

Electronic Publishing in Library and Information Science

JOEL M. LEE
WILLIAM P. WHITELY
ARTHUR W. HAFNER

Definitions

IN THE 1960s, PUBLISHERS BEGAN to use computers to support the production of directories, indexes, and other print publications. Computer assistance saved substantial time and money, and, in conjunction with software for information retrieval systems, gave birth to electronic publishing. Directories, indexes, and other print publications that were produced electronically were then published, accessed, and used electronically and became known as databases.

Databases have had a profound impact on librarianship and have transformed both library user services and operations. Such client services as literature searching are now faster and more comprehensive. Library operations, like interlibrary loans and acquisitions, are now simpler and more effective. Machine-readable databases have also affected the dissemination of professional information. Librarians can find timely information about events and trends in librarianship in a number of databases.

The focus of this paper will be those databases which support library operations and provide professional information. The reader should be aware, however, that databases are just one facet of electronic

Joel M. Lee is Senior Manager, Information Technology Publishing, American Library Association, Chicago, Illinois; William P. Whitely is Director, Department of Marketing and Technical Services, Division of Library and Information Management, American Medical Association, Chicago, Illinois; and Arthur W. Hafner is Director, Division of Library and Information Management, American Medical Association, Chicago, Illinois.

publishing. Indeed, the term *electronic publishing* is used with little precision and may refer to a range of activities that include composing manuscripts, formatting pages, typesetting books, and producing databases. For the reader's information, the various uses of electronic publishing will be briefly described although these are generally outside the scope of this study.

Desktop Publishing

First, the term electronic publishing is used to refer to in-house production of small publications or "desktop publishing." Desktop publishers use microcomputers and peripheral devices—primarily laser printers—to compose and print pamphlets, books, manuals, and other publications in a workstation setting. Desktop publishing combines the output quality and font capabilities of laser printers with the ever-increasing power of microcomputer-based word processing and page makeup software. The method is quite effective and financially viable for many publishers who would otherwise have their publications commercially typeset, at considerable cost, or reproduced photographically from less attractive typed copy.

Publishing Production

Electronic publishing also refers to the production of large typeset-quality publications. A number of mini- and mainframe computer-based systems exist which enable users to input text and/or graphics (by typing, optically scanning, or transmitting from diskette or through telecommunications), to format pages automatically (hyphenate, justify), and to paginate automatically (create headers and footers, page numbers, tables of contents, and indexes).

There are over twenty-four commercial systems which perform these functions, and these can be divided into two basic categories—text systems, and text and graphics systems. The text systems perform formatting and pagination entirely in a batch mode, according to predefined page makeup and type font specifications, and with varying but still limited capabilities for integrating graphic material onto the page. By contrast, the text and graphics systems not only paginate content in batch mode, but also support interactive page makeup with specially designed high-resolution video display terminals. These systems are referred to as WYSIWYG systems—What You See Is What You Get—because the graphic artist is able to see on screen exactly what the print

Electronic Publishing

page will look like and can use the system to make alterations that will be automatically processed later.

Two types of computer architecture support these systems. One older computer architecture uses a network of dedicated terminals connected to a single processor—e.g., the Atex, Penta, and Compugraphics systems. The other, and more current, architecture relies largely on microcomputer workstations which emulate text-input terminals; these workstations may be on a network to the computer, or the data may be transferred into the main system from diskettes—e.g., Xyvision, Textet, and Miles 33. (This workstation approach leads to the occasional use of the term “desktop publishing” in this context.) Output may be to a laser printer, a laser typesetter, or magnetic tape device. In any case, a primary objective of current electronic publishing systems is to reduce typesetting costs by integrating an increasing number of functions formerly performed for the publisher by outside commercial typesetting shops.

Although the primary purpose of these systems has been the production of print materials, the data can be used to generate electronic products as well. As mentioned earlier, electronic database products were developed in the 1960s as byproducts of computerized typesetting.

In order to facilitate electronic production of both print texts and electronic databases, coding systems have been developed for the preparation and tagging of bibliographic records and text manuscripts that will be processed electronically. One recent coding system, the Standard Generalized Markup Language (SGML), is used by authors and editors to mark up texts before production; the standard generic codes can then be automatically translated into the appropriate typesetting and page formatting functions by the production system. One application of SGML is the American National Standard for Electronic Manuscript Preparation and Markup (Z39.59-198X, ISO 8879), a result of the Association of American Publishers' Electronic Manuscript Project.

The term electronic publishing is also used by authors who compose their manuscripts on microcomputers. These authors send both a diskette and a hard copy to their publishers. Ideally, this arrangement should reduce the time publishers must spend rekeystroking text, but because authors and publishers often use incompatible formats, software, and hardware, the problems of conversion can sometimes outweigh other benefits.

Software Publishing

Finally, the use of computers, especially microcomputers, to automate library operational and planning activities, has generated a

publishing medium new to librarianship—the publication on diskette of applications software and templates or programs designed to supplement and accompany printed works. This kind of publishing is better described as software publishing, and, while it is by no means an insignificant publishing activity, it is outside the primary scope of this paper.

TYPES OF ELECTRONIC PRODUCTS AND SERVICES

Although each of the aforementioned activities is occasionally referred to as electronic publishing, the term most commonly refers to machine-readable databases and this will be the focus of the remainder of this article.

Today there are over 3000 databases which offer a variety of information and services to users in hundreds of different subject areas. To make sense of this incredible diversity, most indexes and directories divide databases into logical groups. Every database can be classified as either a reference database or as a source database. These two types of databases are distinguished by the kind of information they contain. Reference databases contain references or citations and refer users to a primary source for more complete information. Source databases contain informative text, raw data, or computer programs and are primary sources of information.¹

The two database kingdoms, reference and source, are often subdivided into database phyla. Reference databases are subdivided into bibliographic and referral databases; source databases are subdivided into textual, numeric, and software databases.² Later, we will consider how each database can be used by library professionals to obtain information and to support library operations.

Bibliographic Databases

Bibliographic databases are reference databases which, as the term implies, can be used to generate a bibliography on a specific topic. Bibliographic databases contain citations to articles, books, reports, and other primary sources of information. In many cases, these databases also contain abstracts for selected items. Some bibliographic databases focus on the professional interests of librarians and information professionals. Examples include INFODATA, INFORMATION SCIENCE ABSTRACTS, LIBRARY LITERATURE, and LISA. These files cite

Electronic Publishing

journal articles, abstracts, and book reviews on library and information science topics. As is often the case with databases, some of these bibliographic products correspond to print publications. LISA, for instance, is an online version of the publication *Library and Information Science Abstracts* as LIBRARY LITERATURE corresponds to *Library Literature*, and INFORMATION SCIENCE ABSTRACTS corresponds to *Information Science Abstracts*.

Bibliographic databases can also be used to support technical activities such as cataloging, acquisitions, and interlibrary loans. OCLC, for example, contains information on the holdings of the OCLC member libraries. The database enables users to borrow materials from other libraries, to generate catalog cards for books and serials, or to produce onsite online catalogs. There are over fifty library holdings databases; other examples include CATLINE, DOCLINE, and RLIN, which collectively contain citations to the holdings of the members of the Research Libraries Group, the National Library of Medicine, the Library of Congress, and many other government, health science, and research libraries.

Referral Databases

Bibliographic databases are one subclass of reference databases; referral databases are another subclass. Referral databases may cite nonpublished or nonprint sources of information such as organizations, individuals, or audiovisual material. ONLINE CAREERS is one example of a referral database that focuses on professional interests and refers its users to employment opportunities in the online field.

A far greater number of referral databases, however, support technical and information-oriented activities. ACCESS, BRS/FILE, CUADRA DIRECTORY OF DATABASES, THE DATABASE OF DATABASES, and PUBLICACCESS MESSAGE SYSTEMS refer users to telecommunication networks, other databases, or public access electronic bulletin boards. Many of these referral databases are also produced as hard copy publications.

Text Databases

Text databases are a subclass of source databases. They contain the complete text of a primary source. These databases are attractive because they can provide quick and direct access to specific passages or articles

that are buried inside voluminous primary sources. Nonlibrary examples would include the full-text newspapers in VU/TEXT or DATA-TIMES. In library and information services, text databases such as ALANET's ALA NEWS BULLETIN, ALA WASHINGTON NEWS-LINE, and INTELLECTUAL FREEDOM ALERT, or Online Inc.'s ONLINE CHRONICLE, target the business concerns of libraries and the professional interests of librarians and online professionals. Other full-text databases—such as BRS BULLETIN and CHRONOLOG NEWSLETTER—lean toward library support by describing databases, discussing search strategies, and announcing new services.

Currently, more and more publishers are creating online versions of their traditional print publications. BRS BULLETIN and CHRONOLOG NEWSLETTER are two full-text databases that correspond to printed publications. Although the ALA WASHINGTON NEWS-LINE and INTELLECTUAL FREEDOM ALERT have relationships with print publications, ALA goes one step further by publishing some information exclusively in electronic formats. Still other publications, such as INFORMATION PUBLISHING: AN ELECTRONIC JOURNAL, only appear in an electronic format with no print equivalent.

The trend toward publishing more information electronically and some information exclusively in that form has led some analysts to predict that print publications will soon become obsolete and disappear.³ Such a future seems unlikely for a format as simple and portable as the book, but certain areas of publishing may well shift toward all-electronic products as pressures for timeliness increase, as technological tools evolve and become more widely available, and as the economics of publishing and distribution change.

Numeric Databases

Numeric databases are another subclass of source databases. They contain statistics or survey data that users often can use interactively. Several organizations collect library statistics including ALA, the Association of Research Libraries, R.R. Bowker Co., and the Center for Education Statistics (CES, formerly the National Center for Education Statistics, NCES) in the U.S. Department of Education. However, none of the surveys is fully available online. Industrious and knowledgeable users can obtain magnetic tapes which contain the raw text of CES publications, but these users must have the hardware needed to store the tapes as well as searching and statistical software. Practically, numeric databases are not available to the average user of library statistics.⁴

USAGE OF LIBRARY AND
INFORMATION SCIENCE DATABASES

Information professionals can use databases either to obtain professional information or to support technical and user services, but technical support seems to account for the greatest volume of database usage in libraries. Two sets of data support this assertion. First, the number of technical support databases far exceeds the number of professional information databases. In the area of technical and user support, Cuadra lists thirty information service directories and more than sixty library holdings databases; in the area of professional information, however, Cuadra lists only sixteen databases. Second, usage data show that technical support databases are used more than professional information databases. In a survey of online professionals conducted by Marquis Who's Who, Inc., users of online services ranked the databases they used most frequently. The top forty databases listed included technical support databases such as OCLC, Books In Print, and CATLINE, but none of the guides to library literature appeared in the list.⁵

Database vendors, such as Dialog, are well aware of the importance of technical support. Indeed, at the 1987 ALA Midwinter Meeting, Dialog announced one new marketing plan that will focus on databases which support a variety of technical services functions including acquisitions and cataloging. This announcement presages the development of more promotional campaigns aimed at technical services staff and others who have not been the primary users of online search services.

Another intriguing feature of library and information science databases is that reference databases outnumber source databases five to one. In other fields such as medicine and business, source databases have been vigorously developed. For instance, Cuadra lists sixty-six biomedicine databases; twenty contain source data including up-to-date information from the Centers for Disease Control (Morbidity and Mortality Weekly Report) and the full text of clinical journals (MEDIS). Martha E. Williams's *Computer Readable Databases* lists more than 200 medical databases of which nearly 40 percent are source databases. In the brokerage industry, all ninety databases on commodities and futures contain original source information. Yet in library and information science, only three of sixteen databases are identified as source databases.

Obstacles to Production and Usage

There are several reasons why producers have not published more source databases in library and information science. First, source databases are expensive to use online. Blodwen Tarter of Information Access Company (IAC), observes that users of source databases often incur expensive connect time charges. Indeed, users are reluctant to browse when the clock is running, and many of the materials available in full-text databases are also available in hard copy on the shelves.⁶ As Tarter reports, full-text databases at IAC have not earned substantial revenues. IAC's experience with general materials in varied subjects can be extrapolated to librarianship and information science.

Second, source databases presuppose an urgent need for information. In some fields, online access is the most effective means of rapid access to urgently needed text. In many fields—such as news, law, and business—data are dynamic and voluminous. In these fields, full-text databases have achieved greater success. Businesses, especially in the for-profit sector, are much less price sensitive than the not-for-profit libraries. Indeed, lawyers, stock traders, doctors, and business people are more willing to pay the price for full-text retrieval especially since information is often tied to profitability.

Third, source databases are complex and can be expensive to produce. Numeric and textual-numeric databases, two other types of source databases, are used widely among stock traders, physicists, and chemists but are absent in library and information science. One factor contributing to the lack of numeric databases may be the way in which library statistics are collected.⁷ In *Sources of Library Statistics*, Lynch notes that the terms used in library surveys are often unclear and that the collection methods lack uniformity. As a result, statistics published by a government agency like the Center for Education Statistics may not be consistent from year to year or from institution to institution to merit electronic compilation or analysis. CES is further impeded by understaffing and frequent reorganization. The R.R. Bowker Co., publishers of the *American Library Directory* (ALD) which contains a significant amount of statistical data reported by libraries, has produced neither a statistical print publication nor a numeric database from this work. The ALD was recently installed online in Dialog, and it will be interesting to see what kinds of data analysis the online version can support.

It would appear that Bowker, a for-profit corporation, is not convinced that the demand for statistics by librarians is sufficient to warrant a substantial investment in developing electronic or print statistical products. The reality of database publishing, much like other forms of

Electronic Publishing

publishing, is that the database must eventually be profitable or, in the case of government- or other grant-subsidized projects, be able to become self-supporting after an initial investment. An industry rule-of-thumb states that it takes from three to five years before a database product shows a profit and even then only after significant investments in design, development, execution, promotion, and customer support. The potential market for library and information science databases, especially for numeric ones, may be perceived by producers as not large enough to warrant the investment.

Some of the major publishers in library and information science have either been slow to begin publishing electronic information or have not entered the online marketplace at all. ALA, for example, only began its ALANET service in January 1984 and established an Information Technology Publishing section in ALA Publishing Services in September 1986. Also in 1986, the ALA initiated a joint venture with Research Publications to produce *The Directory of Library and Information Professionals* scheduled for publication in Spring 1988. It is a comprehensive biographical directory of the information field and is construed more broadly than previous biographical directories in the profession. Plans for the project include not only a three-volume print work, but also a CD-ROM (compact disc read-only memory) to be marketed by ALA Publishing Services. The results of the project will be informative, providing insights into the marketability of a source database on CD-ROM specifically in the library and information science field.

Use of Electronic Publications

Certain libraries appear to make more use of online databases. Chatterton and Pemberton found that about 38 percent of online professionals work in college and university libraries, 26 percent in corporate libraries, 9 percent in government libraries, and only 4 percent in public libraries.⁸ These results suggest that there is a shortage of online services in public libraries. A 1981 study by Mary Jo Lynch confirmed that university and college libraries are more likely to have online services than public libraries. The study also found that public libraries are more likely to have online services than two-year college libraries and school libraries/media centers.⁹ A 1987 survey by the ALA Office for Research found that 35 percent of public libraries serving populations of over 25,000 offer database searching.¹⁰ A 1984 survey of online services in academic libraries reported that over 80 percent of university and over

40 percent of college libraries offered this service, while only about 20 percent of junior college libraries did. However, respondents' projections showed that by 1987 almost 100 percent of university, 80 percent of college, and 70 percent of junior college libraries would offer search services; some libraries in each group reported that such services were available elsewhere in the community.¹¹

In order to establish online information services, a library must invest substantial time and money. Often only large, well-financed libraries are able to afford such an investment. The costs of training time, equipment, and subscription fees—even where actual usage is primarily volume-based—may be prohibitive for smaller libraries. The majority of public libraries may be especially hard pressed to find the necessary resources because of size—80 percent of U.S. public libraries serve populations of 25,000 or fewer and 63 percent serve 10,000 or less.¹²

Human factors also play a role in computer use. Many librarians still resist acquiring and learning to use electronic equipment because the available hard copy versions of indexes, texts, and data forms are familiar and seemingly effective. This resistance is evident even in libraries where operational processes have been automated, often with grant funding, but where research and reference questions are still answered using hard copy materials. Potential users of online information may themselves be intimidated by the enormous number of available options and services.

Moreover, librarians under financial pressure seem to be less heedful to the lure of new information formats and less sensitive to their own need for professional literature. When pressed to make stringent budgetary decisions on resource allocations, librarians will more likely rely on the traditional formats and methods of information service, and will more likely attend to the needs of their clients rather than reserve any significant part of their budgets for professional literature and online information.

PRODUCTION AND DISTRIBUTION OF DATABASES

Electronic databases form a growing industry. From 1976 to 1986 the number of databases recorded in Williams's *Computer Readable Databases* grew from 301 to 2805.¹³ Yet during this decade of enormous growth, options for the distribution of databases remained fairly stable.

Electronic Publishing

Database Vendors

Dialog, BRS, Mead Data Central, and other large commercial vendors dominate the online industry. In 1983, Dialog and Mead alone accounted for 71 percent of total industry-wide online usage and 83 percent of total revenues.¹⁴ Vendors provide a valuable service to database producers because they distribute and, to some degree, market databases, they provide hardware support and search software, and they permit access through telecommunications. Producers prepare the data files, transmit them to the vendors, and receive royalties in return. Library Association Publishing, for instance, produces LISA, but it is made available through Dialog and Orbit.

Producer/Vendors

Other firms both produce and distribute their databases. The H. W. Wilson Company produces LIBRARY LITERATURE, NAME AUTHORITY FILE, JOURNAL AUTHORITY FILE, and PUBLISHER AUTHORITY FILE. Wilson is also the vendor of these products through its own WILSONLINE online service. Other producers may post their own data on their own hardware but allocate other details of access and telecommunications. An example is the OFFICIAL AIRLINES GUIDES ELECTRONIC EDITION, which is available through user transparent gateway on many services.

Value-Added Services and Other Providers

In many cases, database providers offer their users additional value-added services. As discussed earlier, CATLINE, OCLC, and RLIN are online services that provide not only bibliographic references, but also such technical and service support facilities as cataloging and interlibrary loan. The DOCLINE holdings database supports a network for document delivery to health science libraries. Both EBSCO Subscription Services and the Faxon Company, originally founded as library periodical subscription agents, are producer/vendors of bibliographic databases which contain citations to journals published throughout the world. These companies' databases also support full-scale serials acquisitions and control systems.

Other participants in the library electronic publishing industry include not-for-profit associations and other organizations such as the American Library Association (producer of ALANET), government agencies, and national libraries like the National Library of Medicine

(which produces CATLINE). However, the distinctions between for-profit and nonprofit status of the database producer or vendor is a factor of only marginal relevance, for the "rules of the game" in the marketplace generally apply to all participants in the industry.

TRENDS, CHANGES, AND PROJECTIONS

A number of trends in the electronic information industry will have significant impact on libraries and information centers. Librarians will face these trends both as they use information resources for their clientele and as they use information resources to support library operations and obtain professional information.

CD-ROM: Impact and Issues

Major changes in formats and distribution patterns are developing as the entire industry prepares for a shift from online access to optical disc technology. This development is being fueled by two synergistic and complementary trends: (1) a rapidly increasing number of CD-ROM products, and (2) a growing base of installed microcomputer equipment that can be used for both online and CD-ROM access to databases. As the installed base of CD-ROM equipment grows, online vendors and database producers will offer more databases in the CD-ROM format which will in turn help libraries justify the cost of the equipment. "Jukebox" players that can handle several CD-ROMs and networks of CD-ROM workstations will also enhance the marketability of CD-ROM publications.

Some producers are encouraging this growth by marketing computer hardware packaged with their optical disc products. Bowker Electronic Publishing provides several alternatives for acquiring CD-ROM players for their BOOKS IN PRINT PLUS CD-ROM product, and H.W. Wilson Co., as an IBM distributor, has developed its own WILSONDISC workstation. Interestingly, the WILSONDISC software supports both online access to the WILSONLINE files and access to the CD-ROM data.

The growth of the CD-ROM market may have a favorable effect on neglected information services. Because many pages of text can be stored on CD-ROM and because users of CD-ROM pay no separate search fees or connect-time charges, database searching, especially in full-text and source files, may become more attractive in a CD-ROM environment.¹⁵

Electronic Publishing

On the other hand, much depends on the pricing strategies adopted by the producers in marketing their CD-ROMs. The "pay-as-you-go" fee structure of online access may be attractive and helpful to marginal users of certain files and to smaller libraries that cannot justify or support substantial subscriptions to major bibliographic services. However, the "all-you-can-eat" pricing of a CD-ROM subscription, which is comparable to a traditional subscription to a print information service, may be unattractive for the library with more limited funding and a smaller information need.

The Wilson Company's approach to pricing the WILSONDISC product merits attention. Each CD-ROM database is a separate subscription that includes periodic cumulative updates. Also included in the CD-ROM purchase is *free* access to the most current data online through WILSONLINE (telecommunications charges are not covered). This strategy seeks to build vertically upon an existing online user base and to accommodate the inevitable change in balance between online access to optical disc access.

Database producers have some good reasons to adopt CD-ROM as a distribution method for their products. With CD-ROM publications, producers can achieve a more visible role and greater control in marketing and customer support. Moreover, eliminating the online vendor and other middlemen can favorably affect pricing. In fact, the importance of vendors as distributors and marketers could be greatly reduced if more database products are delivered directly to the users on compact discs.¹⁶ In this scenario, online services would contain only those databases not yet published on CD-ROM and those data files which appear between CD-ROM issues.

It is too early to predict what will happen when the CD-ROM industry is able to accommodate nontextual data more effectively, including digitized graphics (in contrast to the photographic image on videodisc), the CD-I (compact disc interactive) format by Philips and Sony for which products should emerge by 1989, and DVI (digital video interactive) announced in 1987 by GE/RCA Laboratories. The firms' strategy is to apply these formats first to broad consumer markets, so that librarians can expect to be collecting CD-I and DVI for their patrons as they now do compact audio discs, before librarians will see materials produced for their professional information needs.

Product Development Trends

Currently, the players in the online field are (1) diversifying lines of business to include both online and CD-ROM products; (2) merging

several databases in one place, especially on CD-ROM; (3) developing offline "front end" software and postprocessing software to formulate and conduct search requests and to format the results; and (4) developing so-called gateway and expert systems, including multiple-file searching techniques. These four developments, though diverse, reflect the industry's fascination with a single concept: integration.

New Product Lines

The diversification of online businesses to include both online and CD-ROM products has already been discussed. A variety of firms are seeking to establish positions and roles in the CD-ROM marketplace. Individual database producers are negotiating independently with CD-ROM production firms for the compact disc versions of their products. For example, the Public Affairs Information Service, which produces the PAIS database available online through Dialog and BRS, has executed an agreement with Online Computer Systems, Inc. to produce a CD-ROM version. Producer/vendors, notably H.W. Wilson, have also launched products with ties to their online (and print) services. Database vendors including both Dialog and BRS are developing plans for participating in the CD-ROM action, either as primary distributors of compact disc products (Dialog OnDisc) or as service bureaus for CD-ROM production.

Producers have merged several products into one online service and/or CD-ROM product as a strategy for building stronger products and services. Online examples include Magazine ASAP and Mead's NEXIS; CD-ROM products include IAC's successful Infotrac. Such merging adds strength to individual database products that might not stand on their own or stand as well. Further, a producer may use merging as a strategy to extend the market for one product by facilitating its use in concert with other products. In the library field, a product released in 1987 is a CD-ROM database of media reviews as an enhanced version of Bowker's BOOKS IN PRINT PLUS. This database pools such Bowker reviewing sources as *Library Journal* and *School Library Journal* with other publishers' materials, including the ALA's *Booklist* and *Choice*.

Preprocessing and Postprocessing Software

Database vendors and independent software producers have been developing microcomputer "front end" (preprocessing) and postprocessing software. These systems and services respond to user demands for greater ease and reduced connect time costs for constructing a search

Electronic Publishing

and to user interest in processing search results into formats suitable for presentation. ProSEARCH, for example, is a software package which was developed for professional searchers by the Menlo Corporation and is now owned by Personal Bibliographic Software Corporation. ProSEARCH can access both Dialog and BRS and allows the searcher to explore all available databases. The software acts as an emulator, translating the original search request into the unique languages of the various databases.¹⁷ Other microcomputer-based front ends include Dialog Link and Wilson's WILSEARCH.

Postprocessing software adds value to the online information by manipulating the search results. Another example from Personal Bibliographic Software is Victor Rosenberg's ProCITE, which describes itself as a "scholar's workstation." ProCITE sorts, manipulates, merges, and formats downloaded search files (including those from ProSEARCH) into customized, user-defined bibliographies. Another, and very powerful, example of a postprocessing, value-added application is Datext, a CD-ROM financial service that merges searching of multiple statistical and other numeric and text databases, manipulation and calculation of statistics, and microcomputer desktop publishing for report output.

Expert Systems and Gateways

A major problem with computer-based information retrieval is that a computer, though fast and tireless, is myopic. Only a fraction of the database, and only a fraction of the search results, can be displayed at any one time. So called "expert systems" attempt to model the human mind and eye, which can survey contents pages, shelves full of books, or files of documents for relevant materials. Current expert systems are an early outgrowth of the still-developing field of artificial intelligence.

The expert systems that are currently available in the online industry are not truly expert but actually a combination of gateway systems and online front-end systems. These systems seek to respond to three increasingly compelling needs: (1) the need for a common command language to simplify searching; (2) the need for centralized access to multiple vendors and databases, or "one-stop shopping"; and (3) the desire to reduce the record-keeping, paperwork, and contracts associated with maintaining separate subscriptions to different vendors' services.

A gateway is simply that—a portal from one computer to another. A gateway system serves as an intermediary between the user and an array of outside hosts. These gateway front ends, designed to be adapted

to the users' level of expertise, are much like microcomputer front-end software in that they assist with database selection and search query formulation. The systems then "open" the appropriate gateway, dialing into the selected host and potentially carrying out the search on behalf of the user. As a result, the user receives assistance with search formulation, avoids the complexities of individual vendor command languages and database file structures, and also gains administratively by having only a single invoice to process. Two primary gateway systems that have appeared in the market are EasyNet, self-described as "an expert system for non-experts," and OCLC LINK, which was released in November 1986 but withdrawn in April 1988.

EasyNet

EasyNet is designed especially for inexperienced end users. It was developed by Telebase Systems with the sponsorship of the National Federation of Abstracting and Information Systems (NFAIS). Telebase claims that EasyNet is "friendly" for all skill levels, but it is primarily designed and most appropriate for the naïve user. EasyNet's software operates on microcomputers and is accessed by dialing a direct toll-free number or through a gateway from another service. EasyNet searches any of over 900 databases available through more than twenty vendors and handles all the billing for any of the vendors accessed (collecting charges onto one bill). Each search retrieves up to ten of the most recent citations for a flat fee (to which are added telecommunications charges). Additional citations are available at an extra charge (as an additional search), and a search with no results incurs no search fee.

For an individual customer, EasyNet eliminates numerous technological and psychological barriers to online searching,¹⁸ not to mention the administrative subscription-related ones. It has been extremely popular in the consumer market under the names I-QUEST and EINSTEIN on CompuServe and in the corporate and online markets under the name InfoMaster which is marketed by Western Union. Despite misgivings on the part of some professional searchers, librarians have been drawn to EasyNet because of the simple subscription arrangements and ease of searching, especially for library patrons as end users.

One appearance of EasyNet in the library community is germane to this discussion. ALANET PLUS is a gateway from the ALANET computer to EasyNet, which presents a set of menus specially designed for the information professional. In addition to providing access to all databases on EasyNet, ALANET PLUS identifies more than forty files in library and information science and related fields (including Wilson's LIBRARY LITERATURE) and prompts the user to assist in the

Electronic Publishing

selection of databases appropriate to the librarian's need for professional information and literature. (It should be added that in addition to its customized version of EasyNet, ALANET has also set up its own gateways to VU/TEXT and EBSCONET.

OCLC LINK

OCLC began major promotions of their gateway, OCLC LINK, in January 1987 but withdrew it from the market in April 1988 due to low usage in relation to required computer resources. Access to Dialog, BRS, VU/TEXT, EBSCONET, and other databases had already been announced and OCLC continued negotiating with other information providers. While the gateway front-end principle might be similar to that of EasyNet, the specific features and functions of OCLC LINK are quite different.

OCLC LINK described itself as "an intelligent gateway" that helps searchers identify online databases, store and edit search results, and communicate electronically with other online users through electronic bulletin boards, online forms, messaging, and conferencing facilities. In promotional materials, OCLC emphasized "preconnection, connection, and post-connection services." OCLC LINK was derived from the iNET 2000 system, for which OCLC acquired a license from Telecom Canada. OCLC's arrangements with information providers and its own telecommunications capabilities provided for advantageous rates for connect time and telecommunications. Unlike EasyNet, OCLC LINK did not actually construct and carry out the search for the user, nor did it eliminate the need for separate contracts and passwords from the various providers. However, it did provide substantial up-front assistance in database selection and online tutorial information on each vendor's command language; using this information and an online editor, the user could prepare the search query before making the connection to the vendor's system. Advertisements for iNET 2000 began to appear in the United States in early 1988, perhaps due to the withdrawal of OCLC LINK.

EasyNet's recent development of techniques that allow for multiple-file searching—Med-Scan on InfoMaster—is in response to another desire of searchers. Conducting the same search on a variety of files has been a tedious and complicated process, but often a much needed one; automatic searching of multiple files facilitates this process. The library and information science files highlighted in ALANET PLUS were scheduled to have a comparable scanning facility in Spring 1988, and additional multiple-file searching processes are under development by Telebase Systems.

While both EasyNet, OCLC LINK, and iNET 2000 are value-added services with much to offer information users, they do not qualify as expert systems. Rather, they are early steps in the lengthy process of developing truly expert systems for access to online information resources.

THE DEMAND FOR INTEGRATION

The several paths described earlier reflect a larger movement, proceeding irregularly, toward a greater degree of integration of access to various information products and services. Whereas the first fifteen to twenty years of the online industry witnessed tremendous proliferation of databases, the negative side effect of this growth was a diminution in the search intermediary's ability to record the burgeoning number of potentially useful databases, much less the complexities of differing access routes to those databases, particularly under the constant pressure of accumulating connect time charges.

At the same time, the use of microcomputers has had significant impact on the online industry as modem-equipped microcomputers quickly began to replace "dumb" terminals for online searching. The 1985 *Directory of Microcomputer Users in Libraries*, prepared by the Technology in Public Libraries Committee of the Public Library Association, a division of ALA, gathered responses from 381 libraries in North America (50.9 percent of them public). Among responding libraries of all types, 47 percent use microcomputers for online searching. Microcomputers were used most for word processing (91.6 percent), data management (83.7 percent), and spreadsheets (72.2 percent), with a total of 91.9 percent of reported usage attributed to library administrative activity.²⁰ While the directory's sample universe may not be comprehensive, the data are still revealing.

In the consumer market, microcomputer-based bulletin boards, with their rather primitive but still attractive facilities for electronic mail and conferencing, became a major attraction to many information professionals. Electronic mail, which had been developing separately but at approximately the same time as the online industry, grew in value as a corporate information tool through such vendors as Dialcom, Inc.; Telemail; GeNie; and others.

Industry Responses

Among the offline responses to this challenge are the front-end and postprocessing software packages that short-circuit the problems of

Electronic Publishing

search query complexity and direct connect time costs, and which add additional value by enhancing the searcher's ability to produce presentable output. Another offline response has been the CD-ROM versions of online databases, followed almost immediately by the merging of separate datafiles to create integrated CD-ROM products.

The online response has been a growing trend in the professional markets—already evident in consumer information markets since the early 1980s—toward integration of multiple online resources and functions. One model, embodied in such consumer “information utilities” as *The Source* and CompuServe, has grown in the professional market; increasingly the online information providers combine a variety of facilities—i.e., electronic mail, database searching, bulletin boards, conferencing, and other functions—as part of a single information service. Among the attractions of such services are the convenience of one-stop shopping noted earlier, the multiuser mainframe environment compared with the single-user microcomputer-based bulletin board, and the greater power and responsiveness of such systems to support a variety of communication and information applications.

Dialcom, Inc. has marketed its integrated service to commercial and government users since 1970 (and *The Source* was founded by former Dialcom staff); ALA's ALANET is a value-added reseller of Dialcom products. Dialog introduced Dialmail in 1985, adding electronic mail, conferences, and bulletin boards to its product line. OCLC LINK and iNET 2000 build upon the gateway concept to include in its services electronic messaging, bulletin boards, and other functions. Western Union was able to expand the limited offering of its EasyLink electronic mail service by adding EasyNet under the name InfoMaster. Both *The Source* and CompuServe have recently made some marketing approaches to associations and corporate organizations, emphasizing the multifaceted nature of the services they offer; this strategy recognizes that the consumer market for information utilities has reached the saturation point, particularly with the added competition of free access microcomputer-based bulletin board systems. These utilities also seek to maximize the use of their computer resources during prime daytime business hours when recreational users are least likely to use them; CompuServe uses its private telecommunications facilities in part to support the telecommunications needs of OCLC, the PaperChase gateway to the National Library of Medicine, and other information providers.

As noted earlier, the preprocessing and gateway services currently available represent early marketplace manifestations of the long-term

process of developing expert systems. While these facilities provide many concrete advantages to their users and substantial financial gain to their purveyors, they are still far from achieving the broader objectives of a common command language and other mechanisms for making the techniques of database query transparent to the user. The complexities of the free enterprise system only add to the difficulty in achieving the level of technological harmony the concept of expert systems implies.

Thus the path toward integration is proving to be a complex one for the information provider, who must balance the pressure for greater accessibility through a multiplicity of routes with the need to maintain identity and strength of market share. This is even more true for users, who must both select from a growing array of online services and offline products and adapt their established operating procedures to these changes. Gateway systems can only mitigate the latter challenge; the working information environment for librarians and information specialists will still remain a complex and dynamic one.

THE FUTURE FOR ELECTRONIC PUBLISHING IN LIBRARY AND INFORMATION SCIENCE

The foregoing analyses of industry and market trends apply generally to the electronic information industry, of which, as we have seen, library and information science is one part. The data on the number and usage of databases in the field suggest that it is only a small part of the whole, and one which is accorded far less priority from its primary audience—librarians and other information professionals—than are databases serving their clients' needs. There is no significant indication that a shift in priorities on the part of information professionals is forthcoming despite the fact that the professional literature is a primary resource for any field that seeks to respond to both a dynamic external environment and internal pressures to adapt to increasingly complex issues and practices.

Thus electronic publishing in library and information science seems destined to play a continual game of catch-up, with information providers in the field applying new techniques and technologies to the professional literature well after they have been applied to the information resources of other fields. The extent to which professionally oriented information products and services succeed in the library and information services marketplace will be one measure of the importance

Electronic Publishing

information professionals give themselves at a time of significant technological, environmental, and professional change.

References

1. Wanger, Judith, and Landau, Ruth N. "Nonbibliographic On-line Data Base Services." *JASIS* 31(May 1980):171-80; and Maloney, James. "Online Information Services." In *ALA World Encyclopedia*, 2d ed., edited by Robert Wedgeworth, pp. 611-21. Chicago: ALA, 1986.
2. Cuadra, Carlos. *Directory of Online Databases* 7(Jan. 1986).
3. Lancaster, F. Wilfrid. "The Paperless Society Revisited." *American Libraries* 16(Sept. 1985):553-55.
4. Lynch, Mary Jo. *Sources of Library Statistics 1972-1982*. Chicago: ALA, 1983.
5. Chatterton, Fred, and Pemberton, Jeff. "The Online Professionals—Who Does What... The Marquis/Online, Inc. Project." *Online* 9(July 1985):15-24.
6. Tarter, Blowden. "The Future of Online Full Text Databases." In *SCIL 1986: Software/Computer Conference Proceedings*, chaired by Nancy Jean Melin, pp. 180-84. Westport, Conn.: Meckler Publishing, 1986.
7. Lynch, *Sources of Library Statistics*.
8. Chatterton, and Pemberton, "The Online Professionals."
9. Lynch, Mary Jo. *Financing Online Search Services in Publicly Supported Libraries*. Chicago: ALA, 1981.
10. American Library Association. "ALA Survey of Public Library Services: A Summary Report." Chicago: ALA, Office for Research, 1987, typescript, p. 2.
11. McKinney, Gayle, and Mosby, Ann Page. "Online in Academe: A Survey of Online Searching in U.S. Colleges and Universities." *Online Review* 10(April 1986):115.
12. See Lynch, Mary Jo. *Libraries in an Information Society: A Statistical Summary*. Chicago: ALA, 1987.
13. Williams, Martha E. *Computer Readable Databases*. Chicago: ALA, 1985.
14. _____. "Usage and Revenue Data for the Online Database Industry." *Online Review* 9(June 1985):205-10.
15. Hensel, Martin. "Data Publishing on Optical Disks." In *CD-ROM: The New Papyrus*, edited by Steve Lambert and Suzanne Ropiequet, pp. 487-93. Redmond, Wash.: MicroSoft Press, 1986.
16. Williams, Martha E. "Highlights of the Online Database Field: CD-ROM and New Technologies vs. Online." In *Proceedings of the Seventh National Online Meeting*, compiled by Martha E. Williams and Thomas H. Hogan, pp. 1-4. Medford, N.J.: Learned Information, Inc. 1986.
17. Hawkins, Donald T., and Levy, Louise R. "Front End Software for Online Databases Searching, Part 1: Definitions, System Features, and Evaluation." *Online* 9(Nov. 1985):30-36.
18. _____. "Front End Software for Online Database Searching, Part 2: The Marketplace." *Online* 10(Jan. 1986):33-40.
19. _____. "Front End Software for Online Database Searching, Part 3: Product Selection Chart and Bibliography." *Online* 10(May 1986):49-58.
20. Burgin, Robert. "The User of Microcomputers Across Types of Libraries." *Library Hi Tech News* 34(Jan. 1987):1, 6-9.

This Page Intentionally Left Blank