records and in the browse files. If a heading is changed in the authority file, the linked bibliographic records must be physically retrieved, modified, and rewritten. This contrasts sharply with systems that store headings in linked authority files only, the bibliographic records having a numeric reference to the authority file heading in place of the actual data. The latter scheme requires only a single update to change the form of a heading in all bibliographic records. However, such a scheme requires many more "read" operations to display bibliographic records than are necessary in ORION, because records must be reconstructed from scattered parts.

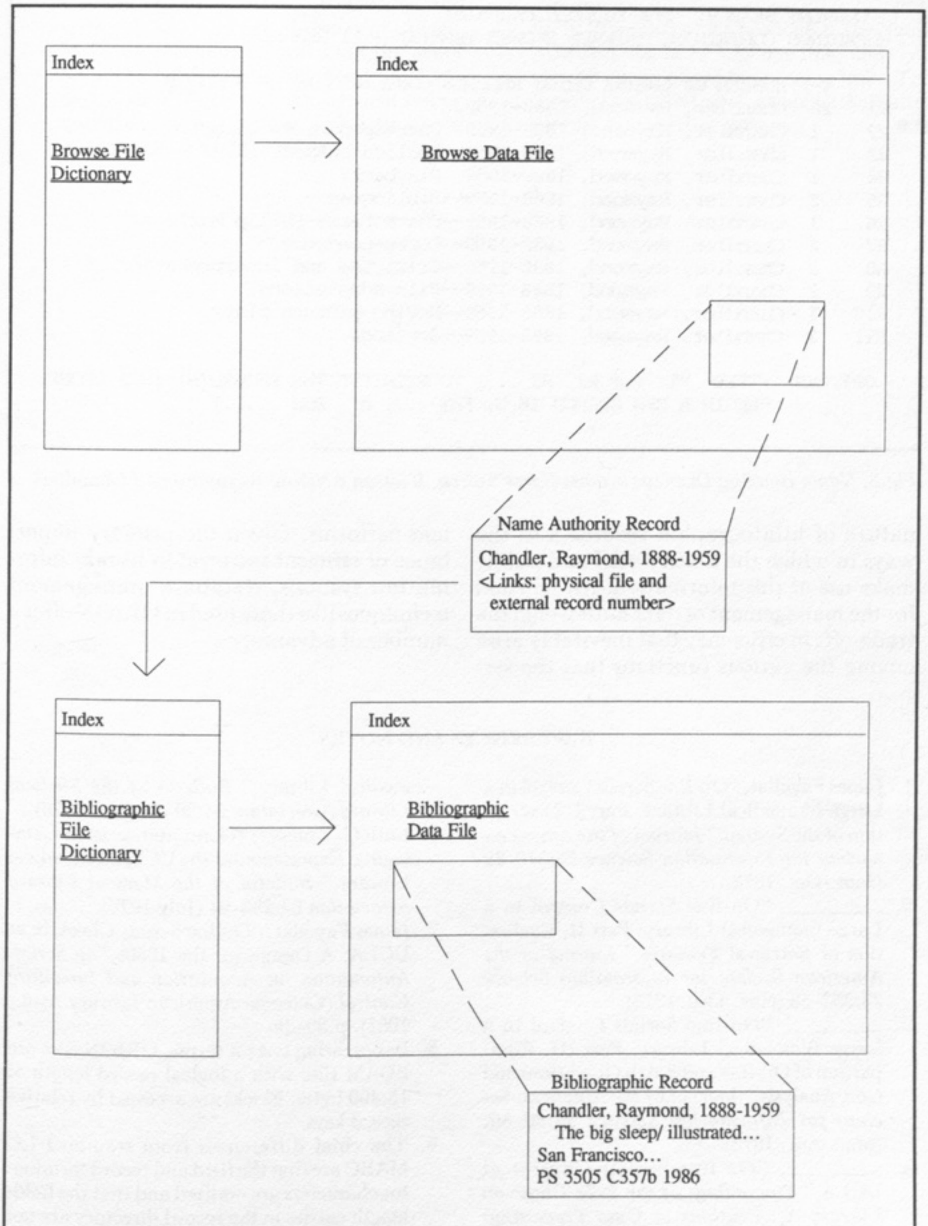
Searching in Browse/Authority Files

The Browse/Authority files allow both of the types of searching described above for bibliographic files. Boolean keyword searching, implemented for the Name and Subject browse files, results in the display of a list of headings that satisfy the search (see figures 4 and 5). The user may then issue a simple command to retrieve the bibliographic records associated with any of the headings.

Exact searching, which is possible in all of the linked files and is the only type of searching implemented for the Call Number file, results in display of a list that represents a window into the browse file. Users can scroll forward or backward in the list and thus in the file. Again, a simple command retrieves the bibliographic records associated with any listed heading, title, or call number.

CONCLUSION

ORION is a powerful and complex system for the storage, maintenance, and retrieval of bibliographic information. Its fundamental design is driven by the unique



A Boolean search in a browse file locates one or more browse/authority records. The browse/authority records contain links to bibliographic records. Each link consists of the physical file name and external number of the linked record. The external numbers are looked up in the bibliographic file dictionaries in order to retrieve the bibliographic records.

Fig. 4. Browse Search and Retrieve (Browse Name Search, Boolean ANDing Raymond *and* Chandler).

```
  CURRENT SEARCH:  bna raymond chandler
 -HEADINGS CONTAINING ENTERED BROWSE TERM(S) - 11 RESULTS

      <-> NUMBER OF ONLINE ORION RECORDS CONTAINED IN EACH GROUP
 R1   28  Chandler, Raymond, 1888-1959.
 R2    1  Chandler, Raymond, 1888-1959. Correspondence. Selections
 R3    1  Chandler, Raymond, 1888-1959. English summer. 1976.
 R4    1  Chandler, Raymond, 1888-1959. Playback.
 R5    2  Chandler, Raymond, 1888-1959--Bibliography.
 R6    1  Chandler, Raymond, 1888-1959--Characters--Philip Marlowe.
 R7    2  Chandler, Raymond, 1888-1959--Correspondence.
 R8    2  Chandler, Raymond, 1888-1959--Criticism and interpretation.
 R9    2  Chandler, Raymond, 1888-1959--Film adaptations.
 R10   1  Chandler, Raymond, 1888-1959--Moving-picture plays.
 R11   1  Chandler, Raymond, 1888-1959--Settings.

  OPTIONS  -TYPE R1  (or R2, R3 ...) TO RETRIEVE THE RECORD(S) IN A GROUP.
           -BEGIN A NEW SEARCH (E.G. FNT ..., or  BSU .....)
```

Fig.5. Name Heading Display (Browse Name Search, Boolean ANDing Raymond *and* Chandler).

nature of bibliographic records and the ways in which the library staff and public make use of this information. All schemes for the management of data must weigh the trade-offs in efficiency that inevitably arise among the various functions that the system performs. Given the primary importance of efficient retrieval in library information systems, database management techniques like those used in ORION offer a number of advantages.

REFERENCES AND NOTES

1. James Fayollat, "On-line Serials Control in a Large Biomedical Library: Part I, Description of the System," *Journal of the American Society for Information Science* 23:318–22 (Sept.-Oct. 1972).
2. ________, "On-line Serials Control in a Large Biomedical Library: Part II, Evaluation of Retrieval Features," *Journal of the American Society for Information Science* 23:353–58 (Nov.-Dec. 1972).
3. ________, "On-line Serials Control in a Large Biomedical Library: Part III, Comparison of On-line and Batch Operations and Cost Analysis," *Journal of the American Society for Information Science* 24:80–86, (Mar.-Apr. 1973).
4. ________, "On-line Serials Control at UCLA," *Proceedings of the 1972 Clinic on Library Applications of Data Processing: Applications of On-line Computers to Library Problems* (London: Clive Bingley, 1972), p.69–81.
5. Louise Darling and James Fayollat, "Evolution of a Processing System in a Large Biomedical Library," *Bulletin of the Medical Library Association* 64:20–24 (Jan. 1976).
6. Ruth C. Traister, "Computer-assisted Cataloging: Experiences at the UCLA Biomedical Library," *Bulletin of the Medical Library Association* 63:283–94 (July 1975).
7. James Fayollat, "Online Serials Check-in at UCLA: A Design for the 1980s," in *Serials Automation for Acquisition and Inventory Control* (Chicago: American Library Assn., 1981), p.20–39.
8. In operating system terms, ORION files are BDAM files with a logical record length of 15,400 bytes. Blocks are accessed by relative record keys.
9. The chief differences from standard LC MARC are that the field and record terminator characters are omitted and that the field-length entries in the record directory are not stored because they can be easily calculated when needed. In their internal format, ORION records also have a 40-character record leader that carries more information than the MARC leader.

10. Readers may see certain similarities between ORION'S database management system and IBM's VSAM (Virtual Storage Access Method). ORION achieves all of the major features of the KSDS (Key Sequence Data Set) mode of VSAM: (1) both random and sequential access, (2) variable-length record storage and update, (3) allocation of free space and regaining of space, and (4) key compression.
11. Fully 70 percent of the headings in the Name Authority file have only one link, and only about 2 percent have ten or more links. The distribution of links in the Subject Authority file follows a similar pattern. ■■

Search Output Enhancement: The 3M Experience

Lynda B. M. Ellis and Kristin K. Oberts

Information Services staff at the Minnesota Mining & Manufacturing Company (3M), headquartered in St. Paul, Minnesota, access a large number of information resources and report the results of their widely ranging searches in uniform, yet flexible formats. Two computer-based tools have been developed for the enhancement of the results of search requests. COVER assists in production of search-result summaries; REFORM assists in editing (including resequencing, deleting, and reformatting) downloaded results from online searches. COVER and REFORM have been used since May 1985 by 3M information services personnel, and COVER has been successfully integrated into daily operation in twelve 3M libraries/information centers in the U.S. and Canada. This report describes the development, implementation, and use of these important tools.

Information services are provided to 3M personnel by a centrally administered staff of about sixty, of whom half have professional degrees. This staff (located, unless otherwise indicated, in St. Paul, Minnesota) is subdivided as follows: seven technical libraries; a business library; an engineering library; a current awareness service; a patent and internal information service; a library technical service group; a computer support group; and a library in Austin, Texas. All of these subdivisions access a large number of information resources in order to respond to requests from users.

The 3M information professional, like all others in this field, must respond to requests for information in a timely manner and with results that are accurate, acceptably complete, and packaged in a form that is intelligible to the requester. In addition, the results should be documented in a form that permits easy updating: if they are electronically downloaded, it should be possible to edit them to meet requester's needs and to print them in a well-formatted report.

Over 13,100 public electronic databases are available from database vendors such as Dialog, STN, Orbit, BRS, etc. Searches run against these databases can potentially produce voluminous bibliographic and nonbibliographic information on virtually any topic. Commercial software can allow a user to download search results from many of these databases (for example, see reference 1).

Our goal was to develop computer-based tools that would give search staff the ability to describe the search to the requester easily. We developed COVER, a package which collects information about a search that includes topic; methodology; sources used, with a description of each; time period for searching each source; guide to the organization of search results, and information on obtaining documents cited in the results. A second, related goal was to en-

Lynda B. M. Ellis is Associate Professor, Division of Health Computer Sciences, Department of Laboratory Medicine and Pathology, University of Minnesota–Minneapolis, and Kristin K. Oberts is Manager of Technical Libraries and Information Resources at 3M corporate headquarters, St. Paul, Minnesota.

hance the quality of online search results. We developed REFORM, a package that includes the capability to merge searches from different databases and vendors; reorder the searches according to the user's requirements; eliminate irrelevant or duplicate citations; and reformat the final report with headings and pagination.

METHODS

COVER and REFORM were developed using the Advanced BASIC 3.0 interpreter on a COMPAQ portable computer. Prior to use in a production setting, COVER is compressed using the BASIC Development System (Betatool Systems), and REFORM is compiled using version 2.0 of the IBM BASIC compiler. COVER and REFORM are used on IBM PC or XT machines with clock-calendar card, monochrome display, and dot matrix printer and with letter-quality printers available when needed. Computers used for online searching also have 1200-baud modems and PC-PLOT communications software (Microplot Systems).

The COVER program collects information and passes it to a word-processing package for final printing or storage. The WordStar/MailMerge, Displaywriter 2, and Microsoft WORD packages were evaluated for use with COVER. All had deficiencies; however, the deficiencies of WordStar were less than those of the other two, and WordStar/MailMerge (MicroPro International) is presently used with COVER.

Planning for the project began in late 1984; systems analysis and programming were started in January 1985. All development, training, and evaluation described below were done in 1985.

Both programs were developed using a rapid-prototype methodology. Two committees, one for each package, were appointed. Each was composed of searchers, search staff management, and systems staff; each reviewed its assigned prototype as it was developed, gave immediate feedback as to its deficiencies, and proposed and assigned priorities to enhancements.

The seventh prototype of COVER and the third prototype of REFORM were considered useful enough for more general distribution. The first general release of COVER (version 1.7) was previewed at the largest 3M technical library for two weeks in April; then three training sessions on both packages were held in early May for users in St. Paul. Version 1.7 of COVER and version 1.3 of REFORM, including complete user manuals, were distributed to all library and patent searchers at that time. A training letter produced by COVER, which was used to demonstrate its features, is shown in figure 1.

A COVER-user questionnaire was distributed in late May, and a final COVER committee meeting was held in early June. The second general release (version 2.0) of COVER, revised on the basis of initial use, was distributed in July, accompanied by a user's manual. The second release changed the way long text was input and added an additional optional comment field and batch mode.

A COVER MANAGER program was developed in August to assist in maintaining the more than 200 standard paragraphs associated with COVER. This program interfaces with WordStar to permit changing old, or creating new, paragraphs, and it can also delete paragraphs.

The last REFORM committee meeting in this development cycle was held in August, and the second release of REFORM (version 1.5), revised on the basis of initial use and including a user's manual, was ready for distribution in November.

RESULTS AND DISCUSSION

Search Results Summary

The COVER program collects information about a search: searcher's name, address, and phone number; requester's name and address; title; date; topic; search method; method of organizing results; sources used; time period for searching each source; instructions on obtaining complete documents; and optional comments.

Names and address are entered once for each searcher who will be using COVER at a particular location and, from then on, are selected from a list of searchers' first names. Source descriptions and instructions for obtaining documents are stored as standard paragraphs that are selected from similar lists of paragraph names. These choices can be automatically repeated for each letter in

3M INFORMATION SERVICES SEARCH RESULTS

TO: JOHN DOE * DIVISION * MAIL CODE
FROM: LYNDA B.M. ELLIS * ISSS * 201-2C-12
TITLE: THIS IS A ONE-LINE TITLE
DATE: July 16, 1985

SEARCH TOPIC: This is a search topic. Like the search title, it is entered in upper and lower case and is limited in size. Unlike the search title, it will retain the case it is entered in and is limited to up to one full screen (24 lines) of information, or 32 lines when printed.

SEARCH METHOD: Unlike the previous fields, this one is optional and can be omitted by typing RETURN alone. It can have the same screenful of information as search topic.

ORGANIZATION OF RESULTS: This field is also optional and is limited to one screen of information.

SOURCES (AND YEARS) INVESTIGATED:
CA SEARCH (March 1, 1984 through January 15, 1985) corresponds to the printed ***Chemical Abstracts***. It covers journal articles, patents, reviews, technical reports, books, conference proceedings, and dissertations in all areas of chemistry and applied chemistry.
The Name One Source (Third Edition, 1984) This description, like search topic, can be up to 24 lines (one screen) long.

HOW TO OBTAIN DOCUMENTS:
Copies of **journal articles** or **conference papers** cited in this search may be obtained by completing a "Copy Service Request" form, Form 7640-I, for each requested item. If you wish, you may photocopy the citation as it appears in your literature search and tape it to the top of the form. Copying of articles which are available in the 3M libraries is done at no charge. Your department will be assessed a service charge for articles which must be ordered from outside 3M. If you have any questions about the local availability of articles or this ordering process, please contact me.

These comments are optional and are limited to one full screen.

If you have any questions about this search, please call me.

LYNDA B.M. ELLIS

Fig. 1. Training letter produced by COVER. The CA SEARCH paragraph under "Sources Investigated" and the "Copies of" paragraph under "How to Obtain Documents" are stored paragraphs.

a batch. Each letter may differ in search title, topic, method, organization of results, time period searched, and comments; but each is identical in searcher, date, sources used, and instructions on obtaining documents.

Date is obtained from the clock-calendar. Requester name and address and search title are entered from the keyboard for each letter; search topic and method, organization of results, and comments are also entered from the keyboard. These fields differ from the previous group in that they can contain up to one full screen (24 80-column lines) of information. Other fields are limited to 256 characters. An example of a screen display for COVER is shown in figure 2.

COVER permits the user to enter, change, and verify the information. It contains online help messages and displays, and after the final verification, users can choose to print the letter immediately, save it for later printing, cancel it without saving or printing, or use it to start a batch of letters. When finished entering the second and subsequent letters in a batch, the user can continue to enter another letter, print the batch, save the batch, or cancel the batch.

When a batch or letter is printed or saved, the information collected by

```
Version 2.1                                          (c) 1985 3M Company
                        COVER LETTER MAIN MENU
Type ? for Help                                       Type ^ to Backup

              1  Produce a Cover Letter

              2  Add or Delete Searcher Info

              3  Print a Previously Saved Letter

              4  Instructions

              5  Quit

              Which do you want?
```

Fig. 2. Initial screen display for COVER.

COVER is passed to the WordStar/MailMerge word-processing package. This package assembles the final letter(s) and either prints them or saves them to disk. Saved letters can be printed directly or edited using WordStar prior to printing. A letter produced by COVER using MailMerge is shown in figure 3. (This letter summarizes the search results shown in figure 6).

COVER MANAGER is a program that assists library management in maintaining the more than 200 standard paragraphs used by COVER. It lists the names, which are displayed for each paragraph in COVER (the display names), and permits a name (and paragraph) to be deleted or added or the text of a paragraph to be changed.

The paragraphs to be added may have been created outside of COVER MANAGER, using WordStar, or can be created through COVER MANAGER's WordStar interface. Paragraphs are edited through this same interface.

If any paragraphs have been added or deleted at the end of a session, COVER MANAGER instructs the user on production of a new list of all standard paragraphs using MailMerge.

COVER Evaluation

COVER was evaluated by four experienced search librarians in 3M's largest technical library. For each use the following data were collected: date; start time; total time for producing one letter; parts corrected during verification, if any; and final disposition of the letter (print, save, cancel, batch). These data were collected for three time periods: prerelease (two weeks in April); first release (May–July); and second release (August–November).

The time spent producing one letter is represented in figure 4. During the first two weeks of COVER use, the median time was 8.6 minutes (from start of COVER to the time when the decision to print, save, or cancel the letter was made). This decreased to 6.6 minutes during the next three months. Indeed, the decrease was even faster than this implies: the median for the first week of use was 9.1 minutes, which decreased to 6.9 minutes for the second week. There was a slight increase to 7.3 minutes over the last four months of use.

The disposition of the letters produced during the test periods is summarized in table 1. From 75 to 90% of the letters produced were immediately printed; most of the remainder were saved for later printing; only 1 to 2% were canceled or used to start a batch of letters.

The number of parts of a letter that were corrected during verification is summarized in table 2. From 70 to 80% of the letters had no corrections during verification; 13–24% had one part corrected; and

3M INFORMATION SERVICES SEARCH RESULTS

TO: 3M INFORMATION USER * 3M COMPANY * 3M LOCATION
FROM: 3M INFORMATION SERVICES STAFF MEMBER * 201 TECH LIB * 201-2S-00
TITLE: LIQUID CRYSTAL POLYMERS
DATE: August 20, 1986

SEARCH TOPIC: Enclosed is the information you requested on the topic of liquid crystalline polymers. The topic included any type of liquid crystal: smectic, nematic, cholesteric, and others.

SEARCH METHOD: The search was carried out using two computerized databases, which are described below. The search terms included all variants of the term 'liquid crystal polymer', as well as controlled index terms assigned by the different databases and terms for specific kinds of liquid crystals.

ORGANIZATION OF RESULTS: The printout is enclosed. It has been organized into alphabetical order by the name of the first author or inventor. 77 references were retrieved in the search and duplicate citations have been removed.

SOURCES (AND YEARS) INVESTIGATED:
CA FILE (1980-date) corresponds to the printed ***Chemical Abstracts***. It covers journal articles, patents, reviews, technical reports, books, conference proceedings, and dissertations in all areas of chemistry and applied chemistry.
NTIS (1980-date) is a database of government sponsored research, development, and engineering reports and studies done by government agencies, their contractors or grantees. Subjects covered include: physical sciences, technology, engineering, biological sciences, medicine, agriculture, and social sciences.

HOW TO OBTAIN DOCUMENTS:
Copies of **journal articles** or **conference papers** cited in this search may be obtained by completing a "Copy Service Request" form, Form 7640-I, for each requested item. If you wish, you may photocopy the citation as it appears in your literature search and tape it to the top of the form. Copying of articles which are available in the 3M libraries is done at no charge. Your department will be assessed a service charge for articles which must be ordered from outside 3M. If you have any questions about the local availability of articles or this ordering process, please contact me.
Government document publications of the United States cited in this search may be obtained by completing a "Government Document Request" form, Form 8155-F, for each requested item.

If you have any questions about this search, please call me.

3M INFORMATION SERVICES STAFF MEMBER

Fig. 3. Example letter produced by COVER for the search results shown in figure 6. The CA FILE paragraph and two paragraphs under "How to Obtain Documents" are stored paragraphs.

5–6%, two or three parts corrected. This is a lower estimate of the number of errors corrected: the user can "back up" at any time to correct a mistake prior to verification and, once in verification, can change one part repeatedly if desired.

The frequency with which a particular part was corrected in all letters with at least one correction is summarized in table 3. Search topic, 24% of all corrections, was the part most frequently corrected.

COVER's evaluation is informative for the implementation of such a tool. No more than two weeks were required for searchers to adjust to its use and for the median time for letter production to reach its minimum.

The slight increase in median time seen in the last four months of use (second release, figure 4) may be real. The second release added one additional comment field. Even if this field were not used, the computer would display its input screen and the user would need to read the screen and press one additional key to omit it. It would also appear (and require a key press to remove) during verification.

Table 1. Disposition of Letters Produced by COVER.

Disposition	Prerelease N	Prerelease (%)	First N	First (%)	Second N	Second (%)
Print	41	(79)	117	(77)	199	(88)
Save	10	(19)	33	(22)	25	(11)
Cancel	1	(2)	1	(1)	0	(0)
Batch	NA		NA		3	(1)
Total	52		151		227	

Table 2. Number of Parts Corrected during Verification.

Number	Prerelease N	Prerelease (%)	First N	First (%)	Second N	Second (%)
0	37	(71)	124	(82)	159	(70)
1	12	(23)	20	(13)	54	(24)
2	1	(2)	7	(5)	10	(4)
3	2	(4)	0		4	(2)
	52		151		227	

Table 3. Frequency of Correction during Verification by Part.

Part*	N	(%)
*Searcher/ Requester	8	(6)
*Title	16	(12)
*Topic	31	(24)
Method	23	(17)
Organization	23	(17)
*Sources/ Time Period	19	(14)
Obtain	12	(9)
	132†	

* = required

†An additional 9 corrections were made to the "optional comments" part added in the second release. This is comparable to the correction rate for title.

It is most unlikely that the longest 10% of time periods are due to extremely long or problematic letters: COVER is used by very busy searchers who are subject to interruptions that may require them to leave in the middle of composing a letter and return later. COVER has no way to distinguish these interruptions from normal pauses.

Another measure of COVER's success is the percentage of letters that are printed directly versus the number that are saved for later printing. While some of the latter are saved for printing on a different printer or because the user doesn't have time to print, some undoubtedly are saved for further editing because of perceived inadequacies in the COVER product. During the last four months of COVER evaluation almost 90% of the letters were directly printed (table 2), an indication that the great majority of letters did not require further editing.

Batch mode was used about 1% of the time during the evaluation carried out in 3M's largest, most diverse technical library. This mode is most useful for more specialized information facilities that search a few sources almost exclusively. For example, at 3M, patent searchers conduct a much higher proportion of their searches in batches, so that twenty to thirty results are received at the same time from one source. Thus twenty or thirty cover letters with the same date, searcher, source, and instructions on obtaining documents must be produced.

The fact that 70–80% of letters produced using COVER have no corrections made on them during verification is encouraging. However, enough are corrected to require a separate verification step and, though search topic is the most frequently corrected part of a letter, errors are corrected in all parts, and thus all need to be verified (table 3).

COVER is a successful tool for increased library productivity. While it takes more time than a note, a professional-quality letter can be produced in a majority of cases in

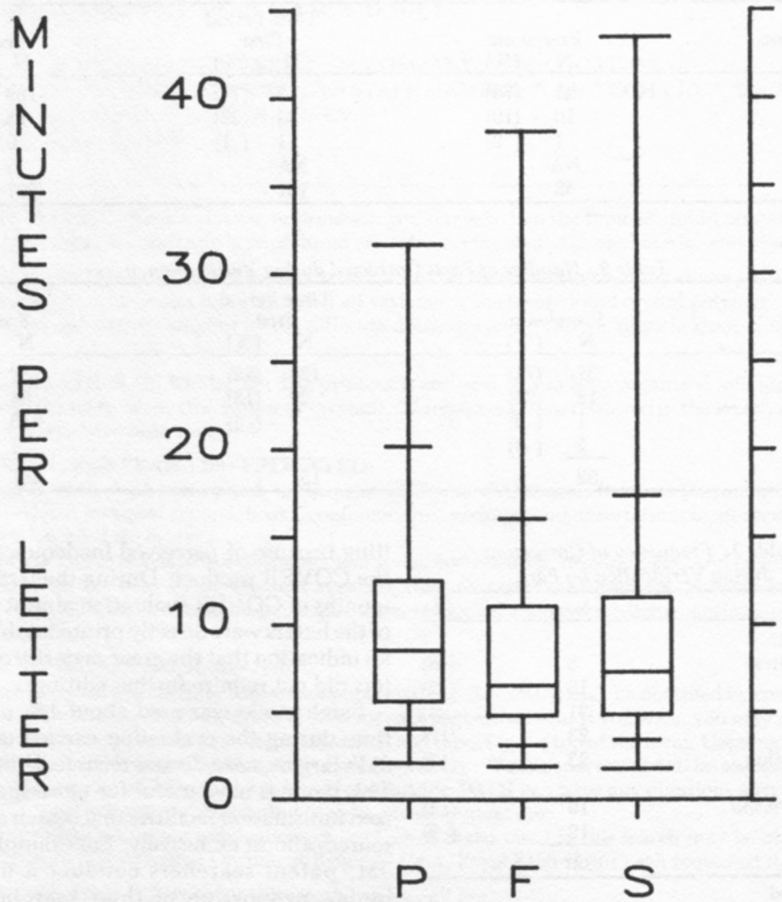


Fig. 4. Time needed to produce one letter using COVER during three different time intervals: P = prerelease, two weeks in April 1985 (n = 52); F = first release, May–July 1985 (n = 151); S = second release, August–November 1985 (n = 227). The column for each time interval has horizontal bars (from the top) at the maximum value; 90th, 75th, 50th, 25th, 10th percentiles; and the minimum value. The rectangle in each column encloses the middle 50% of the data. The horizontal bar in the rectangle is the median.

five to twelve minutes from its start to its printing or saving. Its success is measured in its daily use in twelve libraries/information centers in the U.S. and Canada. COVER MANAGER permits the library administration to maintain the hundreds of standard paragraphs and insures that COVER can be kept up-to-date.

Search Results Editing and Reporting Package

The REFORM program accepts files of downloaded search results and converts them into its internal format. Once in internal format, files can be merged, erased, or selected for further use, which includes temporary or permanent deletion of records, restoration of temporarily deleted records, deletion and restoration of tagged fields within records, and assignment of priority codes. Records may be displayed on the screen and printed on a printer or to a file—though they can be seen on a screen display, temporarily deleted records are not printed on a printer or to a file.

Records within files can also be sorted or ordered by their user-assigned priority codes. In addition, the present formats of certain more frequently used, tagged databases from selected vendors have been incorporated into REFORM such that it can sort files from these database-vendor combinations based on other keys, including name of first author or inventor, date of publication, company name or corporate source, or patent country and number.

A sample screen display of a REFORM record is shown in figure 5, a sample page of a REFORM report, in figure 6. (This report is summarized in the letter shown in figure 3).

REFORM performs its design criteria successfully, but it is not yet in daily use. The staff of 3M's Information Service are currently developing downloading guidelines that will enable them to incorporate REFORM more effectively into daily use.

REFORM can be compared both with commercial general-purpose word-processing packages and software for managing personal bibliographies. Its major advantage over a word-processing package is the ability to sort records by several fields, expand tags reversibly, and delete complete or partial records, so that the same search can easily be output in several different formats. A word-processing package may be preferred for simple editing and printing of unsorted references without change in format.

Packages for managing personal bibliographies, such as SCI-MATE Personal Data Manager,[1] can accept bibliographic information in predefined formats and output them in different formats, optionally sorted by different attributes. However they require specific information about the format of a file before it can be accepted and may require that the file be edited through a text editor or word processor before it can be used.

REFORM's major advantage over such packages is its ability to accept files from its standard database-vendor pairs without preediting or special formatting instructions. However, the formatting information is still required: rather than being entered by the user, it is part of REFORM. Thus the REFORM program and/or its associated data files must be changed whenever a new database-vendor pair is added or a format changes. Such changes are not difficult but require knowledge of the BASIC programming language and access to REFORM's BASIC source code and a BASIC compiler. The expense of this maintenance requirement is small compared with the expense of purchasing twelve copies of *and* updates to a commercial package and with the time that would be required by each of twelve or more searchers to enter and update the formatting information themselves.

The letters produced by COVER and the reports produced by REFORM can greatly improve the quality of search-output results. Other workers reporting on alter-

```
Type ? for Help                                        Type ^ to Backup
              REFORMAT SEARCH RESULTS - FILE COMMANDS
Version 1.5                                     (c) 1985 3M Company

NEW        OLD        ERASE        MERGE        INSTRUCTIONS
QUIT

     WHICH DO YOU WANT:
```

Fig. 5. Initial screen display for REFORM.

3M INFORMATION SERVICES SEARCH RESULTS — Page 1

Liquid Crystal Polymers

Alphabetized by Name of First Author or Inventor

Record from CA-CAS
Document Title
Optical textures observed during the shearing of thermotropic liquid-crystal polymers
Author (Inventor)
Alderman, N.J.; Mackley, M.R.
Corporate Source or Patent Assignee
Dep. Chem. Eng., Univ. Cambridge
Corporate Source or Patent Assignee Location
Cambridge CB2, 3RA, UK
Document Source
Faraday Discuss. Chem. Soc., 79 149–60, 7 plates
Document Type
J
Publication Year
1985
Language
Eng

Record from NTIS-DIALOG
2/4/4
Title
—Exploratory Development of Foams from Liquid Crystal Polymers
Title
—(Note) Final rept.
Author
—Chung, T.S.
Corporate Source
—Celanese Research Co., Summit, NJ.
Sponsoring Organization
—National Aeronautics and Space Administration, Washington, DC.
Publication Year
—Jan. 85
Language
—English
Contract Number
—NAS1-17290

Fig. 6. First page of a REFORM report of the search summarized in figure 3.

native methods for improving search output have noted that changing to a more professional format results in many favorable comments and reduces the number of questions on search strategy, databases used, and time periods covered.[2-5] We expect similar benefits from COVER and REFORM.

CONCLUSIONS

COVER and REFORM, two valuable tools for search output enhancement, have been developed. COVER has been evaluated as effective and is in daily use in twelve facilities in the U.S. and Canada. Using it, a search librarian, in the majority of cases, can enter the information necessary for producing a professional-quality search-result cover letter within five to twelve minutes.

Still to be determined are the cost-effectiveness of REFORM and the maintainability of both these tools. Their applicability for use in other information center environments is unknown.

ACKNOWLEDGMENTS

This work, which was performed while the first author was on leave from the University of Minnesota, would not have been possible without the support and encouragement of many 3M staff members, including Barbara Peterson, Karen Flynn, David Schrader, Christine Portoghese, and the searchers in the 3M Technical Library. Further information on COVER and REFORM can be obtained from Kristin K. Oberts, Information Services 201-2S-09, 3M Center, St. Paul, MN 55144.

REFERENCES

1. C. Stout, "Sci-Mate: A Menu-Driven Universal Online Searcher and Personal Data Manager," *Online* 7, no.5:112–16 (Sept. 1983).
2. C. Riley, M. Bell, and T. Finucane, "Elimination of Duplicate Citations from Cross Database Searching Using an 'Intelligent' Terminal to Produce Report Style Searches," *Online* 5, no.4:36–41 (Oct. 1981).
3. C. W. Grotophorst, "Another Method for Editing Downloaded Files," *Online* 8, no.5:85–93 (Sept. 1984).
4. R. Miller, "Designing Your Own Low Cost Front-End Software," *Online* 9, no.2:94–98 (Mar. 1985).
5. C. Tenopir, "Software for Online Searching," *Library Journal* 110, no.17:52–53 (Oct. 1985).

The Engineering Information System: A Guided Tour

Scott Deerwester

The Engineering Information System (EIS) is an interactive information retrieval system that was written at Purdue University between 1979 and 1983. In this article we concentrate on features that are unique to EIS and that have been shown to increase its perceived utility to users. The features reviewed here are

- Appropriate help at any point in a search session with a single keystroke. This means, for example, that a user could ask for help in the middle of typing a keyword, read through the help message, and complete the keyword.
- The abstraction of the bibliographic database software as a virtual retrieval machine. A user interface accomplishes searches by interacting with such a machine, allowing multiple user interfaces to access the same databases, and a single user interface to access very different databases.
- The incorporation of stemming into both the user interface and the underlying database software. Documents that contain words that are morphologically related to words in users' queries are retrieved as well as those containing the precise words.
- The user interface incorporates a window interface that runs on any cursor-addressable terminal.
- The entire system is built on a widely distributed network. It is possible for a user interface running on any machine on the network to access databases on any other machine. Retrieval servers, which are implementations of the abstract retrieval machine, effect retrieval in response to requests by user interfaces over the network.
- In the internal search software, lists of retrieved documents (hit lists) are implemented by an extension of the UNIX stdio **FILE** structure. Short lists are handled in core. When a list exceeds a threshold, the system automatically creates a file in which it stores the list. When a list is deaccessed, the file is automatically deleted. The merger of in-core lists with temporary files is seen as an important example of data abstraction.

The remainder of this article is divided into three sections. We first discuss the user interface and its features. Next, we present the retrieval server interface and discuss how user interfaces and retrieval servers interact. Finally, we discuss the database software upon which the retrieval servers are based. Technical and functional overviews of EIS are available from Purdue.[1,2]

USER INTERFACES

The primary user interface is based on a window system that incorporates menus and an extensive and extensible help system. The first message that users see informs them that help is available by typing ?. Figure 1 shows the initial main menu that a user would see upon invoking EIS after having read the initial window.

Because people have a relatively limited amount of short-term memory, the user interface is built in a way that augments it. The user "believes" that there are two windows on the screen, one containing an informative message and the other containing a menu. Users don't need to remember the contents of the message because enough

Scott Deerwester is a member of the Graduate Library School faculty, University of Chicago.

```
-----------------------------------------------------------
| EIS User Interface, v2.0                                |
| (-----------------------------earch                     |
|   |■Start new search             |                      |
|   | Display search results       |                      |
| T | Or                           |.  You may            |
| a | And                          |any point.            |
|   | And Not                      |                      |
| T | See current window           |October 29.           |
|   | Remove saved record frame    |                      |
|   | Change database              |                      |
|   | Exit                         |                      |
|   |______________________________|                      |
|_________________________________________________________|
```

Fig. 1.

clues are present on the screen to remind them. The design of the windows, i.e., the characters that represent the windows' borders, were chosen to maximize the illusion of one window covering another. The first principle behind the user interface is thus:

Use the screen as a visual cache for short term memory.

The second principle is similar. Users ask for help because they need it to accomplish some task. Help should therefore be available anywhere, even in the middle of performing a task. Further, the presentation of helpful information to users should not erase from the screen all of the visual reminders of what they were doing prior to asking for help. The second principle is thus:

Relevant help should be available anywhere with a single keystroke.

The principle is implemented by reserving an ASCII character (?) as the help character. If, for example users were to press ? while in the main menu, as above, they would see the screen depicted in figure 2. The message **Hit any character to continue** appears at the bottom of the screen.

The window system upon which the user interface is based uses the UNIX *curses* package and consists of approximately 1,500 lines of C. The interface itself consists of about 2,100 lines of C and 220 lines of YACC. A more extensive review of the user interface is given elsewhere.[3]

RETRIEVAL SERVER INTERFACE

The retrieval server is a computer program that performs all retrieval-related functions on behalf of a user interface. The retrieval server itself is never seen by the user, who perceives all functions as carried out by the user interface itself. The server may or may not be running on the same machine as the user interface. This is discussed later.

There are two aspects of the retrieval server interface, apart from the search software per se, that are unique to EIS. First, retrieval servers are implementations of an abstract retrieval machine (ARM). In order to use any particular access retrieval mechanism, it is necessary to implement the ARM using this mechanism. In other words, it is necessary to write a program that uses this retrieval mechanism that appears to the user interface as if it were an abstract retrieval machine. If the retrieval system can be packaged in this way, then it doesn't matter to the system as a whole what the specifics of the mechanism are. The language in which the user interface communicates with the retrieval server is the ARM's "instruction set," by analogy to actual computers' instruction sets. There are three such implementations in current use at Purdue.

```
-----------------------------------------------------
| EIS User Interface, v2.0                          |
| (-----------------------------earch               |
| | Start new search               |                |
| | D---------------------------------------------------
| T| O| In a menu, the following keys are valid:        |
| a| A|                                                  |
|  | A| j        move down                              |
| T| S| k        move up                                |
|  | R| J        move to the bottom                     |
|  | C| K        move to the top                        |
|  | E| /c       move to the option starting with       |
|  |__|          'c', where 'c' is a letter             |
|_____| RETURN   select an option                       |
      | Ctrl-x   get out of the menu                    |
      |_________________________________________________|
```

Fig. 2.

Second, access to a retrieval server is obtained over a local area network. User interfaces have, for each bibliographic database, a machine name, server name, and database name. When a user wishes to access a given database, the user interface connects to the machine upon which the database resides and invokes a retrieval server to initiate access to the database. Knowledgable users can even use the user interface to interact with their own *refer* databases using a *refer* retrieval server.*

Abstract Retrieval Machine

The abstract retrieval machine was designed to maximize the flexibility of retrieval access and minimize the amount of information that must be communicated between a user interface and a retrieval server. The ARM instruction set is shown in figure 3.

In figure 3 x, y, and z represent integers in hexadecimal using ASCII digits and letters, and s is a string of alphanumeric ASCII characters. Hexadecimal is used because it requires fewer bytes to represent numbers than decimal but is still human-readable, although the only humans who actually have occasion to read ARM instructions are programmers debugging servers. The ARM instruction set is never seen by users. The commands that can create a hit list are OR, AND, AND NOT, and SEARCH. Each of these returns an integer id (also represented in hexadecimal) that is used subsequently to refer to the hit list. By "return" we mean that the server passes a character string to the user interface that represents a return value. A function within the interface interprets this character string and returns the corresponding value to its own caller. To give an example, a user query **university and (purdue or chicago)** might result in the following dialog between a user interface and retrieval server.

Interface	Retrieval Server
university;	**1**
purdue;	**2**
chicago;	**3**
|2,3;	**4**
&1,4;	**5**

Normally the Boolean operators are destructive, meaning that they deallocate their operand hit lists. This is, however, "settable." The *howmany* operator returns the size of a hit list. A user interface might request the size of the hit list to report to the user using *howmany*, thus:

Interface	Retrieval Server
?5;	**7B**

indicating that there are 123 records ($7B_{16}$) in the resulting hit list. To actually retrieve

**Refer* is a program to perform bibliographic functions related to word processing.

& *x*, *y*;	*and*
| *x*, *y*;	*or*
~ *x*, *y*;	*and not*
@ *x*;	*free hit list*
? *x*;	*howmany*
+ *x*, *y*, *z*;	*seek*
# *x*, *y*;	*get record*
s;	*search*

Fig. 3.

a record, the user interface asks for the id of the hit at the current location in a specific hit list, using the *get hit* (.) instruction and requesting the record corresponding to this hit with the *get record* (#) instruction. The second operand of *get record* is a flag indicating whether a full or abbreviated record should be returned. The server returns an integer byte count, specifying how many bytes are in the record, and then the contents of the record.

To continue with the above example, the following dialog would retrieve the first record:

Interface	Retrieval Server
.5;	6F24
#6F24,0;	FD

followed by the 253 (FD_{16}) bytes contained in the record.

Nonsequential access of hit lists is accomplished by seeking desired locations in the hit list using *seek* (+). The operands of *seek* correspond to those of the UNIX **lseek** system call. The first operand is a hit list. The interpretation of the second depends on the value of the third. If the third operand is 0, the second is the number of the hit to which to seek. If it is 1, then the second is interpreted as a relative offset, in hits, from the current location in the hit list. If the third operand is 2, then the offset is from the end of the hit list. A user interface could request that the current position of a hit list be placed at its beginning as follows:

Interface	Retrieval Server
+ 5,0,0;	0

Advantages of a Virtual Machine Implementation

There are three major advantages in implementing a retrieval server as a virtual machine. First, the server is stateless. This means that the server does not need to keep track of the state of a given interaction with a user interface. The responsibility for keeping track of which hit lists are in use belongs completely to the user interface. The server allocates and frees hit lists in response to requests from the user interface.

Second, the number of bytes that must be passed between the two programs is small, since the ARM language is very terse. Since this language is used only to communicate between the two processes—the user interface and the retrieval server—it does not need any facilities for user-friendliness. Online help, error feedback, and similar functions that require large amounts of interaction with the user's terminal are all handled by the user interface and never even seen by the retrieval server.

Third, it hides the implementation of the retrieval mechanism from the user interface, enforcing a very desirable degree of modularity and allowing users and system maintainers to simply plug in different versions of either servers or interfaces, knowing that compatibility with the remaining modules with be maintained. This has been well tested, both while using several user interfaces with the same server and while using several servers with the same user interface.

Network Interface

The retrieval machine is invoked by a special set of functions, in the user interface code, that replace the corresponding standard functions. There is a C function that corresponds to each of the operators in the ARM instruction set. The *and* operator, for example, corresponds to a function called **and()**. The user interface is thus written as if it were simply calling the functions that comprise the retrieval system. In fact, these functions are dummy functions that simply pass the requests on to the retrieval server, await a reply, and return (in the sense of a function return) the reply to their callers. If, for example, the user interface invokes the **and()** function, the special version of this function that is compiled into the interface passes a request to a retrieval server, waits for the reply, and returns the appropriate value to the caller. Had the "real" version of **and()** been compiled in instead, it would have performed the same function directly, rather than pass a request to a server. Neither the users nor, in fact, the user interface itself are able to tell whether these operations are performed internally by the user interface or by calls to an external server.

The user interface has three pieces of information for each database:

- Machine name
- Retrieval server name
- Database name

The user interface accesses, by default, a database containing all of the monographs in the engineering library. This database is updated weekly to incorporate tables of contents for all newly acquired monographs. Keywords from the tables of contents are included in the indexing of the documents. The user may choose to access other databases as well, including their own private databases. When a user initiates access to a database, the interface invokes an internal function, **initdb()**, with the name of the database as an argument. The function consults a table of databases and issues a remote command on the database's machine invoking the appropriate server with the name of the database as an argument.

An earlier version of **initdb()** used a different strategy: since all of the engineering library databases reside on a single machine, the user interface would first attempt to initiate access to the database locally—i.e., on the machine on which the interface itself is running—and if this access failed, attempt access on the engineering library's machine. The current strategy allows databases to reside on any number of machines.

Once a retrieval server has been invoked, all communication between the server and the user interface is done using ASCII character strings, in the ARM instruction set. Upon invocation, the server returns an integer indicating whether it successfully opened the requested database. If not, the user interface reports the error to the user.

In practice, users do not know whether the databases that they are using are on their machine or a remote machine. The amount of information that passes between the server and the interface is so small that it essentially happens instantaneously. The only bottleneck is in passing record texts across the network. By using appropriate "read-ahead," this is also perceived by users as very quick. As soon as a document is written to the screen, the user interface requests the next record from the server—so that by the time the user is finished reading the current record on the screen, the next record is normally already available and can be written quickly to the screen.

BIBLIOGRAPHIC DATABASE SOFTWARE

Although there are several kinds of access mechanisms—i.e., different retrieval systems—available to the users, a single mechanism serves the bulk of the requests. The native retrieval system, also labeled EIS, consists of approximately 6,500 lines of C and some thirty-five functions. In addition to the functions that correspond to ARM instructions, these include functions used to build databases, utility functions used in various ways, etc. In this section we discuss both the EIS database software and the alternate access mechanisms.

EIS Database Software

The first version of EIS was inspired by the Boston Children's Museum Information System (IS), written by Bill Mayhew in the mid 1970s. EIS uses a hash table as its primary access mechanism, giving remarkably fast access time.* From the time that a user types a query to the time that the system reports the number of documents retrieved by the query is almost never more than a few seconds. The disadvantage of using a hash file is that pattern matching searches are not possible. A user could not, for example, find all documents containing a word beginning with **alph**.

A second feature of the search software

*A hash table is an access method wherein the address of an item can be computed on the basis of the key under which the item is indexed.[4]

mitigates the effect of this lack. The system incorporates a stemming function that is used both in indexing and retrieval.[5,6] When documents are indexed, the stem of each word, as well as the word itself, is associated with the document as a keyword. A document that is indexed under **libraries**, for example, would automatically be indexed under **library** as well, and similarly for documents indexed under **librarian**, etc. During retrieval, the user interface expands each search term by performing a search for the exact term and for the stem of the term and merging the resultant hit lists. Thus, if a user entered **libraries** as a search term, the user interface would automatically expand the expected

search ("libraries") into
or (search ("libraries"), search (stem ("libraries")))
where **stem ("libraries")** returns the stem of the word "**libraries**," namely "**library**."

This is, of course, "settable" by the users. Quoting a word in a query causes the system to match on the precise word, not expanding the search to include the word's stem.

A third important feature of EIS' database software is its handling of lists of hits. Hits—or primary keys—are represented as long integers, and a hit list is thus simply a list of long integers, in which each hit (i.e., each long integer) uses four bytes.* The operations that must be performed on a hit list are

create	Create a new hitlist
put	Insert a hit into a hitlist
get	Get a hit from a hitlist
seek	Move to a given point in a hitlist
free	De-allocate a hitlist
howmany	Return the size of a hitlist

The analogy of operations on a file is appealing, since each of these operations has an equivalent in character I/O using files. One option, therefore, is to simply put each hit list in its own file, and use the I/O primitives provided by the operating system to implement the required capabilities. There are, however, two prohibitive costs in doing so. First, a given UNIX process can only have a limited number of files open at a time, and when the number of hit lists exceeds this number, the system must take some corrective action. Second, actually interacting with hit lists—especially short ones—using disk I/O is extremely and unnecessarily slow. It would be much better to be able to deal with short lists in memory and write only long lists to disk.

*More efficient representations are, of course, possible, but this is not considered here.

The solution is to write a special interface to the system buffered I/O routines. The system routines use in-core buffers for the bulk of their access requests, filling or flushing the buffer as needed. By interposing routines that intercept requests to fill or flush buffers and create files on the fly to handle requests that will not fit in core, the system routines can be used to handle all hit list requests. A casual inspection of the internal software would lead to the conclusion that nothing more than normal UNIX I/O was used to manipulate hit lists when in fact a relatively small number of hit lists actually have files associated with them. The resulting increase in performance and the ease of interaction have both proven important in the lifetime of the system.

Alternative Database Software

There are two other servers available to EIS users. The first, called the text server, uses a very simple access mechanism whose dictionary is a sorted ASCII file searched using a standard binary search. Records are stored sequentially as a normal text file, separated by blank lines. This server has been used for small, volatile databases, for example the list of reserve materials. The records can be modified with a text editor, and a modest collection reindexed in a half-hour. The disadvantage, of course, is that the search time is O(log(n)) rather than O(1) for EIS.

The third server interacts with *refer* databases.[7] This server presents some special problems. *Refer* was written assuming that all retrievals would be for a single, known item. It was designed to be a preprocessor for *troff*, a text-formatting program. Users embed a search request in a text file. Running *refer* on the file results in the request being replaced by an actual citation in the formatted output text. Under this constraint, *refer* works well—searches are accomplished quickly with little space overhead. If *refer's* search strategy is used for unknown item searches, the performance rapidly degrades.

Nonetheless, it has proven useful to users to be able to use a full-feature information retrieval system to interact with their own databases of bibliographic records. This is especially true given the typically small size of such databases, with at most a few hundred citations.

CONCLUSIONS

Functionally, EIS has been successful in the four years since it was completed. It has become the primary means by which Engineering faculty, students, and staff search the Engineering Library's resources. The library staff also make ample use of its facilities and have published a number of papers on its use. Technically, it incorporates innovations that have proven successful and could be used in other systems. The implementation of a retrieval server as a virtual retrieval machine has had particularly good consequences, both in ensuring a clean interface between modules of the system and in allowing the system to take full advantage of local area networking.

The version of EIS described here is not available outside of Purdue University, although it is possible to use the system over dial-up lines. A second version of EIS was written during 1984 at Colgate University. The second version has somewhat more functionality, allowing weighted, in addition to standard, Boolean searching. Functionally, this means that searches result in hit lists that are ordered in decreasing order of estimated relevance, rather than a set of documents that are simply partitioned into relevant and nonrelevant. If interest warrants, this version may be made publically available.

REFERENCES

1. Alan Alexander-Manifold, *EIS—An Overview* (West Lafayette, Ind.: Purdue Univ. Libraries).
2. Library Systems Department, *EIS User's Manual*, (West Lafayette, Ind.: Purdue Univ. Libraries, 1983).
3. Edwin D. Posey and Charolotte A. Erdmann, "An Online UNIX-Based Engineering Library Catalog: Purdue University Engineering Library," *Science & Technology Libraries* 6, no.4:31-43 (Summer 1986).
4. Ellis Horowitz and Sartaj Sahni, *Fundamentals of Data Structures* (Potomac, Md.: Computer Science Pr., 1976), p.456–71.
5. Scott C. Deerwester, "The Implementation of a Stemming Algorithm in an On-Line Information Retrieval System," *Proceedings of the 1981 ASIS Midyear Meeting*, Durango, Colo., 1981.
6. Julie B. Lovins, "Development of a Stemming Algorithm," *Mechanical Translation* 11:22–31 (1968).
7. Bill Tuthill, "Refer—A Bibliography System," in *UNIX User's Manual, Supplementary Documents* (Berkeley: Univ. of California Computer Science Div., July 1983).

Communications

Using a Personal Computer to Provide Online and Executable Documentation for Searching Bibliographic Databases

William Leigh and Noemi Paz

A bibliographic database searcher working in an organization may devote much time to "chauffeuring" colleagues in the use of a system that he or she mastered some time ago. Expert searchers often work with many systems containing differences in syntax and function, but some of these systems may be consulted quite rarely. Expert searchers often must dig back into notes and even into the system manuals themselves to refresh their memories concerning some aspect of the system that they or someone else now wishes to employ. Extended research (of a particularly painful sort) may even be required to exercise a system function that is used regularly but seems to have slipped the mind at the moment.

These problems may be alleviated through "executable documentation" residing on a personal computer. This technique affords the developer a semiformal medium for organizing, recording, annotating, and distributing guidelines and instructions for the use of database systems. This documentation may include both prose narrative for the user and formal commands for the computer system. Online access to documentation is concurrent with use of the database system itself, and this documentation may be prepared for the searcher's own later use or as a tutorial for colleagues.

William Leigh and Noemi Paz are both professors in the Department of Computer Science at the University of Southern Mississippi, Hattiesburg.

EXECUTABLE DOCUMENTATION

This technique requires that a personal computer with a communications program be used as the search vehicle and that a "memory-partitioning" program, sometimes called an "integration" program, be available on the computer. A large variety of such programs is available commercially at this time. (We use one called Memory/Shift from North American Business Systems; others are available for all major brands of personal computers.)

A memory-partitioning program on a personal computer allows two or more separate programs to reside in memory simultaneously, in separate "partitions." The user can alternately select among the resident programs: when one of the programs is selected, it executes and may be used as though it were the only program residing on the computer.

Many of these memory-partitioning programs support the capability for copying text from one partition into another partition as though it were entered from the keyboard, and this feature makes possible the executable nature of our documentation technique. Using this facility, a piece of the documentation, which is an example command in the formal search language of the database system, can be copied from the documentation text in one partition and, without rekeying, submitted to the subject system accessible in the other partition.

Setting up the PC for executable documentation involves several steps: the communications program is invoked to run in one partition under the memory-partitioning program, a word-processing program is invoked to run in a second partition, and documentation for the database system is supplied as a file to be read by the word-processing program and manipulated by the user employing its facilities.

Now, when the memory-partition program is supporting the execution of both the communications and word-processing programs, and when the user has read the appropriate documentation text file (the "script") into the word processor, the elements for executable documentation are in place. With the press of a key, the user may transfer attention from the documentation text to the actual use of the database system via the communications program. (See figure 1 for a diagram of the relationship between these components.)

When viewing the script, users may employ the search facilities of the word processor for quicker scanning, insert their own notes into the documentation text, or copy selected portions of the text, such as search commands, into the memory-partitioning program's transfer buffer. After the user's memory is refreshed by consulting the script, after a command that is particularly difficult to retype, or after an extended piece of alternative source code or data is copied into the transfer buffer, attention may be shifted back to the database system. Work can be resumed until the documentation text must again be consulted, and the copied command line can be submitted to the bibliographic database system.

An alternative method for executing the documentation is to use the word-processing program's facilities for copying from the text into a work disk file. After switching to the partition containing the communications program that is connected to the database system, the communications program's function for "uploading" a disk file can be employed to transmit the commands or data excerpted from the documentation script to the database system. This alternative technique is useful when the memory-partitioning program does not support "cut-and-paste" between partitions or when a large piece of text, such as a data file, needs to be transferred.

SCRIPT DEVELOPMENT

The problems of developing scripts are largely the same as those encountered in writing hard-copy computer system documentation. Special problems include those posed by the limited screen of the personal computer and the planning that must be done for the executable feature.

As in hard-copy documentation, it is necessary to establish conventions for coding the scripts. Computer-system dialogue examples and formal command examples must be clearly differentiated from annotations supplied for tutorial and information purposes. It is good practice to include an explanation of each script's conventions. (The example script supplied in figure 2 contains such a convention section; the example conventions are limited by a necessity to use only a monochrome computer display and standard characters.) The availability of appropriate equipment would allow color, bold printing, highlighting, italics, and special graphic symbols to be used in scripts to good effect.

It is particularly important to identify

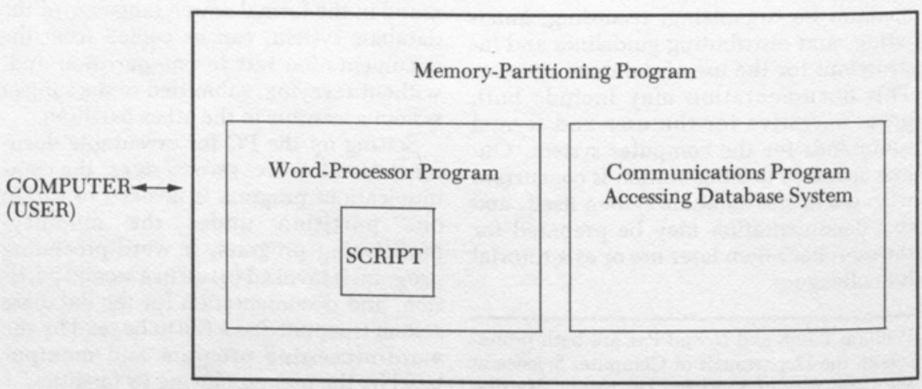


Fig. 1. The Components in Executable Documentation.

procedural steps clearly. This may be done by numbering the steps in the procedure in order of execution, which is more important for longer procedures and those where it is crucial that no steps be skipped and that they be done in the specified order.

Blocks of text that may be transferred out of the text to files or to the other partition must be furnished in blocks uncluttered by other documentation—this permits their transfer in a single operation.

The allowance for the exploitation of the word-processing features is an opportunity rather than a problem in script preparation. In particular, the "search" ability of the word processor may be used to implement a crude index within the script. A section of the script can identify text strings which, when searched for, will cause associated subsections of the script to be displayed. (See figure 2 for an example of this technique.)

The simple script in figure 2 was prepared in accordance with these listed guidelines: an embedded index and a script convention section are included, and the content of the script was adapted (from Borgman and others, 1984).

USING AN "OUTLINING" PROGRAM

This tool can streamline the develop-

```
[  Script File Name—brs

[  Immediately below is a table of contents for this script. Use your word processor's search facility with the
[     word shown in brackets {  } as a search argument (include the brackets) to find the section of the script
[     you desire to consult.

[  {adj}        —Adjacency searching.
[  {exp}        —Subject-term expansion.
[  {script}     —Explanation of conventions used in preparing this script.

   [  {adj}        —Performing adjacency searching on BRS.

   [               "Adjacency" refers to searching on multiple terms that must appear immediately adjacent
   [               to each other.
   [

   [               This is contrasted to searching on multiple terms joined with and, where the terms need
   [               not appear adjacently.
   [

   [               Adjacency searching can reduce the problem of irrelevant retrieval.
   [

BRS—SEARCH MODE—ENTER QUERY
   1_:      higher and education
   RESULT      83637                             [  This will search for two single
                                                 [     free-text words that may appear
                                                 [     somewhere in the same unit
                                                 [     record.

   2_:      higher adj education
   RESULT      78034                             [  In this search the two words must
                                                 [     appear side by side in the same
                                                 [     unit record.

   3_:      higher-education
   RESULT      75102                             [  Use of a multiword subject
                                                 [     heading from ERIC thesaurus.
                                                 [     Note that the words are joined
                                                 [     with a hyphen.

   4_:      . . off
```

Fig. 2. Example of Script for Searching Techniques on BRS.

```
[ {exp}       —Subject-term expansion on BRS.

[             This technique is useful for identifying
[                all forms of a term.

BRS/PSYC/1967—APR 1983
BRS—SEARCH MODE—ENTER QUERY
 1_:  root marriage                          [ The command is "root"
                                             [   followed by the term to be
                                             [   expanded.
R1    MARRIAGE                               3922  DOCUMENTS
R2    MARRIAGE-ATTITUDES                      257  DOCUMENTS
R3    MARRIAGE-COUNSELING                     937  DOCUMENTS
R4    MARRIAGEABILITY                          16  DOCUMENTS
R5    MARRIAGES                               712  DOCUMENTS

      [ {script}      —Explanation of conventions used in preparing this script.
[     System commands and example dialog and data to be transferred are to the left of the [ symbols.
[        Annotation is to the right of the [ symbols. In the left section commands keyed by the searcher
[        are in lowercase text. Computer-supplied text is in uppercase.
[     Note: the user might enclose comments and additions to the script in < > symbols so that they
[        may be recognized as user-supplied.
[     Tokens to be used as finders are coded {finder-token}. These tokens, including {  }, are to be
[        used as search arguments in the operation of your word processor for finding the sections of the
[        script so marked.
```

Fig. 2. Continued.

ment and presentation of scripts for executable documentation: the outlining program is used as a replacement for the word-processing program in executable documentation; it resides in one partition, the communications program in the other. (We use an outlining program known as ThinkTank from Living Videotext, but several are available commercially.)

Outlining programs allow the organization and access of text according to a user-designed hierarchical structure. When accessing the text that is stored in an outline structure, various levels and depths of the outline may be selected for display. In this way the searcher-user can navigate quickly and effectively in a complex script, seeing detailed instructions only when requested.

Figure 3 shows a top-level outline display for searching information supplied in a 1986 article by Mourey. The first three levels of the outline are displayed here. The + (plus) sign indicates that there is material in the outline under the heading. The – (minus) sign indicates that there is no more information under the heading. Deeper levels of this outline-script under headings marked + might contain text scripts similar to the word processor–based one in figure 2.

```
+ Expert searcher's advice on starting points for
  searches involving statistics.

+ Author Search
   + SCISEARCH
   – SOCIAL SEARCH
+ Cited Reference Search
   + SCISEARCH
   – SOCIAL SCISEARCH
+ Software package Search
   + Locating package
   + Review
+ Topic Search
   + MATHSCI
   + DIALINDEX
+ Problem Solving Search
   – COMPENDEX
   – NTIS
```

Fig. 3. Display of High-level Search Strategy Online Script.

Executable documentation with the outlining program uses the capabilities of the personal computer better than the word processor implementation. This outlining capability allows script development and use to greatly exceed the utility of hard-copy documentation.

DISCUSSION

A personal computer can be used to provide online and executable documentation for database searching. The documentation is "online" in that it may be consulted concurrently with the conduct of a search. This documentation is "executable" in that appropriate parts of it may be transferred to the database system and executed there and in that its contents, in some ways, can be interpreted by the word-processing or outlining program to control its presentation.

Documentation may consist of notes made by the searcher to be used as reminders and as "canned" procedures at a later time; or it can be of a more tutorial nature, prepared for persons other than professional searchers using a database system. The documentation script can be tailored to the purposes of the intended user; an index to subsections of the script can be provided easily. This technique uses only a personal computer and common, commercially available software and readily admits user extension and modification.

This approach is especially effective for casual users of database systems, that is, those who use a particular system infrequently. In this situation sketchy notes and examples suffice for a script. In the case of novice users, more elaborate and carefully prepared scripts must be provided, though this is perhaps more easily accomplished under this technique than any other. It is easily envisioned that database vendors might make special-purpose tutorial scripts available.

This technique can be more effective than hard-copy documentation techniques with the same purpose. The convenience of online documentation and the resulting frequency of use are frequently described. Documentation supplied under this described method can be distributed conveniently on diskette; can reside on the personal computer that is used as a terminal for conducting searches; can be easily modified and extended by the user; and, if so distributed, can be easily downloaded when it is needed so that current, common versions may be conveniently and centrally maintained. The use of an outlining program for the development and access of scripts brings to bear much more powerful organizing and presentation techniques than are possible with hard-copy documentation.

This method also offers an opportunity to interpose between the users and the bibliographic database systems a more uniform interface style and set of conventions than would normally be present in a collection of documentation manuals supplied by the vendors of database systems. A user who has mastered the executable documentation concept and a set of scripting conventions is prepared to use all the systems and system features that have similar scripts prepared. This can effectively integrate the use of disparate systems, establishing commonality between differing interface styles and semantics.

This method offers utility in cases where long, difficult commands may be keyed or where sets of data must be modified. This last advantage has no counterpart in hard-copy documentation.

A limitation of the technique is that its users must be adept with their personal computer's word-processing, outlining, memory-partitioning, and communications software. This requirement is met by a high and rapidly increasing proportion of present and prospective bibliographic database users in modern organizations.

ACKNOWLEDGMENT

The authors gratefully acknowledge the NASA Office of Technology Utilization, which partially supported this work.

BIBLIOGRAPHY

Borgman, C. L., D. Moghdam, and P. K. Corbett. *Effective Online Searching* (New York: Dekker, 1984).

Mourey, D., "Conquering the Fear of Searching for Statistical Information," *Online* 10, no.2:59–64 (Mar. 1986). ■■

Octanet—An Electronic Library Network; The First Four Years: Summary and Evaluation

Claire Gadzikowski

Each of the four Octanet modules (interlibrary loan, electronic messaging, reporting, and union list production) is described. This summary is followed by a report on a user-satisfaction survey evaluating the Octanet system. A final summary includes information on the number of institutions, the hours of usage, and requests input during each of the four years of Octanet operation.

INTRODUCTION

Octanet is a computer-based telecommunications network for the delivery of services to the Midcontinental Region (Colorado, Kansas, Missouri, Nebraska, Utah, and Wyoming) of the National Library of Medicine's Regional Medical Library network. It is a joint project of three entities: the Midcontinental Regional Medical Library Program (MCRMLP), Washington University School of Medicine Library's PHILSOM serials control network (PHILSOM), and the Medical Computing Facilities at Washington University in St. Louis, Missouri (MCF). A detailed description of the design and development of the network may be found elsewhere.[1] Briefly, the MCRMLP manages and coordinates network operations. Washington University provides the PHILSOM database, which was originally designed and is still used for serials record control; PHILSOM is the source of holdings information for the routing of Octanet ILL requests. MCF wrote the computer programs and provides expertise in the installation and local supervision of telecommunications equipment.

Octanet has functioned as a prototype for DOCLINE, the nationwide automated interlibrary loan system of the National Library of Medicine (NLM). Initial development of the project was funded by approved reallocated funds from the NLM contract that supports the MCRMLP. Since July 1983, Octanet has been funded on a complete cost-recovery basis except for management personnel time, which is still covered by the MCRMP contract with NLM. All other costs of the system are paid by its users, who are billed at an hourly rate for their online time.

At present, Octanet is connected to DOCLINE by means of a computer-to-computer interface, which is discussed later in this paper. In the future, the prototype will be phased out, and Region 4, like the other six regions in NLM's national network, will participate directly in DOCLINE.

Claire Gadzikowski is Special Projects Coordinator at Midcontinental Regional Medical Library Program, McGoogan Library of Medicine, University of Nebraska Medical Center, Omaha.

SUMMARY OF SYSTEM CAPABILITIES

During the four years 1982–85, Octanet services consisted largely of the automated routing of interlibrary loan requests from regional borrowers to their resource libraries and the NLM. An electronic messaging system has been a functional component of the ILL module and has also been used by many individuals in user institutions for messages not related to ILL.

The system also generates reports for management and for system users and has produced batch-mode union lists for four health sciences library consortia in the Midcontinental Region.

Octanet is a command-driven system. The main system prompt appears immediately upon completion of log-in and whenever the user presses CTRL/C to change function:

FUNCTION >

In response to this prompt, the Octanet user enters a two-letter mnemonic that represents the system operation or "function" to be performed. For example, the mnemonic for regular request input is *KA* (all ILL functions begin with *K*), and the mnemonic for the electronic message function is *MX*.

In addition to the function mnemonics,

the Octanet lexicon includes four-letter institutional "location codes" (e.g., NEBØ for the University of Nebraska Medical Center library), "user codes" that are personal names, and one-letter symbols for transaction types (e.g., *A* for a request input using function *KA*, *P* for "photocopy sent," *R* for "item received"). Location codes and user codes are used to gain access to Octanet and to indicate the destination for an electronic message. For more about transaction types, see New Services below.

Interlibrary Loan

Octanet interlibrary loan functions may be divided into two classes. The first involves the expediting of operations done manually by ILL personnel: request input and record keeping. The second classification includes new services not available in a manual ILL system.

Expediting Manual Operations

Request input is primarily handled by two Octanet functions: the regular-input function (*KA*) and the express-input function (*KE*). The former includes full prompts for the PHILSOM number of the journal title and for all of the fields included on an ALA interlibrary loan request form. The regular-input algorithm also displays full routing information for the request. Express input substitutes two-letter mnemonics for all prompts and provides an abbreviated display. For an experienced operator, express input may reduce input time as much as 50 percent.

Octanet borrowers are assisted in their record keeping by three system-generated reports. The Borrowers Report is provided monthly by the RML; this report contains a listing of the request numbers of all filled requests, grouped by the location codes of the lenders who filled the requests. (Octanet lenders also receive a similarly formatted report, the Lenders Report, grouped by the location codes of their borrowers). A quarterly List of Titles Requested provides all users with a frequency tabulation of the journal titles they have requested, sorted by volume year within journal title. A second quarterly report, the Turnaround Report, provides information on turnaround time (from request input to receipt of item).

The Borrowers Report relieves the Octanet user of the necessity of keeping manual records to document lender billing for ILLs. The List of Titles Requested provides copyright compliance information for borrowers and also acts as a basis for collection development for many Octanet libraries.

The Octanet system does not provide automated billing, either for filled loans or for online time. Each of the system lenders bills borrowers separately and keeps separate billing records. All Octanet users are billed monthly for their online time by the MCRMLP. At present, the charge is $27.50 per user hour. This figure was arrived at by averaging all telecommunications and computer access costs billed to the MCRMLP for a six-month period and dividing that amount by the average number of user hours per month over the same six-month period. This user-charging algorithm has resulted in complete cost recovery.

New Services

An automated ILL system is able to provide services not available in a manual system. For Octanet users, these services include automatic routing, a request status check, and computer-to-computer routing of requests to DOCLINE, NLM's automated, interregional interlibrary loan request system, if the request cannot be filled by an Octanet lender.

All requests are automatically routed first to the borrowing library's resource library, then to other resource libraries in the region on a lowest-traffic basis. Requests are actually sent only to those lenders who have the requested volume and issue in their PHILSOM holdings string for the requested title. If no Octanet lender holds the requested item, the user is informed by the Octanet system that the item was "not found in NET"—i.e., not found in the holdings of any Octanet lender. The user may then cancel the request or elect to have it routed to NLM via DOCLINE.

In a manual ILL system, borrowers have no way of checking the status of a request except by a phone call. Status information is available to Octanet users by means of an online transaction log (see figure 1).

When each line is read horizontally, the transaction log illustrated tells us that the request went from borrower MDT0 (a borrowing library in Nebraska) to lender NEB0 (the University of Nebraska Medical Center), who filled it with a photocopy (transaction type *P*). MDT0 then acknowledged receipt of the item (transaction type *R*), using an Octanet function designed for that purpose. If MDT0 had checked the status of the request before the item was received, the first two lines of the transaction log would have provided the information that the request had been filled by NEB0.

If the request had been sent to NLM, the transaction log would indicate that location RML0 had received the request and referred it to location NLM0 "via the RJE" (remote job entry).

The RJE connects directly with DOCLINE and reformats Octanet requests so that they are acceptable to that system. The actual transfer to DOCLINE of requests "not found in NET" is accomplished once a day by means of an "OUTWATS" line between St. Louis and Bethesda. Members of the ILL department at Washington University are responsible for this operation.

The RJE was developed during late 1983 and implemented in January 1984. Previous to that time (for almost two years), members of the ILL department at the McGoogan Library, University of Nebraska Medical Center, manually re-input all requests "not found in NET" into DOCLINE. This manual operation took three to four hours per day; the transfer of requests from Octanet to DOCLINE via the RJE takes about ten minutes.

Electronic Messaging

Function *MX* was used more than 11,000 times during 1982–85. It has carried everything from personal messages to job notices, from workshop announcements to "Season's Greetings." It is used often in support of the ILL module, sometimes to request books and occasionally to carry information from one member of the regional chapter of the Medical Library Association to another.

Messages are instantaneously transmitted online. Any given message may be sent to one user, to a number of specified users, or to all system users. The receiver may delete a message from his/her user file as soon as it is read or may elect to keep the message to reread or print out later. All messages, whether deleted by the receiver or not, are stored on the system until Octanet management erases them. Messages that have been read and deleted by the receivers are erased every day; unread messages more than two weeks old are normally erased about once a month.

System-Generated Reports
Reports for Network Management

Chief among these is a monthly report of system usage, sorted by user code (individual personal name) within location code. Per-hour charges to Octanet locations are based on this report.

Several management reports are run quarterly. These reports document lender fill rate, reasons for referral, and processing time (thruput). In addition, a new quarterly management report was run for the first time on requests input during the October–December quarter of 1985. This report, called The Serials Report, lists titles and intraregional referral information for all requests forwarded to NLM via the RJE. Included are requests routed directly to the RJE and requests that were initially routed to one or more resource libraries but could not be filled there; reasons include "in the bindery," "on loan," "lost or missing," etc. The RML will use this report to analyze referral patterns between Region 4 and NLM

19-JAN-86	MDT0	CHRIS	NEB0	A	BORROW VIA KA
20-JAN-86	NEB0	DANA	MDT0	P	PHOTOCOPY SENT
22-JAN-86	MDT0	CHRIS	NEB0	R	PHOTOCOPY RECEIVED

Fig. 1.

and the effect of these patterns on the regional fill rate.

Reports for System Users

These reports were described above, in the section on expediting ILL record keeping.

The Octanet Union List Module

The relationship between Octanet and PHILSOM has made it possible for four consortia within Region 4 to obtain printed union lists of their holdings. These consortia are in St. Louis, Missouri; in Denver, Colorado; in the area of Salt Lake City, Utah; and in the state of Wyoming. More than 100 institutions participate in these four consortia and benefit from the union list module.

Union list consortia have the option of contracting with PHILSOM to input their holdings into the database. Methods of annual updating are also negotiable with PHILSOM. The hard-copy listings are produced by a local St. Louis printer in formats similar but not identical to those of PHILSOM printed products.

Holdings input for the purpose of producing a consortium union list are not automatically available for online activity in PHILSOM or Octanet. However, since the union list holdings are already in the database, they can be made available online, enabling the consortium members to lend and borrow among themselves on Octanet. No additional cost to the consortium is involved except the cost of online time used for ILL operations. To date, two consortia (St. Louis and Wyoming) have taken advantage of this option—St. Louis on January 1, 1985, and Wyoming on January 1, 1986. At this writing, fifteen institutions participate in the St. Louis online consortium, and twelve institutions participate in the Wyoming online consortium.

EVALUATION

In April 1984, the MCRMLP surveyed 307 Octanet users at the 252 user institutions that had borrowed on Octanet during the October–December quarter of 1983. Responders were asked to rate Octanet functions and services on a scale of 1 to 5:

5 I think it's great.
4 I like it.
3 neutral response
2 I don't like it.
1 I think it's awful.

The user evaluation survey had a response rate of 53.1 percent, with 163 users returning completed surveys.

REQUEST INPUT FUNCTIONS. Results indicated that those users who liked express input or thought it was great (86 percent of those who responded to this survey item) tended to be neutral or unenthusiastic about the more time-consuming regular input function. Users who liked regular input, on the other hand, tended to leave express input unrated (missing value), suggesting that they had not tried it. Although not surprising, these results emphasize that some individuals are intimidated or otherwise put off by a system function that requires the rapid recognition of mnemonics: express input substitutes two-letter mnemonics for regular-input prompts such as "volume number" and "issue number."

ACKNOWLEDGMENT FUNCTION. Second in popularity to express input was the function that permits users to acknowledge receipt of a requested item, thereby enabling the system to report the turnaround time of the request. Of those who rated the acknowledgment function, 75 percent of liked it or thought it was great. This was a quite a surprise to Octanet management, since users had complained at first about having to perform this function in order to get turnaround data. However, the function had been considerably abbreviated a short time before the survey was taken, and it was assumed that this change was largely responsible for the unsuspected popularity of this routine operation.

STATUS CHECK. Sixty-nine percent of the Octanet borrowers who rated the request status check (transaction log display described above) either liked it or thought it was great.

USER REPORTS. In their responses to the user evaluation survey, 66.9 percent of the users who rated the List of Titles Requested indicated that they liked it or thought it was great. Informal discussions with Octanet users have indicated that this report assists in collection development decisions. Both the Borrowers Report and the Turnaround Report received a rating of 4 or 5 from more than 50 percent of survey responders.

It is not immediately clear why the reports received different approval ratings. None received more than a 4 percent disapproval rating (1s and 2s) from those who rated them. However, these user reports received 3s from about 30 percent of responders, indicating no opinion or a neutral response from almost one-third of those who rated the item.

This level of neutral response was characteristic of the responses to all survey items: the average for all items was 38.1 percent neutral responses, with a few items ranging as high as 50–55 percent neutral responses. The average for disapproval ratings (rating of 1 or 2) for all items was only 7.2 percent. This suggests that many of those who completed the questionnaire did not have strong positive or negative opinions about the majority of the functions and reports being rated.

SYSTEM USE

More than 600 institutions have Octanet location codes. These include about 200 online locations, who input their own requests, and more than 400 locations without online capability, who send their requests to their resource libraries for input. Active location codes in an average month include about 170 online locations and about 90 "non-online" locations.

During the past four years, there have been some changes in the volume of usage of the Octanet system, in terms of hours used and in terms of requests input. As table 1 shows, an increase in the latter may or may not be accompanied by an increase in the former. Between 1983 and 1984, there was an increase in the average number of requests per month, but the average number of hours used per month went down. Both requests input and hours used increased between 1984 and 1985, but the increase in requests greatly outweighed the increase in hours. The former was primarily due to the implementation, on January 1, 1985, of the first "mini-Octanet"—that is, the use of Octanet by fifteen St. Louis medical libraries to lend and borrow among themselves as a consortium.

Table 1 also shows the percent of increase in requests each year from 1983 through 1985, again reflecting the effect of the implementation of the St. Louis online consortium in 1985.

It should be noted that the figures for minutes per request were computed from system reports of hours used. These figures include online time for lending functions and electronic messaging. The actual time used for ILL borrowing operations cannot be determined; it can only be stated that the majority of Octanet online usage is for interlibrary loan borrowing activity. Were it possible to determine how much time was spent performing lending operations and using the message function, the figures for minutes per request would be lower than those shown in table 1.

CONCLUSION

In summary, Octanet usage has increased each year, while the time used per

Table 1.

	Average Requests per Month	Percent Increase in Requests	Average Minutes per Request	Average Hours per Month
1982	2619		*	*
1982–83		44		
1983	3768		5.5	348
1983–84		13		
1984	4249		4.3	304
1984–85		38		
1985	5853		3.7	360

*Not available

request has decreased. User response to the system has been predominantly favorable. Although the statement from many users that "my turnaround has been cut in half" has not been documented, an informal consensus would seem to indicate that automated interlibrary loan routing in the Midcontinental Region has been a highly successful venture during the first four years of its implementation.

REFERENCES

1. M. F. Johnson, Jr., and R. B. Pride, "Octanet—An Electronic Library Network: Design and Development," *Bulletin Medical Library Association* 71:184–91 (Apr. 1983). ■■

Reports and Working Papers

Network Advisory Committee Report from December 1986 Meeting

The Network Advisory Committee (NAC) has been engaged in a two-year process since May 1985 to develop a vision statement for networking and action plans for achieving that vision. Toni Bearman, executive director of NCLIS, had asked NAC to assist in redrafting the networking section of the NCLIS document, *Toward a National Program for Library and Information Services: Goals for Action*. This section includes Objective 8: "Plan, develop and implement a nationwide network of library and information services" and "the Nationwide Network Concept." Congress had requested that NCLIS revise the document, and NCLIS determined that while most of the document was still valid, the networking section needed revision in view of technological developments and in view of the fact that "the nationwide network concept" called for a network established from "the top down."[1]

NAC began with a program devoted to the history of networking, held in May 1985, to identify key factors influencing the present and future directions of networks.[2] Barbara Markuson presented an excellent overview of network development, while Sue Martin asked whether networks were really needed. Noel Hanf reviewed legal aspects of networking, while Ron Miller discussed the impact of technology on networks.

There was no consensus regarding the characteristics and goals of networks. Change is the one constant: change in players, change in roles, and change in relationships.

Major issues identified during the two-day NAC meeting were as follows:

- Definition of the goals and characteristics of networks
- Governance and the role and relationships of local, regional, and national networks
- Standards and standardization, quality control for products and services
- Planning and coordination for linking networks, statistics, and planning data for costs, markets, users, and research and development
- Economic and legal issues including ownership of data, restrictions on use of data, competition with commercial services
- Need for strong leadership

NAC members agreed to:

1. Assist NCLIS in developing a strategy, from a networking perspective, to revise the 1975 program document for incorporation in NCLIS programs and plans for the proposed 1989 White House Conference on Library and Information Services.
2. Identify a "common vision" for the networking community to guide future planning.
3. Strongly urge the Secretary of Education and directors of other appropriate agencies to carry out their important responsibilities for gathering and disseminating library and networking statistics.
4. Become a catalyst to convince the library and information community of the importance of networking (i.e., "personalize" it).
5. Address the impact of local systems developments on nationwide networking.
6. Commission a paper on the future of print materials.
7. Urge federal support for networking and library services.
8. Examine networking in other fields and identify potential implications for library and information networks.

These items formed the focus for the next several meetings. The immediate goal was to develop a common vision statement and an action plan for achieving that vision.

A second meeting in December 1985 was focused on developing a common vision statement.[3] JoAn Segal urged that networking be integrated into librarianship and not treated as something independent and apart. Rowland Brown and Dick McCoy both emphasized the growing demands of and role of library users in networking services. Other perspectives were provided by Joe Howard, Henriette Avram, Lois Ann Colianni on national library services; Jim Nelson on state library services; and Lou Wetherbee on regional networks. Jim Gavan discussed local system needs; Charles Bourne, the private sector view; and Barbara Cooper, the user's view.

Prior to the meeting a questionnaire was sent to NAC members. The questionnaire defined networking as:

> A vehicle to assure the nation's libraries and other information providers can, through the assistance of technology, meet the needs of library users through an appropriate combination of:
> 1. local resources and
> 2. state, national, and international resources.

Six basic assumptions were also given:

1. Sharing of resources among libraries is necessary to meet users' needs.
2. The ability to share resources depends on continued participation and the willingness of libraries to cooperate.
3. Evaluation of the short-term advantages of local networking must be done within the framework of resource sharing upon which libraries must inevitably draw.
4. It is neither efficient nor economical for libraries to duplicate effort for certain functions.
5. Technology makes rapid and efficient distribution of information possible.
6. Networking promotes the most efficient and effective use of technology.

Members were asked to give

- their vision of networking
- strategies for achieving the vision
- barriers to achieving the vision

Among the strategies suggested were the following:

- Gain bibliographic control of that not already covered
- Involve publishers and information providers; they play an ever greater role
- Continue parallel paths, but increase coordination among them
- Encourage forums for discussion; implement a series of major discussions as a follow-up to the NAC program
- Reexamine and clarify the role of the players
- Seek an increased willingness to cooperate; analyze the motivation; note the benefits and advantages; determine real costs
- Encourage more involvement of state and national libraries
- Determine the relationship of the local system to the national system
- Control expectation levels
- Involve private, not-for-profit, government, and public sectors.

Among the barriers listed by the respondents were the following:

- Lack of perceived value in cooperation
- Redundancy of efforts; consumption of scarce resources
- Failure of networking to reach the vast majority of libraries
- Failure to establish a mechanism for collecting holdings information from locally processed cataloging
- Lack of standards and protocols, as well as failure to adhere to those that currently exist; incompatibility; reversion to local standards
- Failure to integrate the end-user into network activities
- Lack of fiscal strategies; critical network economic issues need to be addressed
- Controversy over contracts, copyright, and royalties
- Lack of public awareness of network activities
- Lack of understanding of the value society places on information
- Uneven access to information
- Lack of understanding of the politics of sharing; vested and proprietary rights; turfs; territorial rights
- Insensitivity to interdependence
- Technological threat to cooperation

The third meeting in July 1986 was devoted to five major topics in support of the vision statement: standards, linking, education, public policy, and economics.[4] A background paper on each area was sent to members prior to the meeting. Working groups discussed in detail the five areas and identified major issues and potential barriers. Some thirty-eight items were identified (see appendix A). The five areas fell into four categories: actions, studies and research, organizational issues, and communications. The Program Committee reviewed the thirty-eight issues and reorganized them for the December 1986 meeting.

NAC first reviewed and accepted the revised vision statement at the December 1986 meeting and adopted four premises as further clarifying the vision statement (see appendixes B and C). A number of significant points were made in this carefully crafted statement. These include

1. The nation's library: an aggregate of all information resources.
2. Access is important, but must be at a cost supportable by society and the individual.
3. Resource and information sharing is essential to access.
4. There will be no central, monolithic network.

Four task groups were appointed to deal with four consolidated groups of issues: organizational, communication and liaison, program, and major projects. Each group further defined the tasks and assigned one or more individuals to complete the tasks. The tasks and the individual assignments are given in appendix D.

The two-year process has resulted in a vision statement and an action plan for achieving that vision. Among the tasks is Task 24b, which will explore, with COSLA and state library associations, meetings at state and regional levels to review the vision statement and to develop action plans at state, regional, and local levels. Along with Tasks 1, 7, and 9, these steps will provide a solid basis for the revision of the NCLIS document and for the proposed White House Conference.

Appendixes:

A. Issues
B. Premises
C. Statement of a Common Vision in Networking
D. Summary Action Plans

REFERENCES

1. NCLIS, *Toward a National Program for Library and Information Services: Goals for Action* (Washington, D.C.: NCLIS, 1975), p.48–60.
2. Network Advisory Committee, "Key Issues in the Networking Field Today," Network Planning Paper, no.12. (Washington, D.C.: Library of Congress, 1985).
3. Network Advisory Committee, "Toward a Common Vision in Library Networking," Network Planning Paper, no.3 (Washington, D.C.: Library of Congress, 1986).
4. *Library of Congress Bulletin* 45, no.41:345–47 (Oct. 13, 1986). ■■

APPENDIX A. ISSUES LISTS

Communications Issues Identified 860711

1. Articulation of networking and nation's library and resource sharing implications.
2. COSLA/state library organizations meetings on the benefits of networking for uninvolved public libraries.
3. Presentations at national meeting of ALA divisions.
4. An LSP/OSI brochure.
5. Communication of networking standards activity to NAC members and broader networking community.
6. Development of a model program regarding networking benefits for trustees.
7. NAC reaffirm its position on equity of access.
8. NAC promote and publicize activities related to access and information service at the federal, state, and local levels.
9. NAC encourage preparation of papers and meetings on:
   a. Ethics/privacy related to for-profit participation.
   b. Beyond bibliography.
10. NAC update the "Networking for Networkers" brochures.

APPENDIX A. ***Continued***

11. NAC review and prepare a meeting possibly at ALA, on the Office of Technology Report on Intellectual Property Rights.
12. NAC and other groups develop joint meetings on Intellectual Property/Economics/Future of Networking.

Study and Research Issues Identified

1. Implications of sharing/not sharing information.
2. Impact of technology on resource sharing.
3. Research into methods of reducing telecommunication costs.
4. Identify networking standards needed and identify gaps (and prioritize).
5. Survey library schools/curriculum re networking.
6. Archives of electronic data: access/preservation/control.
7. Present and future broad concepts of information networks.
8. Economic definition and scope of networking.
9. Analysis of economic issues reports and organizations activities (CCC) involved with intelligent property issues.

Educational Issues Identified

1. NAC/COSLA state conference on NAC issues/this meeting regional networks do similar.
2. NAC organize continuing education in economics of networking.

Action/Advice

1. NAC support for-profit libraries in multitype.
2. NAC support funding for telecommunication and equipment in LSCA.
3. NAC seeking funding for NAC projects.
4. Formalize representatives on NAC from NISO.
5. Recommend to NISO networking standards.
6. Establish NAC Standards Committee.
7. Establish NAC Organization Committee.
8. Encourage COSLA to issue networking handbook/guidelines.
9. Request SLA inform NAC re implementation of NCLIS/SLA Task Force recommendations.
10. NAC members and member organization participate in activities re access and information services.
11. ASCLA publicize in local papers—benefits re networking, in advance of LSCA reauthorization.
12. COSLA inventory of LSCA benefits/accomplishments.
13. COSLA overview of state and multitype networking plans.
14. LS/COSLA/NCLIS—satellite for educational information services.
15. Establish NAC Committee on research related to networking.

APPENDIX B. PREMISES

1. NAC reaffirms and articulates support for the concept of "the Nation's Library." Such a concept should not imply a physical collection or institution as much as it should access to information regardless of where it resides. To be emphasized here are the resource sharing implications inherent in the concept. NAC should encourage an examination of these implications.
2. NAC underscores the importance of standards in networking. Without standardization the means to access the Nation's Library will be greatly hampered. The networking community must be made aware of standards, how they are developed and adopted, and their centrality in making networking viable. Standards are arguably the most crucial technical aspect of networking, and without them and adherence to them, national networking will remain fragmented at best.
3. NAC emphasizes the importance of equity of access for all the Nation's citizens. Equity of access does not automatically guarantee unrestricted access by everyone to all resources in the Nation's Library but it does imply the availability to all of a certain level of information necessary to support scholarship, recreational use, information needs and interests, and government efficiency. NAC should contribute to the formulation of what such a policy means.
4. NAC's agenda highlights the benefits of networking to the Nation and its libraries. It is through networking that the concept of the Nation's Library can become a reality, providing the means by which users throughout the Nation can have equitable access to information.

APPENDIX C. STATEMENT OF A COMMON VISION IN NETWORKING

A common vision of networking and shared operational objectives is both possible and desirable. However, the diversity among libraries and other information providers and the variety of economic and political factors influencing them is so great as to make impractical a monolithic, nationwide network. To promote the concept of "The Nation's Library" as the aggregate of all available information resources and to bind present and future efforts together in principle and philosophy, the Library of Congress Network Advisory Committee recommends the following statement as a common vision of library and information networking.

Our common vision of networking is an environment in which libraries can provide each individual in the United States with equal opportunity of access to resources that will satisfy their and society's information needs and interests. All users should have access on a timely basis to the information they require without being faced with costs beyond their own or society's means.

To realize this vision, there must be technical and intellectual sharing of resources among the public and private sectors; local, state, and federal governments must fulfill their various responsibilities to individuals and society; and the diverse missions of the several types of libraries must be accommodated. As this vision becomes a reality, there will emerge a diverse but coordinated structure of networks rather than a monolithic one. Active research, rapidly developing technology, collaborative leadership, common standards, and shared communications will provide means by which the system will be further shaped as an interlocking series of local, state, regional, national, and international relationships that are capable of serving the nation's information needs.

APPENDIX D. SUMMARY ACTION PLANS

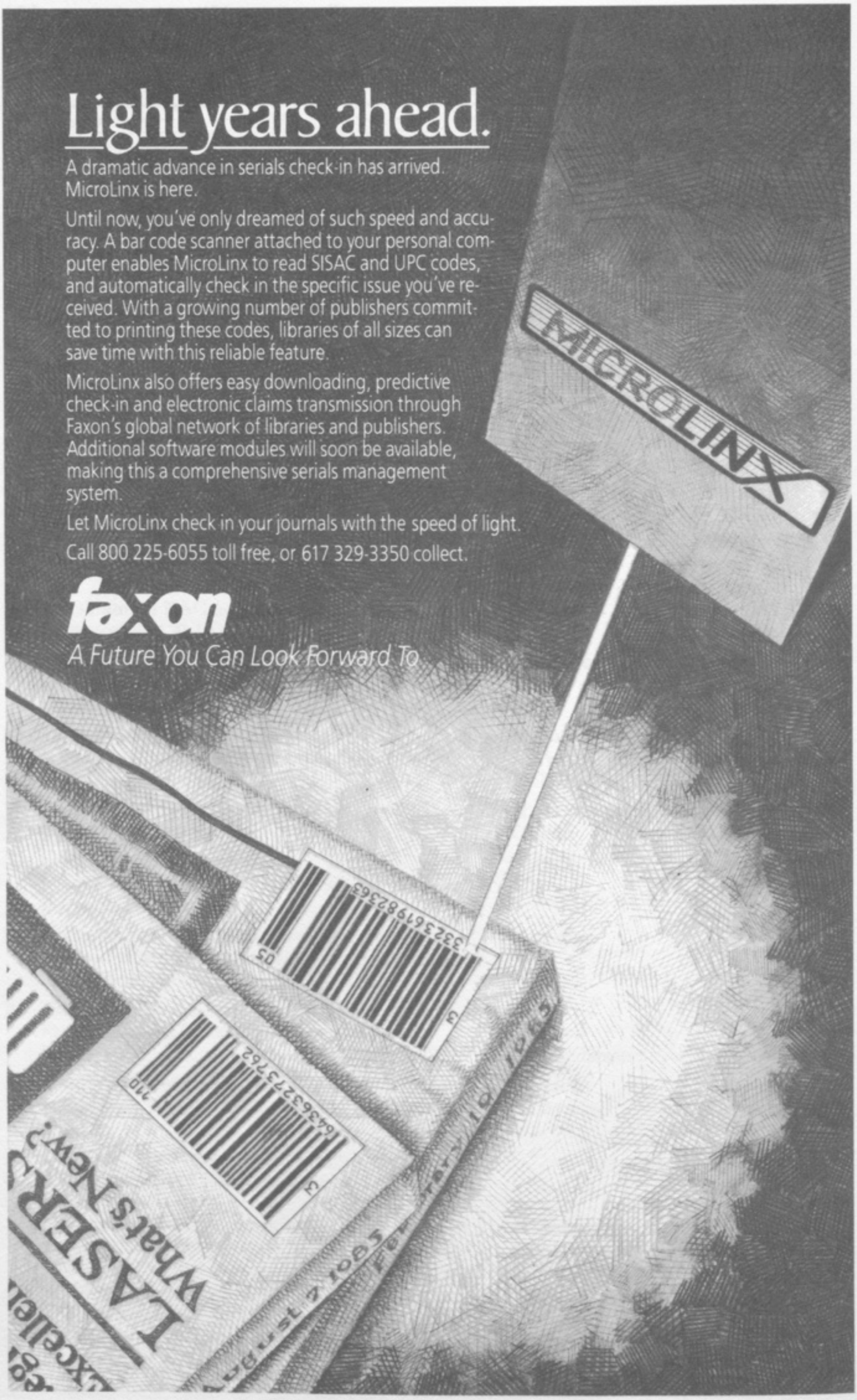
Task	Who	When
1. Create a practical handbook for networking: Plan for Developing	Team 1	Apr. 1
2. Create a model program for Trustees (see also task 1: Plan for Developing)	Team 1	Apr. 1
3. Develop a clear statement of support for including for-profit libraries in LSCA reauthorization	J. Schubert B. Dillehay w/SLA, ALA	Report by Apr. 1
4. Include in LSCA reauthorization funds for equipment: Determine what type of equipment	M. Buckland	Report by Apr. 1
5. An overview and inventory of LSCA multitype plans, programs (relate to table 2)	See tasks 1 and 2	Report by Apr. 1
6. Inventories of successful LSCA program	V. Arterbery	Report by Apr. 1
7. Conference replicating or extending NAC discussion by regional or specialized groups: Identify issues and groups to address	C. Bourne	Report by Apr. 1
8. Regular presentations on NAC meetings at ALA and other national conferences. Combine with 9 and relate to 11.	C. Bourne	Report by Apr. 1
9. Publicize NAC activities through closer cooperation with other organizations: ASCLA, COSLA, state and regional networks	C. Bourne	Report by Apr. 1
10. Publicize NAC activities through participation of NAC members in network-related programs	Team 1	Report by Apr. 1
11. Publicize NAC activities through other mechanisms	Team 1	Report by Apr. 1
12. Invite NISO executive director to report on network-related standards at NAC meetings	S. Paul	Spring '87 NAC Meeting/Review Spring '88 Mtg.
13. Each NAC participant to provide a brief summary of activities: network projects, standards, recent publications, significant and ongoing programs	NAC members	Continuing

APPENDIX D. ***Continued***

Task	Who	When
a. Develop form	J. Riley S. Harriman	March '87
b. Request report with announcement of next meeting	S. Harriman	March '87
c. Chair Q&A sessions on reports at spring meeting	D. Brunell	April 29
14. Increase awareness of networking issues	NAC members	Ongoing
a. Notify Sigrid of appropriate items	NAC members	Ongoing
b. Provide bibliographic citations for items relevant to their constituency	NAC members	Ongoing
c. Deposit reports whenever possible with ERIC	NAC members	Ongoing
15. Prepare proposal and source of funding for a study on cost of library resource sharing and cost of not sharing: letter proposal for CLR	C. Payne M. Avram CLR	Draft by ALA/ Revised Draft by Feb. 15
16. Commission a study on the impact of new technology		
a. Review results of OCLC Users Council Committee studying this topic: Obtain Report	M. Jacob	When Available Report on Status by April 29
b. Obtain for NAC members Fred Meyer Trust Report on CD-ROM technology	M. Stussy	Dec. 11 (Will be sent to Members)
17. Commission a study on reducing library telecommunication costs (too time-, technology-, and area-dependent)	No action	
18. Survey library school networking curriculum: Letter to ALISE asking them to undertake	M. Jacob H. Avram ALISE	Jan. 1, '87
19. Inventory electronic archives: Letter to NCLIS asking them to undertake	H. Avram CCLIS	Jan. 15, '87
20. Commission a study on extension of networks beyond bibliography: Develop a NAC program on this topic as a first step	S. Paul	Spring '88 (Progress Report by Apr. 22)
21. Commission a study on the economic size and shape of information networks (seen as too diffuse and with no real application)	No action	
22. Hold NAC meeting on OTA report, "Intellectual Property Rights in an Age of Electronics and Information"	C. Henderson	April 22–Apr. 24
23. Hold NAC meeting on library school networking curriculum: Await ALISE response on survey (task 18)	M. Jacob H. Avram	Fall '88
24. Hold NAC meeting to discuss state networking developments and role of state agencies in fostering networking		
a. Joint COSLA/NAC meeting	R. Akeroyd	Fall '87(?)
b. Encourage state agencies to hold meeting with state library associations to discuss NAC Vision Statement and implementation at state and local levels	R. Akeroyd, COSLA, State Agencies	Midwinter, Report Apr. 22
25. Commission a brochure on "linking" (not just LSP): Develop a proposal and identify funding sources	M. Gikas K. Schmidt	In '87 (Progress Report Apr. 22)

APPENDIX D. ***Continued***

Task	Who	When
26. Recommend role for NAC in the standards process	NAC members S. Paul	Spring '87
27. Hold networking conference in preparation of CCLIS II and update *Networks for Networking*	F. Grisham P. Andre B. Dillehay B. Oakley	Fall '88 or '89 (Progress Report Apr. 22)
28. Define role and organizational structure for NAC (Ranges from abolish NAC to a funded organization)	V. Arterbery W. Shaw	Apr. '88 with Mtg. in '87 and Draft
29. Develop a prioritized list of networking research needs and communicate it to library schools and research organizations (target groups: CLR, ALISE)	NAC Committee R. Miller	Report '87


Light years ahead.
A dramatic advance in serials check-in has arrived. MicroLinx is here.
Until now, you've only dreamed of such speed and accuracy. A bar code scanner attached to your personal computer enables MicroLinx to read SISAC and UPC codes, and automatically check in the specific issue you've received. With a growing number of publishers committed to printing these codes, libraries of all sizes can save time with this reliable feature.
MicroLinx also offers easy downloading, predictive check-in and electronic claims transmission through Faxon's global network of libraries and publishers. Additional software modules will soon be available, making this a comprehensive serials management system.
Let MicroLinx check in your journals with the speed of light.
Call 800 225-6055 toll free, or 617 329-3350 collect.
faxon
A Future You Can Look Forward To.
MICROLINX
LASERS
What's New?

News and Announcements

CLSI Signs Contract with the New York Public Library and Brooklyn Public Library

The City of New York approved a 7.4 million-dollar contract with CLSI, Newtonville, Massachusetts, to automate circulation operations in the branch libraries of the New York Public Library and the Brooklyn Public Library. The contract calls for a phased implementation of the system into the eighty-one NYPL branch libraries in the boroughs of the Bronx, Staten Island, and Manhattan and the fifty-eight branches plus the Central and Business Libraries of the Brooklyn Public Library.

Initially, the libraries will be installing two separate circulation systems, housed at one data-processing site, LIONS, located on the fourth floor of the Donnell Library Center on West 53d Street. Each of the two institutions has an annual circulation of approximately ten million. Together, they serve a population of 5.5 million.

After completing three-quarters of the implementation, CLSI will install an intersystems link that will enable the two systems to communicate with each other.

CLSI will use an X.25 packet switching network for communication between terminals and the central processors. The configuration has been developed with the cooperation of NYNEX, the parent company of two former regional subsidiaries of AT & T. ■■

People's Republic of China Signs with Washington's Western Library Network

The People's Republic of China has signed an agreement for a Washington State agency to supply the National Library of China with computer software. Under the agreement, China will use software developed by the Western Library Network (WLN) to automate the library's card catalog.

WLN, a division of the Washington State Library, provides computer-based services and products to some 340 participating libraries in and around the Pacific Northwest. In addition, WLN licenses its software to library networks in other parts of the country and around the world. WLN employs approximately seventy people and provides service to several million patrons of public and academic libraries.

China's purchase of WLN software is part of a modernization program being undertaken by the National Library of China. Officials in China have said they hope to have a new system in full operation by 1990.

Formally opened in 1912, the National Library of China embodies the progress of Chinese civilization over the past 5,000 years. Library officials say the facility holds thirteen million items, including books and works of art dating back to the sixteenth century B.C. ■■

Personal Bibliographic Software and UMI Article Clearinghouse Announce Automatic Document Ordering

UMI Article Clearinghouse and Personal Bibliographic Software (PBS) have jointly developed a method by which PBS software can automatically generate document orders for the UMI Article Clearinghouse. The ordering feature enables information professionals to create clearinghouse document orders easily using PBS software for the IBM PC and Apple Macintosh.

UMI Article Clearinghouse is a document delivery service offering copyright-cleared access to more than 10,000 periodicals. Subject coverage is broad, ranging from popular to highly technical publications, with concentrations in business and management, computers and electronics, and medicine.

Pro-Cite, the bibliographic database management system from PBS, now enables the researcher or librarian to generate correctly formatted document orders automatically. These orders, when saved to disk, can then be sent electronically to the UMI Article Clearinghouse through UMInet, UMI's own electronic mail system.

Pro-Cite is a specialized database- and text-management program for bibliographic information. It features variable-length fields and records; powerful searching, sorting, and indexing capabilities; and automatic bibliography production in a variety of formal styles. With the companion Biblio-Link programs, users can transfer records downloaded from online database services directly into a Pro-Cite database.

With the new document order capabilities, PBS software can perform every step of the process from online search to electronic document order. Using Pro-Search, information professionals can find and download the citations they need, and then use Biblio-Link to convert and transfer the records to a Pro-Cite database. From Pro-Cite, they can format the citations as Clearinghouse orders for transmission to UMI. The high-quality document copies are then shipped by UMI within twenty-four to forty-eight hours of order receipt by first class, overnight mail, or facsimile.

Document ordering is also available through Pro-Search, using the document-ordering features of the BRS and Dialog online services. In addition, Pro-Search enables a user to set up the clearinghouse as the default document supplier. Pro-Search is the online searching aid from PBS, providing easy, cost-effective searching of BRS and Dialog databases. ■■

Circulation Added to Innovacq Library System

Innovative Interfaces, creators of the highly acclaimed Innovacq Library System for acquisitions and serials control, has developed a circulation module for the Innovacq/Innopac system. The Innopac circulation module is now in operation in beta test at Uniformed Services University of the Health Sciences in Bethesda, Maryland.

The Innopac circulation module is the fourth major component of a library system chosen by more than seventy libraries throughout North America. The four Innovacq/Innopac modules—acquisitions, serials control, online catalog, and circulation—can operate independently or together as an integrated system.

Innovative Interfaces of Berkeley, California, has developed and marketed library automation products since 1978. Its library systems and interfaces can be found in more than 150 libraries. ■■

The British Library Awards Contract for Catalog Conversion

The British Library has awarded a major contract for computerizing its general catalog to Saztec Europe. Work started in February and will be completed early in 1991. The data will be keyboarded from the 360 volume-printed edition in Ardrossan, Scotland, and about seventy staff are expected to be employed by Saztec on the project.

When the conversion is completed the machine-readable catalog will be among the largest single bibliographic databases in the world. Readers in the British Library reading rooms will have direct online access to all the records of the library's uniquely rich collections, linked to an automated system for ordering up the books they select from the catalog. Remote users will be similarly able to consult the catalog through Blaise-Line, the library's existing online information retrieval system.

The library considers a computerized catalog of its holdings, with benefits both to managing the collections and the provision of services to users, to be essential when it starts to occupy its new building at St. Pancras in 1991.

Saztec Europe has also been awarded exclusive worldwide rights for ten years for publishing the converted catalog on CD-ROM or other optical storage media. The British Library will receive royalties from the sales.

The net cost to the library for the achievement of the converted catalog is about 1.8 million pounds at current prices spread over four years. ■■

DDC Editorial Support System Installed at Library of Congress

A microcomputer system has been installed at the Library of Congress to support editorial work on the next edition of the Dewey Decimal Classification. The system is the result of four years of intensive planning and development by Library of Congress staff, Forest Press (publisher of the Dewey Decimal Classification) and Inforonics, a Massachusetts computer software and systems firm specializing in bibliographic systems. Installation of the system began at the Library of Congress in June 1986 and was completed in November.

The system has three basic modules: input, search, and editing. Editorial staff may access all words and numbers, in the schedules, tables, and index of the Dewey Decimal Classification. Tag-prompted workforms and full-screen editing capabilities are provided. The software, which is written in the *C* programming language, runs on an AT & T microcomputer using the Unix operating system.

The system will enable editorial staff in the Decimal Classification Division of the Library of Congress to edit and update the classification quickly and efficiently, to produce exhibits for advisory committee meetings, and to generate a computer tape that will be used to photocompose both the unabridged and abridged editions of the DDC. ■■

Princeton Signs with Carlyle

Princeton University has contracted with Carlyle Systems, Inc. to purchase an automated library system. Public access terminals were installed in the university's Firestone Memorial Library in the first phase of providing online catalog access. Carlyle's cataloging input/edit and authority control subsystems will be added in the next phase of the installation.

Princeton had tested a smaller Carlyle system for several months before making the decision to purchase a full system. ■■

Northwestern and TBG

Northwestern University and TBG, Inc. have agreed to discontinue their discussions regarding TBG's acquisition of Northwestern's NOTIS system. NOTIS, a library automation software package for research libraries, was first developed in the early 1970s. ■■

INNOVACQ Posts Strong Sales

In 1986, new sales of the INNOVACQ library system have brought the total number of installations to sixty-five, representing more than a 35 percent increase over the number installed at the end of 1985.

New 1986 INNOVACQ installations include

Adelphi University Library, Garden City, N.Y.

American University Law Library, Washington, D.C.

Arizona State University Library, Tempe

Ball State University Library, Muncie, Ind.

California Polytechnic University Library, San Luis Obispo

Federal Reserve Bank Library, New York, N.Y.

Graduate Theological Union Library, Berkeley, Calif.

Northeastern Illinois University Library, Chicago

Indiana University of Pennsylvania Library, Indiana, Pa.

San Jose Public Library, Calif.

University of Arizona Library, Tucson

University of California–San Diego Library

University of California–San Francisco Library

University of Dayton Library, Ohio

University of Nevada Library, Reno

Westchester Library System, Elmsford, N.Y.

Winthrop College Library, Rock Hill, S.C.

INNOVACQ is a multiprocessor turnkey system developed by Innovative Interfaces Inc. of Berkeley, California. ■■

Recent Publications

Book Reviews

Cook, Michael. *The Management of Information from Archives.* Aldershot, Eng., and Brookfield, Vt.: Gower, 1986. 243p. (ISBN 0-566-03504-9).

For many years now, it has been evident to American archivists that their profession was changing. While, for most people, the words *archives* and *archivists* still conjured up images of musty repositories maintained by equally musty staff members, those in the profession knew that a revolution was occurring. The development of records management programs coupled with the application of computer technology to archival services were just two indications that the theory and practice of archives was undergoing a thorough reassessment.

This reevaluation was not taking place solely in the United States. Archivists around the world grappled with the issue and came to many of the same conclusions as their American counterparts. The new ethos is admirably set forth in this new book by Michael Cook.

Cook, university archivist and senior lecturer in archives studies at the University of Liverpool, opens his study by observing that many of the practices and techniques he describes are not new. What is different, however, is a "profound change in attitudes and values . . . which are derived from newly evaluated goals and objectives." Cook applauds the movement away from isolating archivists and archival administration and forging closer ties with library and information sciences. He notes that "today the world of archive administration is becoming much more open to ideas from the world of the other information services." This belief—that the archival profession should align itself with other information management services—is shared by archivists on both sides of the Atlantic. Many professionals feel that for archival services to prosper they must be perceived not only as historical study sources but also as providers of planning information for business and government. With this in mind, Cook sets forth a new definition of archives:

> Archives may be defined as information media which have been generated from within the organisation, and the management of which has been delegated to a specialist service. The purpose of this delegation is the preservation of the materials and the exploitation of the information in them.

The remainder of his study describes and explains "the practices of records and archives management, in the light of an enhanced emphasis on the provision of information as a commodity to a body of users." In a succession of chapters, Cook discusses the history of archives services, records management, acquisition and archival appraisal, archival description, and the application of computer technology to archives services. Though Cook draws most of his examples from the British archival tradition, his analyses are sound, pragmatic, and applicable to practices in the United States. Particularly noteworthy are the chapters on archival arrangement and description, which benefit greatly from Cook's work with the Archives Description Project at the University of Liverpool as coauthor of its report, *Manual on Archival Description*. The need for common standards and practices in archival description has long been a concern of American archivists. Important also is Cook's call for investigation of user response to archival finding aids. In fact, Cook devotes the final chapter of his book to a discussion of user services, noting that "by comparison with other information services, archives have not given much emphasis to the study of user needs."

The chapters on the application of computer technology to archival services offer a general overview but have two drawbacks. The first is that the majority of examples cited are taken from British application, and, on the whole, the British archival profession does not seem to have made as much progress in this area as its American counterpart. Second, advances in the field are occurring so rapidly that, by necessity, discussions of this nature are dated within a short time. But these hindrances should not detract from the usefulness of the chapters as a beginning point for the study of computer systems within an archival setting. Moreover, Cook deserves a special mention for noting the importance of considering staff attitudes and training when planning for a new system. As he succinctly observes, "Staff indifference or even hostility are often sufficient to cause the failure of the system."

Ultimately, the value of this book is the fact that it forces the archival specialist to reexamine familiar theories and practices in a new light. As society demands more and more from its information services, archivists will have to break free of the constraining image of the past and "push towards viewing archives work as a branch of information management." To refuse to do so relegates the profession only to the past with no place in the future.—*Nancy F. Lyon, Yale University Library, New Haven, Connecticut.* ■■

Davis, Steve, and Candy Travis. *The Electric Mailbox: A User's Guide to Electronic Mail Services.* Dallas, Tex.: Steve Davis, 1986. 290p. paper, $19.95 (ISBN 0-911061-14-2).

This book is a very useful tool for those who have never used electronic mail (e-mail) or for those who want to compare or select a new service. The authors provide an overview of electronic mail communications, guidelines for choosing a service, and descriptions of necessary hardware and software. Steve Davis and Candy Travis, experienced e-mail users, have been writing and publishing books on computer-related topics for more than four years.

The chapters on hardware and software are a good introduction to and overview of the telecommunication process. In addition, a shopper's guide is provided for modems and telecommunication software packages for the most popular PCs. There is a brief description of using e-mail on the road, which includes the use of portable PCs and acoustic modems for pay phones and hotels and motels that have hard-wired telephones.

The most useful information in the book describes and compares the major electronic mail utilities: Source, CompuServe, Delphi, ECHO, Dialmail, MCI Mail, EasyLink, InfoPlex, GTE Telemail, RCA Globcom, OnTyme, GE Quik-Comm, and Dialcom. Each chapter provides an overview of the e-mail utility, subscription and rate information, and step-by-step instructions for accessing and using the service. Easy-to-follow procedures help the user access the service, send and read mail, use mailing lists and electronic filing, and upload mail. Additionally, there is a review of the most common commands and shortcuts. Voice mail and developing communication technologies such as FM radio signals and satellites are described in succeeding chapters.

Telephone numbers for telecommunication networks are listed in the appendixes. Additionally, a special free bonus is offered by the authors and publisher: more than $200 worth of discounts and special introductory offers on e-mail and online database services, communications software, modems, and computer books.

The book is clear and easy to comprehend and use; however, the typeface and the poor print quality contribute to eye strain. This work is a very good introduction to e-mail and comparative information on the major e-mail services. And, although it is informative, it will be outdated very quickly as e-mail services constantly change. For example, the book is dated 1986, but it already contains some outmoded information, such as the discussion (and listing of telephone numbers) for Uninet, which is now a part of Telenet. The same is true of the discussion of ITT Dialcom, which is now owned and operated by British Telecom. However, I would highly recommend this book as a good introduction to e-mail and a comprehensive comparison of major e-mail services.—*Jenny McGee, MEDLINK, New England Journal of Medicine, Waltham, Massachusetts.* ■■

Kantor, Paul B. *Costs of Preservation Microfilming at Research Libraries: A Study of Four Institutions.* Washington, D.C.: Council on Library Resources, 1986. 32p. paper, $3 prepaid.

This study, commissioned by the Council on Library Resources, analyzes the costs associated with preservation microfilming in four very different library settings. Microfilming programs at the University of Chicago, Columbia University, Library of Congress, and New York Public Library were examined. Using data collected through work samples and logs, as well as information supplied by the libraries and gathered through site visits and interviews with key personnel, Paul B. Kantor has identified the range of costs for eight separate tasks in the process. These tasks are identified as: (1) decision making, including bibliographic searching; (2) physical preparation for filming; (3) record keeping; (4) preparation of bibliographic targets; (5) filming; (6) film processing; (7) quality control checks for film resolution, density, and textual completeness; and (8) administration, including such activities as training, personnel management, and secretarial duties. (The tasks of identification and retrieval of materials to be filmed are not analyzed.) Although confidentiality requirements precluded any presentation of tabulated, comparative results for these eight tasks in each library, a range of costs is provided, and the library with the lowest cost for each task is identified. Thus, the procedures of the library with the lowest cost for a given task could serve as a possible model for greater efficiency.

The study found the greatest unit-cost variations in the library-specific processes of decision making, physical preparation, and review, with smaller variations in the costs of actual filming work. As the costs for record keeping and administration were found to contribute more than 35 percent of total labor costs, these managerial areas appeared to hold great potential for savings. Kantor concludes that greater efficiency can be found in the development of large processing centers and careful attention to the problems of record management and administration.

Despite the mass of data collected and the useful, conceptual framework that divides the overall process of preservation microfilming into individual tasks, the lack of specific details concerning each library's procedures weakens the study's conclusions. Thus, a major question of the nature of the standards for preservation microfilming and the uniformity of following these standards is not addressed. If all of the libraries studied are producing an archival negative, a master negative, and a positive service copy and all are following the same production and quality control standards, then the data for each library should be comparable. However, if they are not, and a distinction between microfilming and preservation microfilming can be drawn, then comparisons of data collected become questionable. Similarly, the exact nature and differences between record keeping and administrative procedures at each library are left quite vague. Although it is posited that cataloging costs, including the minimal-level cataloging done at the Library of Congress, are not included in the study, the analysis of record keeping *does* include time spent entering data into the RLIN database. Unfortunately, the details of this aspect of record keeping remain unclear in the study. As the variation in record keeping is the largest cost variation found in the study, the exact nature of this task and any differences between each library's procedures must be examined carefully. Are searching functions combined with queuing of records in RLIN? Is only queuing considered? or, Is the updating of queued records also included? What impact do these record-keeping variables have on administrative costs, especially training? Without clear distinctions between tasks and full details of procedures followed to accomplish them, it may be premature to identify a preferred method for efficiency.

Additional, fundamental questions must be addressed and differences between materials being microfilmed should be explored before broad conclusions on greater efficiency are drawn. Are there significant differences between policies for retention of the original item after microfilming, and what effect would such differences have upon record keeping and filming speed? Are such policies dependent upon subject parameters of materials being filmed? How much record keeping is involved in the replacement of missing pages, and did the rates for this specific problem differ between libraries? These kinds of questions do not negate the value of Kantor's study, but, rather, underscore it, for it is precisely this kind of extensive cost analysis that encourages an examination and improvement of current procedures. Any lack of specificity in the delineation of individual tasks should only encourage and not hinder further investigation of cost-effective procedures. Moreover, the revelation of cost-effective methods in certain tasks, such as machine production of eye-visible targets at the Library of Congress, is not obscured by numerous potential variables.

Costs of Preservation Microfilming at Research Libraries: A Study of Four Institutions is, therefore, a helpful work for any research library that is pursuing or planning a preservation microfilming program. Given the magnitude and urgency of the problem of preserving the intellectual contents of millions of brittle books, the need for cost efficiency cannot be ignored. Kantor's study provides a useful methodology for examining costs, identifies key tasks in the process of preservation microfilming, and offers suggestions for promoting greater efficiency and further research.—*Ted Kuzen, University of Virginia Library, Charlottesville.* ■■

Robinson, Lawrence. *The Facts on Fax.* Dallas, Tex.: Steve Davis, 1986. 274p. paper, $19.95 (ISBN 0-911061-15-0).

When dealing with a new technology, it is often difficult to glean quickly enough nontechnical information to understand the field. Lawrence Robinson relieves this problem for telefacsimile with *Facts on Fax*, "a how-to book about the new high-tech world of facsimile communications . . . , a simple, yet comprehensive guide that can show you how to design and build an effective facsimile . . . system." In meeting this description of his book, Robinson draws upon his experience as coordinator of South Central Bell's large telefacsimile network, although his expertise never limits his use of clear, simple language to explain concepts and features. Many readers will be grateful that prior knowledge of communications technology is not needed, but many will also wish that the text would progress into more technical details at times. As might be expected from the author's background, the book is oriented toward the business world, a focus that does not prevent it from being of value to those examining telefacsimile in a library context.

The first four chapters of the book present the author's convincing and well-stated arguments for the use of telefacsimile. He rejects the view that telefacsimile is only a temporary "bridging" technology to an imminent paperless society. Instead, he argues that it is a very flexible and accessible form of electronic mail for a world that will not soon abandon its paper documents. Building on this premise, he compares telefacsimile to a number of rapid-communication alternatives. Many libraries may not be ready to agree with Robinson's business-oriented assumption that a rapid-communication system is essential and cost-effective. His information and evaluations, however, will aid libraries in deciding if use of such a system is justified.

The remaining chapters of the book are devoted to selecting and operating telefacsimile terminals. Chapter 5 presents the first steps to be taken in choosing a terminal. It deals at some length with whether to pick the older, slower, analog technology available at lower prices or the newer, faster, and more versatile digital machinery. This is followed by a discussion of the new terminals that code and store data for delayed automatic transmission and have the capability to interface with computer systems. The chapter concludes with a rather brief discussion of whether to lease or buy equipment. Chapter 6 deals with specific features available on terminals. Here Robinson's ability to write in layman's terms is especially valuable. Considerations such as power needs, machine size, and service agreements are also covered. After chapter 7, which is devoted to designing a telefacsimile network, chapter 8 presents a great deal of information on selecting a vendor, stressing the questions to be asked before signing a contract. The following two chapters discuss running a network and upgrading network machinery. As with chapter 7, the subject matter may seem more applicable to a business than to a library. Libraries, however, can extract valuable ideas that may improve efficiency and cut costs. A final chapter examines future possibilities for the technology. Of particular interest is the examination of the predicted capabilities and limitations of the new generation of Group IV terminals.

The book concludes with 77 pages of what is termed a "bonus section." This contains Robinson's personal evaluations of specific terminals. He has grouped the machines according to the volume of activity he feels they can best handle. Even though it contains only a sampling of available machines and the models presented will be superseded, this section is useful as it shows the features discussed earlier as they appear, or fail to appear, on actual terminals. This section is illustrated with photographs of the models reviewed.

Robinson's unquestionable enthusiasm for telefacsimile leads him to adopt a tone that reads more like a sales pitch than a detached evaluation. Worse, it leads him to pass lightly over such problems as telephone signal distortion and automatic document feed malfunctions, leading one to assume that the technology is largely trouble free. The book's style is weakened by the repetition of arguments and information when one clear exposition with later references would be adequate. These flaws, however, do not destroy the value of the work and one hopes a better-organized and edited version will eventually appear. For detailed, library-oriented information, this work cannot replace Judy McQueen and Richard Boss' "High-Speed Telefacsimile in Libraries," *Library Technology Reports*, 19:7–111 (1983). Instead, *Facts on Fax* complements McQueen and Boss with comprehensive, general, and current information presented in a very accessible manner.—*Steven A. Brown, Science Library, University of Georgia, Athens.* ■■

Saffady, William. *Optical Disks for Data and Document Storage.* Westport, Conn., and London: Meckler, 1986. 94p. paper, $29.95 (ISBN 0-88736-065-3).

Optical Disks for Data and Document Storage is a survey of products for data and document storage. It is the second in a series of annual surveys of optical disk technology product developments and applications. The focus is on products and services available since 1984 in each of three application areas: Optical read-only memory (OROM), read/write optical media and drives, and turnkey optical disk systems.

The work is divided into three parts based on the three application areas. Part 1 discusses "Read-Only Optical Disks," defined as optical disks that must be created by a mastering process and, therefore, do not support direct data recording by users. They are well suited to database dissemination, software distribution, and technical publishing. The digital videodisc is the first technology included. There is a detailed, technical description of the medium, including the problems associated with storing digital data on an analog medium. Specific companies offering premastering services are identified, and specific equipment needs are noted. CD-ROM technology is also described in detail. There is a comparison of CD-ROM and optical videodiscs that is valuable for understanding their differences. Specific system software for using each technology is identified and described briefly. A variety of applications is described in detail, including specific vendors and development costs. One of the most interesting aspects of the discussion is the comparison with magnetic storage, micrographics technologies, and online services.

"Read/Write Optical Media and Equipment" is the focus for Part 2. Read/write optical disks can be used for the direct recording of information from a variety of sources including magnetic tapes and video scanners. Both analog and digital optical disk systems are discussed in the context of their usage in the broad area of information management applications. In the section on recording technology, there is a technical description of commercially available recording technologies and a general discussion of the emerging erasable optical technology. Magnetic-optical recording, which is in the research stage, is also discussed. On the topic of read/write drives, data from both research reports and commercial product development are included. Specific vendor products for microcomputer, minicomputer, and mainframe computer installations are described. In the discussion of read/write optical disk applications, emphasis is on the complexity of systems implementation. Currently applications are typically custom developed for particular clients by systems integrators who have the engineering expertise necessary for a successful installation. A number of these companies and their projects are identified.

Part 3, "Turnkey Optical Disk Systems for Document Storage and Retrieval," begins with a discussion of the "turnkey concept." Defined as preconfigured combinations of optical disk and computer components designed for the office environment, these are far less expensive than customized systems.

The section describes the characteristics and capabilities of a number of available systems and compares them with computer-assisted microfilm retrieval (CAR) systems. The discussion focuses on specific elements of optical systems including scanners, optical recording media, document display and printing, software, and other features. For each element there is a description of the hardware and operational characteristics, with information on specific vendors as well.

This publication packs a large amount of information into very few pages. The technical descriptions of the various optical media are easy to understand, and the inclusion of vendor products and services is very useful. Although aimed at technical specialists, the work can be used by those without a specific technical background. Citations to specific works in progress are found in each section and are very useful for gathering additional information. There is also a substantial bibliography.

As with any publication dealing with new technology, timeliness is a problem. Most of the information cited is from 1984 and 1985. In the two years since, the technology has changed and so have the costs. However, this should not significantly reduce the usefulness of the work, since most of the technical descriptions are basic to understanding optical technology.—*Pamela Q. J. Andre, National Agricultural Library, Beltsville, Maryland.* ■■

Small Scale Bibliographic Databases. Ed. by Peter Judge and Brenda Gerrie. Library and Information Science Series. Sydney, Australia: Academic, 1986. 198p. (ISBN 0-12-391970-3).

This book is a very useful compendium of practical experience in the creation of relatively small-scale bibliographic databases. The word *relatively* in the sentence above is carefully chosen. Some of the applications and databases discussed in the book are definitely nontrivial. The title of the book is therefore rather too modest, almost misleadingly modest. The book can be profitably read by persons interested in more than small-scale applications. That focus is both the book's strength and its weakness.

The individually written chapters on topics relating to the creation of databases are all of rather high quality for a book of this sort. Some of them, like "Subject Control" by Brenda Gerrie, are in fact of very general applicability, not merely to small-scale systems. The inclusion of such pieces however is not readily apparent from the book's title. Conversely, the reader who browses in this book expecting to find specific information comparing hardware and software options for microcomputer implementation is going to be disappointed. Most of the applications used as examples are in fact minibased, not microbased.

The real topic of the book is more the creation of relatively small abstracting and indexing databases than the creation of corporate or in-house databases. If approached from that point of view, the book is a good read, presenting some imaginative database applications. A further attractive feature of the book is that it is written by Australians, and therefore the examples used are, for us non-Australians, fresh and interesting.

The chapters entitled "Small Scale Databases," General Overview of a Database System," and "Management with a DBMS" provide a good overview of data systems and data design and are particularly recommended. That overview successfully serves two purposes: as a review from which the relatively experienced can profit and as a good introduction for the novice. Its only major failing is that it does not describe for the uninitiated what the differences are in capability between the DBMS systems that originate in the business data-processing tradition and the data management and retrieval systems that originate in the information-retrieval tradition. These differences are very great and very important, for they are in effect reflections of the capabilities of the packages developed in those traditions. This failure is a distressingly common one, however, found in most books on the topic, not just this one.

Overall, the book is well worth an examination by anyone involved with the design or establishment of "relatively" small-scale bibliographic databases.—*Michael E. D. Koenig, TRADENET, New York, New York.* ■■

Nonprint Reviews

Textbank/PC. Produced by Group L Corp., 481 Carlisle Dr., Herndon, VA 22070.

Hardware requirements for the package under review are IBM PC/XT, IBM PC/AT, Compaq, ITT XP, ATT 6300, Zenith Z200, Wang PC in IBM emulation mode, or any other 100 percent IBM-compatible personal computer; operating system DOS 2.0 or later version with a hard disk; 640K memory; IBM-compatible monochrome or color monitor with color adapter, enhanced or professional graphics adapter. $995.

The first thing that strikes the user of the text retrieval software Textbank/PC is that it does not resemble any type of PC-based DBMS software currently on the market: it is not a dBase type product; that is, it is not a relational database with an emphasis on record management functions, such as report generation, updating, or labeling. Nor does it bring to mind any of the software specializing in bibliographic file management—formatting records according to a particular bibliographic style, simple searching of ASCII files, downloading records from online vendors and so forth. Finally, in many ways Textbank/PC differs significantly from its competitors—text retrieval software like In-Magic, SIRE, Stairs, BRS/Search, and Finder.

Why is Textbank/PC so different and what is the advantage of this difference? Let's begin by understanding what text retrieval software does and what we librarians want it to do. That is, what are the requirements for good text retrieval software in the bibliographic environment? As its name implies, text retrieval software (TRS) must handle text well. For librarians this may mean a number of things, but the stuff of our everyday experience is pretty well defined by a variety of bibliographic records—MARC records, online bibliographic citations with their fixed- and variable-length "fields"—and by a few means of "capturing" these records—downloading ASCII files or manual input using a text editor. For those of us who dream of ways to automate the entire process of finding information, good TRS simplifies the downloading of machine-readable ASCII files into the DOS environment of our hard drives and then allows us to search and retrieve the filed information in a manner with which we are accustomed—for example, through the command languages and search/retrieval algorithms of an online service like Dialog or BRS. Throw in the ability to manage well the DOS environment of the hard drive for quick response, and you have a great TRS package.

Until recently, the only candidates for this TRS pantheon were the various products developed for mainframe computers and now marketed in PC versions. In-Magic, Finder, BRS/Search, Stairs, SIRE—all are excellent TRS packages that have journeyed from mainframe to micro, bringing with them the unique advantages of mainframe-searching and database-building capabilities. Unfortunately, many of these packages still bear some of the disadvantages of their mainframe parents—difficult command language and even more difficult and time-consuming procedures for the importing of foreign files, which require the creating of a database according to predesignated field arrangements. Whenever this is the case, text editing, a process that is always time-consuming and often moderately difficult, is always the rule.

Textbank/PC changes all of this. Using distinctively new programming that allows

Textbank/PC to instruct itself to recognize textual "zones" of an ASCII file, Textbank/PC enables us to create databases "after the fact." Coupled with full-text search/retrieval software that nearly matches the power of Dialog or BRS and is considerably easier to learn and use, Textbank/PC clearly offers a unique, and in many cases, superior TRS product.

BUILDING AN INFORMATION BASE

Textbank/PC supports two types of database—the designers prefer the term "information base"—building. Information bases may be built from scratch, using pattern recognition rules, or they may be constructed using preexisting templates supplied by the program or used in a previous information base. I will illustrate briefly the steps in building each type.

Building an information base from an existing template is quite simple. You begin with an ASCII file—for instance a PsychInfo search downloaded from BRS—located on any drive or hard drive directory. At the main menu you select the building option and the program leads you through a series of steps—locating the target file, specifying stop words and so forth. At this point you also select which template you need to use. The present version of the program supports all Dialog databases that offer the tagged field format (format 4) and all BRS databases. Once a template is selected you can instruct the program to accept all fields of the database, or you may select specific fields that you have downloaded. After this step, the program builds the information base. For my information base of 388 brief records (author, title, source, descriptors; about 170,000 bytes) from BRS' PsychInfo database on a PC XT (640K), it took just under two hours to build, and now it occupies about 400,000 bytes of the 20 MGB hard drive—a 116 percent overhead.

Building an information base from scratch is a little more difficult, but not much. The first step is to study the textual structure of the document, in this case, a communication research study guide within an information base of library study guides. The guide was typically structured: entries were listed by section—indexes, yearbooks, etc., and an index was also included.

Now the idea of Textbank/PC is to define textual zones that may be inclusive or exclusive. In the communication guide we have the guide itself (a guide zone); the various sections (section zones); the entries within each section (entry zones), which were further subdivided into call number and summary areas (call number and summary zones); and the index (index zone).

Graphically the structure of zone relationships could be expressed,

```
GUIDE (information base name)
  Guide
    Section
      Entry
        Call number
        Summary
      Index
```

where entries at the same level of indention are mutually exclusive zones and indented entries are inclusive zones.

Once you have a sense of how a text is analyzed into zones, the next step is to enter the build program and go through the steps by which the program parses the file using a number of pattern recognition rules. For example, to recognize where a section zone begins, you instruct the program to look for a line of boldface letters indented twenty-four spaces from the margin. To instruct the program where the zone ends, you either select another pattern rule or elect "peer"—an instruction that indicates a zone ends where the next instance of the section zone begins. Once you have selected all rules for zone recognition, the program checks over your selections and notes any errors in syntax construction. You also have the option to examine the ASCII file to see how your rules have "hit," that is, which textual zones have been created. The same function also allows you to view the ASCII file as Textbank/PC sees it—filled with ASCII characters that indicate spaces and other codes—control codes from word-processed documents for example. After checks have been made you are ready to build. In this case my 33,000-byte study guide took about twenty-four minutes to build and occupies about 75,000 bytes on the hard disk—a 133 percent overhead.

SEARCHING AN INFORMATION BASE

For seasoned online searchers, searching an information base is easy and actually quite fun. For the novice, a very well thought-out command language and instructional manual take a sensible approach to full-text and zone delimited searching and should present few difficulties.

As should be apparent, the search/retrieval design philosophy behind Textbank/PC was to develop a fully searchable text analyzed into textual zones for potentially more precise retrieval. As a result, both in building information bases and searching them, the emphasis has been on full-text retrieval—every meaningful unit can be searched.

To this end, Boolean and proximity operators are supported as well as full truncation (prefix, suffix, and infix), expansion, and range searching. The Boolean operators are the standard AND, OR, NOT, as well as the oddly useful XOR—either *x* or *y* but not both. Proximity can be by word or zone—for example, the term *communication* within three words of the term *anxiety* or *bible* within two sentences of *catholic*. Three wild cards are used in truncating—an any-character wild card, a one-character wild card, and a range wild card (e.g., 198 [5-7]). Terms can also be expanded through truncation and then a macro made to consolidate them all in one search term, for example, all forms of *river, canal, lake, stream* can be merged into the single macroterm *water*.

As with online retrieval software, search steps can be combined by Boolean operators and delimited by zone. Search sets can also be sorted, saved, downloaded to disk or printer.

Searching the PsychInfo information base, then, was little different than searching on BRS. Terms can be limited to specific zones that correspond to BRS fields or searched full text. There are some limits to the number of combinative sets in the same work area—you can only combine five sets at any one time. Zone searching also presented the peculiarity that if you search the descriptor zone you may initially only display the retrieved zones, requiring you to use the zoom-out function to view the wider context—that is, the complete record (in the template all fields are structurally equal and are subordinate only to the entire record). Generally, however, as an experienced online searcher I had no problems in searching the information bases I had created.

COMMENTS

Group L's insistence on the term *information base* is appropriate when we consider the design and applications of Textbank/PC. As I mentioned earlier, Textbank/PC specializes in text retrieval, not records management or bibliography, although in some respects it could be so adapted. Moreover it is designed to work in the PC environment by users who want easy yet powerful access to their information but who are not necessarily trained in the convoluted complexities of mainframe languages or do not have the time, inclination, or experience for fancy database maintenance procedures.

Textbank/PC has a few drawbacks. Proximity searching could be made bidirectional and more flexible (presently you must OR together the searches "*A* within six words of *B*," "*B* within six words of *A*" to locate *A* and *B* within six words of each other, no matter what order the terms). The work area could be expanded to include more than five sets at a time. Editing of information bases is also somewhat cumbersome, requiring deletion, exiting to a text editor, and then reappending the records. Finally it would be nice to improve the already excellent online and offline help (both the manual and a great customer service department) by developing some form of programmer's toolkit or online accessible program that would help in the difficult steps of creating an information base from scratch—perhaps a reveal-codes program that would in turn help select appropriate pattern-recognition rules.

By this last comment I do not mean to suggest that Textbank/PC is difficult to use—quite the contrary. PC user-friendliness is Textbank/PC's major virtue. From the simple installation procedure through information base building and searching, Textbank/PC renders a very complex process very simple—just take a look at the file structure in the root directory Textbank/PC creates. Yet with this simplicity you receive state of the art database software with mind-boggling possible applications. I plan to use it for a subject specialist inhouse database of select articles for communication research, machine-readable bibliographies for a PC at the reference desk, and for a general database of book and article citations that I can search and then upload in ASCII format to other software programs. I have also shown graduate students with whom I work the principles of online searching—without the costs of going online.

I am sure readers will find other applications. Why not? With Textbank/PC any ASCII file, no matter when or how it was created, is a candidate for a fully searchable database.—*Dennis Brunning, Arizona State University, Tempe.* ■■

Other Recent Receipts

Listed here are books and other publications received for review that are of potential interest to LITA members. Some of these materials may be reviewed in later issues of ITAL.

Breslin, Jud. *Selecting and Installing Software Packages: New Methodology for Corporate Implementation.* New York: Quoroum, 1986. 242p. $39.95 (ISBN 0-89930-158-4).

Brophy, Peter. *Management Information and Decision Support Systems in Libraries.* Aldershot, England, and Brookfield, Vt.: Gower, 1986. 158p. $39 (ISBN 0-566-03551-0).

Computers and Computing Information Resources Directory: A Descriptive Guide to Approximately 6,000 Live and Print Sources of Information on General and Specific Applications of Computers and Data Processing, Including Associations and User Groups, Consulting Organizations, Research Organizations and University Computer Facilities, Libraries and Information Centers, Trade Shows, Exhibits, and Association Conventions, Online Services and Teleprocessing Networks, Directors, Journals, and Newsletters. 1st ed. Ed. by Martin Connors. Detroit: Gale, 1987. 1,271p. $160 (ISBN 0-8103-2141-6).

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).

Dollars and Sense: Implications of the New Online Technology for Managing the Library. Ed. by Bernard F. Pasqualini. Chicago and London: American Library Assn., 1987. 118p. paper, $10 (ISBN 0-8389-3338-6). "Proceedings of a Conference Program held in New York City, June 29, 1986, [by] Machine-Assisted Reference Service Section, Reference and Adult Services Division, American Library Association."

Goldfield, Randy J. *Office Information Technology: A Decision-Maker's Guide to Systems Planning and Implementation.* New York: Quorum, 1986. 226p. $35 (ISBN 0-89930-108-8).

Human Aspects of Library Automation: Helping Staff and Patrons Cope. Ed. by Debora Shaw. Clinic on Library Applications of Data Processing: 1985. Champaign, Ill.: Graduate School of Library and Information Science, Univ. of Illinois at Urbana–Champaign, 1986. 129p. $15 plus $1 postage (ISBN 0-87845-072-6).

Kies, Cosette. *Marketing and Public Relations for Libraries.* Scarecrow Library Administration Series, no.10. Metuchen. N.J., and London: Scarecrow, 1987. 202p. $14.50 (ISBN 0-8108-1925-2).

Kusack, James M. *Unions for Academic Library Support Staff: Impact on Workers and the Workplace.* New Directions in Information Management, no.10. New York, Westport, and London: Greenwood, 1986. 108p. $25 (ISBN 0-313-24991-1).

National Directory of Bulletin Board Systems, 1986/1987. Ed. by Ric Manning. Westport, Conn. and London: Meckler, 1987. 40p. paper, $19.95 (ISBN 0-88736-092-0).

Online Public Access to Library Files: Second National Conference. Ed. by Janet Kinsella. Oxford, Eng.: Elsevier, 1986. 141p. paper, (ISBN 0-946395-25-X).

Reid, William H. *Four Indications of Current North American Library and Information Doctoral Degree Programs.* University of Illinois Graduate School of Library and Information Science Occasional Papers, no.176. Champaign, Ill.: Graduate School of Library and Information Science, 1987. 83p. paper, $3 plus $.50 postage (ISSN 0276 1769).

Scham, A. M. *Managing Special Collections.* New York and London: Neal-Schuman, 1987. 201p. paper, $35 (ISBN 0-918212-98-7).

Smith, Milburn D., III. *Information and Records Management: A Decision-Maker's Guide to Systems Planning and Implementation.* New York: Quorum, 1986. 285p. $39.95 (ISBN 0-89930-111-8).

Yates-Mercer, Penelope A. *Private Viewdata in the UK.* Aldershot, England, and Brookfield, Vt.: Gower, 1985. 206p. paper, $55.95 (ISBN 0-566-02521-3). ■■

Letters

To the editor:
I regret that my article "Using CCF: The Common Communication Format" in the December 1986 issue of *ITAL* contains a factual error. It states that the standard ANSI Z39.2 requires that field tags consist of three characters and indicators two. This is not true of the standard, though it is true of the US/MARC formats, which follow the Z39.2 standard. The standard itself permits tags and indicators to be fixed by the designers of the individual format. I apologize for the error.—*Peter Simmons.* ■■

Compact discs are unequalled in their capacity for information storage. The extraordinary advantages of CD-ROM are now available from EBSCO Subscription Services through our new division, EBSCO Electronic Information.

A world of information in the palm of your hand.

Through EEI, EBSCO develops, markets and provides subscription services for a comprehensive range of CD-ROM databases. Because it's EBSCO, you can depend on easy ordering, updates and renewals.

EBSCO makes it easy.

EBSCO simplifies the purchase and use of CD-ROM technology. We offer the complete system including hardware, or the product compatible with your existing system. All your service needs can be met with a single toll-free phone call to EEI.

Call EBSCO today for the technology and the service that will take you into the library of the future.

EBSCO SUBSCRIPTION SERVICES

EBSCO Subscription Services
P.O. Box 1943
Birmingham, Alabama 35201
205/991-1182

ei

EBSCO ELECTRONIC INFORMATION