
Online Searching Technique and Management

Edited by James J. Maloney

Online Searching Technique and Management

Edited by James J. Maloney

AMERICAN LIBRARY ASSOCIATION • *Chicago 1983*

Preface

These papers are based on the proceedings of a day-long program of the Machine-Assisted Reference Section (MARS) of RASD, held on July 11, 1982, at the annual conference of the American Library Association in Philadelphia. In two separate sessions the program addressed the needs of the novice searcher and of the manager of an online search service. The morning portion of the program was devoted to an introduction to online searching for the novice searcher. Papers discussing the management of an online search service were delivered in the afternoon.

The program resulted from a decision of the MARS Executive Committee to respond to librarians who attended MARS committee meetings seeking advice or guidance on a variety of recurring questions, such as:

- What is the relation between database searching and the more traditional library services?
- When does one decide to conduct an online search?
- How does one choose a database vendor?
- What is database searching used for?
- How does one resolve the unique financial considerations prompted by database searching?
- How can an online search service be funded?
- What are the characteristics of a good search analyst?
- What preparation and training are necessary for the professional librarian to become a search analyst?

The question What is database searching used for? is asked not only by novice searchers and librarians initiating online database access, but also by librarians seeking innovative users for an established online search service.

Questions of this nature are the focus of considerable attention and effort on the part of committees within MARS. These committees regularly gather information in order to study issues or problems as they af-

fect database searching, to make recommendations or establish guidelines, or to serve as advocates on behalf of the community of database searchers. The framework for the 1982 MARS program was provided by one of those committees, the Education and Training of Search Analysts (or MARS/ETSA) Committee. Under chairs Greg Byerly (1981) and Rebecca Whitaker (1982), the MARS/ETSA Committee developed "An Introduction to Online Searching: A Suggested Outline" (included here as the appendix). This outline served as a checklist for librarians by profiling the basic concepts of database searching and elaborating the applications of this service to library work. The work of the MARS/ETSA Committee served as an ideal starting point for the members of the 1982 MARS Program Committee to select an outline of topics for the program, since the program would address the questions that were most frequently asked regarding the basic concepts and applications of database searching.

With three exceptions, all papers for the program were chosen by the Program Committee during the 1982 ALA Midwinter Meeting from nearly one hundred abstracts that had been submitted in response to a general "call for papers" in the fall of 1981. Because of the wealth of their experience, wit, and insight into the field of database searching, Sara Knapp and Peter Watson were approached by the committee to deliver keynote addresses for the morning and afternoon sessions, respectively. Carol Fenichel volunteered on the basis of her extensive teaching and writing experience in the field of database searching to fill any gap that the program committee might have. The committee responded by having her address the difficult topics of databases and database producers and of database vendors and search services.

This work addresses the basic concepts and applications of database searching, the various institutional considerations such as policies, procedures, record keeping, and promotion as they relate to a search service, financing, and the training of database searchers.

As a text on database searching, the work is a unique publication insofar as it is drawn from the experience, background, and perspective of a variety of people. Sally Knapp's view on how the public and private sectors should determine future development of online searching reflects the convictions and principles of a librarian who has actively engaged and shaped this issue. Randolph Hock's chapter, "Who Should Search?" is a thorough and emphatic analysis of the attributes of a good searcher, from someone who has personally trained over 2000 searchers. To quote Hock, "I feel strongly that the future of reference librarianship is going to be determined to a large degree by the use of online systems, that the quality of reference service will largely depend upon the quality of

searching, and that the quality of searching will be dependent upon who is searching." Carol Fenichel's paper exudes with confidence that "soon it will be feasible to do all reference work online." The selection of equipment for online searching is expressed by Janet Bruman with the clarity and detail of someone with considerable professional experience.

This work also represents the contributions of many librarians who have actively participated in the work of MARS committees. In addition to the contribution of the MARS Education and Training of Search Analysts Committee previously referred to, much of the work of MARS Cost and Financing Committee is embodied in the paper presented by Nancy Grimes, the chair of that committee, entitled "Online Reference Services: Costs, Budgets, and Financial Management." The essence of a work based on the 1982 MARS program lent itself well to the inclusion of the two separate, selected bibliographies of materials recently compiled by the MARS Committee on Measurement and Evaluation of Service and the MARS Use of Machine-Assisted Reference Services in Public Libraries Committee.

Credit for the program is given to the members of the 1982 MARS Program Committee whose energy and enthusiasm for the project made the program possible. The committee members included:

Nathan Baum, Reference Department, Melville Library, State University of New York Stony Brook, New York

Linda Friend, Patee Library, Pennsylvania State University, University Park, Pennsylvania

Rachel Gade, Ramsey County Public Library, Roseville, Minnesota

Bernard F. Pasqualini, Free Library of Philadelphia, Philadelphia, Pennsylvania

Maria J. Soule, Learning Resource Center, Florida Keys Community College, Key West, Florida

Credit is also given to the MARS Executive Board and, in particular, to Guy T. Westmoreland and to Pamela C. Sieving, for the inspiration which they provided for the program. Special thanks is also extended to Andrew Hansen, RASD Executive Director, for the patience, practical advice, and effort which he extended on the committee's behalf.

JAMES J. MALONEY

Contents

Preface

PART 1. TECHNIQUE

Online Searching: Past, Present, and Future 3

SARA D. KNAPP

Online Searching in Perspective: Advantages and Limitations 16

JANE I. THESING

Initial Considerations for an Online Search 22

CAROL M. TOBIN

Databases and Database Producers and Vendors 26

CAROL HANSEN FENICHEL

Physical Requirements: Terminals, Printers, and Furniture 35

JANET BRUMAN

The Interview Process in Online Searching 50

LAWRENCE R. MAXTED

The Mechanics of Online Searching 57

KRISTINE SALOMON

PART 2. MANAGEMENT

Library Organizational Patterns in Online Retrieval Services 75

PETER G. WATSON

Who Should Search?: The Attributes of a Good Searcher 83

RANDOLPH E. HOCK

Library Applications of Database Searching 89

M. R. DUSTIN

The Policy Manual for Online Reference Services 103

GERTRUDE FOREMAN

Forms and Record Keeping for Online Searching 107

GAYLE MCKINNEY

Costs, Budgets, and Financial Management 123

NANCY E. GRIMES

Methods of Funding 135

JOHN EDWARD EVANS

Options in Training and Continuing Education 149

JAMES RETTIG

Publicizing an Online Search Service 160

MARYJANE S. COCHRANE

The Impact of Online Search Services on Libraries 166

REBECCA WHITAKER

APPENDIX

An Introduction to Online Searching: A Suggested Outline 173

MARS EDUCATION AND TRAINING OF SEARCH ANALYSTS
COMMITTEE

BIBLIOGRAPHIES

Online Reference Service: How to Begin 181

EDITED BY EMELIE J. SHRODER

*The Evaluation of Information Retrieval Services:
A Selected Bibliography* 189

COMPILED BY RASD-MARS COMMITTEE ON
MEASUREMENT AND EVALUATION OF SERVICE

CONTRIBUTORS 193

Part 1

TECHNIQUE

Part I

TECHNIQUE

SARA D. KNAPP

Online Searching: Past, Present, and Future

Definitions

Online searching has been a subject of interest to librarians for over a decade. In 1982, literally millions of searches were conducted and the ranks of online searchers continue to grow. What is online searching?

Online searching has been defined as a means of finding desired information, usually bibliographic references, by using a computer in an interactive mode.

A *database* has been defined as a collection of information called machine-readable records. In our work, this often corresponds, in content, to a printed index.

Every day hundreds of searchers independently but simultaneously search the vast array of databases made available to them online by producers and vendors in this country and abroad.

Growth of Searches and Databases

The growth of online searching has been remarkable and continues to climb. In 1974, Martha Williams reported that 700,000 online retrospective searches were conducted via online services in the United States and Canada. In 1976, 1,200,000 were reported and the figure had jumped to two million by 1977. That figure had doubled again for an estimated four million searches by 1979.¹ In 1982, Williams reported that the rate of growth of online use is declining slightly "but in absolute numbers the growth continues."²

Paralleling the growth in the number of searches performed was the increase in the number of available databases. Martha Williams's *Computer Readable Data Bases* listed 301 U.S. databases in 1975, 528 by 1979, and the 1982 edition lists over 750 available for public access, worldwide.³

History

The origins of the computer go back two centuries to Leibnitz's ideas for a calculating machine and, more recently, to Babbage's arithmetical precursors, Hollerith cards, otherwise known as punch cards, first used in 1890 for the U.S. Census. During World War II, computers were used secretly for cryptography.

As with many other aspects of our present life, the potential of computers began to be realized in that great surge of development following World War II. The application of computers to handling large amounts of diverse information in novel ways struck the imagination of many, and some accounts foreseeing direct use of large banks of data by untold people made their way into popular and professional literature.⁴

One of the applications of computers during the early 1960s, computer-assisted typesetting, led to the development of the online databases of today. Publishers of large information services, such as those which abstracted journals, discovered that computer-assisted typesetting had a useful by-product. The information contained in the publication was in a form which a computer could read and manipulate. Publishers of these information services had indirectly created a database through this method of typesetting at virtually no additional expense or effort over that involved in normal production of a printed index.⁵

These databases were made available to the public during the early 1960s through the resources, generally, of university computer or information centers. These centers conducted searches of all or specific portions of databases that were loaded onto the center's own computer. The centers also conducted current-awareness searches, in which new material that was added to a database at periodic intervals was searched against the information profile of a client researcher. These current-awareness searches are commonly referred to as *selective dissemination of information* or *SDI* searches.

Under the present method of conducting a database search, the librarian interacts directly with the database through a computer terminal. The results of the search can be known immediately. In contrast, clients of computer or information centers of the early 1960s would often wait for days or perhaps weeks after submitting a request to see the results of the search, because the center batched, or combined, several requests in order to reduce the operating cost of each search.

Several early services were operational in batch mode. In 1964, the National Library of Medicine (NLM) began offering a batch service on requests that had been referred to it.⁶ In batch-mode searching, an entire search request is submitted, neither too broad nor too narrow, and then the searcher waits two weeks for the search results. Batch searching is

rather like playing the blindfolded game of pin-the-tail-on-the-donkey; you make what you hope is a stab in the right direction, but you lack the feedback to know how to adjust your strategy.

Actually, online searching was first investigated in the 1950s and was first publicly demonstrated by the Systems Development Corporation (SDC), in 1960, but that system lacked the ability to use Boolean operators or to back-reference previous search statements. SDC began development of the precursor of its retrieval language, ORBIT, in the mid-1960s. This also was later improved and used in NLM's Elhill system for MEDLINE which in the 1969-71 period was operated by SDC.⁷

Lockheed began in 1965 to develop its DIALOG system to provide an online system for NASD and began regular searches services by 1967.⁸

In 1968, the State University of New York (SUNY) Biomedical Communication Network became operational, offering online access to MEDLARS, NLM Current Catalog tapes and a database of medical library monograph holdings. The network used an early system, IBM Document Processing.⁹

Since I learned searching on that old, rather primitive system, I would like to share some recollections of what searching was like then. First of all, the terminals were hard-wired to the computer. No dial-ups, you just left your terminal on all day. (Every now and then it would type a message such as that the system would be down tomorrow or that the SUNY Central Staff wished you a Merry Christmas.)

The system allowed you to back-reference search statements but you had to put in your entire strategy before you got any results. To change it you had to re-enter the entire strategy. You couldn't print descriptions for citations so you couldn't always tell which part of your strategy was causing problems.

We had an old clunker of a terminal. It was noisy and printed only a 110 baud (10 characters per second), which we thought was just fine. Response time of a minute or two didn't seem so bad, either.

In reminiscing with other searchers I have wondered which of our expectations of that period had or had not been fulfilled and how the outcomes of those expectations that might help us look at the predictions of today.

I know some of us expected online searching to grow far faster than it did. It seemed so obviously wonderful to me that I was amazed at the lack of interest and even opposition to it among some contemporary librarians. I think one can say that change is more likely when those responsible for it perceive it as being in their own interests. Many librarians and administrators did not so perceive online searching.

Another librarian told me she had thought we'd be able to integrate searching into local operations, like being able to have call numbers on searches and run them against a local holdings database. Integrated searching operations, too, have waited on technology and incentive for development.

Others of us hoped for standardization among database elements. Little has been done in this area, probably because until now many producers have seen little value in it. Some vendors who recognize users' needs have compensated for the failure of producers to standardize.

But a number of changes were necessary for the evolution of our present systems. One of the most significant was the development of inexpensive dial-up communications linking remote terminals to the large computers of the online vendors.

Packet Switching and Telecommunications Networks

This development of packet switching and telecommunications networks now makes possible worldwide online access via networks such as Tymnet and Telenet. Being hard-wired to the old SUNY BCN (Biomedical Communication Network) computer meant that one line from the computer was dedicated to us and our terminal was dedicated to it. It limited the number of users who could be accommodated.

Tymnet, first offered by Tymshare, Inc., in 1969, was the first widely available data network which made possible the present flexible and more economical mode of communication. "Based on 'packet switching' these systems divide the input flow of information into small segments or packets of data which move through the network in a manner similar to the handling of mail but at immensely higher speeds."¹⁰ By so doing, the same lines can be used for many communications from many widely dispersed users at the same time. The Tymnet network was based on minicomputers which linked various terminals to central computers. The network switches, which were linked by voice-grade telephone-type lines, could store and then forward the data from node to node, using the fastest route to a destination.¹¹ Today, the networks—Tymnet, Telenet, Uninet, and others, too—translate signals from all kinds of dissimilar terminals.¹²

Development of National Search Services

The emergence of packet switching and telecommunications networks made possible the further development and expansion of the national search services which we know today.

Before the advent of Tymnet in 1969, accessibility to online databases was limited. Lockheed had developed DIALOG, which made a handful of scientific and technical databases available to the National Aeronautics and Space Administration (NASA). A version of DIALOG had also been developed by Lockheed for the European Space Research Organization, which was later used by the European Information Retrieval System (IRS). SDC's ORBIT system was available to the U.S. Air Force over telephone lines. The SUNY BCN was providing medical libraries with access to databases created by the National Library of Medicine.

By 1971, NLM's MEDLINE had become operational as a commercially available database. In 1972, both DIALOG and SDC began to enlarge the small selection of their scientific and technical databases, by making their services available to a wider market of users. BCN, which was the predecessor to Bibliographic Retrieval Services (BRS), ceased in May 1977 when BRS became operational.

Of the national search services referred to, DIALOG, SDC, and BRS are commercial vendors of databases. Database vendors contract with the producers of databases, such as the National Library of Medicine or the Educational Resources Information Center (ERIC), to make the databases which they create available through their systems. A vendor system encompasses the computer equipment used to store multiple databases, the staff of technical and public services personnel employed to maintain that equipment and to facilitate its use for the public, and the software developed to enable users to search those databases.

Database producers are the federally supported institutions, agencies of the federal government, professional associations, and corporations that produce the machine-readable records that usually correspond to a printed index. Most database producers simply make their machine-readable product available to the public through one or more of the database vendors. Some database producers such as the National Library of Medicine, however, have made their databases available through their own vendor systems, as well as those of the commercial vendors. This marketing practice has worked well for a producer such as NLM which, in response to the information needs of the medical community, has developed several different databases sufficient to warrant maintaining a separate vendor system. There is a trend among database producers, which I will discuss later, to develop their own vendor systems which allow them to market their databases directly to the public.

The momentum of developments in database searching during the 1970s continued with the emergence of online users' groups in the mid-1970s. In 1976, requests to join RASD's Information Retrieval Committee were so numerous that the chair, Peter Watson, proposed forming a

Machine-Assisted Reference Services Discussion Group, and so MARS was born at the ALA Conference held in Chicago that year. By 1978, MARS had so many members with so many interests that it was able to become an RASD section with its own committees. Two new publications, *Online* and *Online Review*, emerged in 1977, followed by *Database* in 1978. In 1982, *RQ* began carrying reviews of databases.

Recently we have witnessed the widespread conversion to 1200 baud, a development that has reduced connect time for searching. Conversion to 1200 baud has led to changes in pricing by database producers. Many of them are adding "per hit" charges, which means that the searcher pays whenever he or she finds something in the database, as well as for the time connected.

The Future

At least one expert predicts that in the next ten to fifteen years major changes in the way we interact with computers will be possible. These include natural language capabilities, speech understanding, and speech generation. In addition, computing costs will be low enough to make these features feasible.¹³

Another expert predicts an end to offline printing, with new pricing schemes to accommodate that change and the documentation of search results by microcomputers for subsequent printing and manipulation.¹⁴

End Users and Information Professionals: Changing Roles

Will our patrons learn to search for themselves? What will this do to us? There is absolutely no doubt that direct searching by users is in the offing. Already, user-friendly interfaces have been developed. Easy systems using lists of choices and simple commands are available to home users, and some are being marketed to libraries as well. Admittedly these are not yet very sophisticated or powerful systems, but far greater possibilities are now being developed. Systems linked with full text databases, which will allow users to browse easily, can literally put a library of information at one's fingertips. Using the kind of associational trails dreamed of by Bush and others since the 1940s, users will be able to follow their own logic and their own associations to the exact information sought.

The users will be ready for them. Some of our generation may have faced terminals with trepidation, but that will certainly not be the case with today's youngsters who use microcomputers in schools and computer games at home and who find terminals and computers quite a normal part of everyday life. One college president is predicting that

within a few years college freshmen will be required to have their own microcomputers.¹⁵ Even today's adults find user-friendly systems comfortable and enjoyable to use. The recent report of the "paper chase" self-service system indicated a fantastic response by users of a medical library.¹⁶

What of the role of libraries and librarians? Where will we be? Experts seem to differ. One well-known writer points out that, although the do-it-yourself trend is widespread, we frequently pay others to mow lawns, paint houses, repair our cars, and even do our taxes. Presumably there will always be those who want to have someone else do their searching. Others point out that searching is very personal. They foresee such sophisticated and yet simple-to-use systems taking users beyond citations to actual data with so much assistance that neither searchers nor libraries will be needed.

The difference between these views may just be a question of how far in the future you care to look. When we think of doing searches in the near future, we think in terms of present systems. It seems incredible that very many end users would want to learn the myriad conventions, codes, access points, and controlled vocabularies we now use in searching. User-friendly terminals and systems will be quite different and it is likely that as they evolve, more and more of the searches we now do will be more easily done at the user's own terminal and convenience. Nor will users have to wait for documents referred to by databases, since they will also be online. Present systems are indeed cumbersome, unwieldy, and demanding of the user.

In the immediate future, however, there are at least five roles for information professionals:

1. To continue to provide searches and service to those who are disinclined or unable to do them for themselves¹⁷
2. To do the difficult, unusual, or offbeat searches not so readily handled by the first generation of user-friendly systems. An important part of this role may be in assisting users to analyze their search requests and in helping them to think about their questions¹⁸
3. To educate users about available information services and help them learn to use them¹⁹
4. To act as advocate for the public, especially our own user communities, in identifying and working to obtain the kinds of services needed that do not exist and in evaluating the effectiveness of those that are in existence²⁰
5. Another role, suggested by Information Industry Association (IIA) President Paul Zurkowski, is to be a sort of coordinator of locally produced databases.²¹

Producers and vendors have a tremendous interest in promoting end-user searching. When everyone has their own terminals, the market will be almost unlimited. Intermediary searching can never promise the volume of business that end-user searching will provide. So there are tremendous economic incentives for bypassing intermediaries. And technologically, the combination of artificial intelligence and microcomputers as terminals promises to make searching very, very easy.

Private/Public Sector

The question of public or private funding of the production, distribution, and use of online information systems will surely have a profound effect on the future of online services. The economic argument over whether those who use a service should pay for it began with the issue of user fees in publicly supported libraries and now extends to the propriety of government as database vendor, for example in providing the MEDLINE network, or as the producer of such widely used databases as ERIC, AGRICOLA, MEDLARS, NTIS, and the National Institute of Mental Health.

Advocates range along a continuum **from** those who believe government has an obligation to provide **certain kinds** of information services to those who would merely refrain from restricting the government role to those who would carefully limit the role of government. There was disagreement on these fundamental issues between members of the National Commission on Libraries and Information Science Public Sector/Private Sector Task Force and I am sure there is disagreement among ALA members. The report of the NCLIS Task Force seems to have as one of its most important but controversial elements the encouragement of private enterprise in the development of information as a national resource. Reliance on the marketplace would make the criteria of value economic rather than political. Reducing the value of information solely to economic criteria is controversial: we might argue that unless everyone has the dollars to spend on information, economic factors aren't always a fair measure of the public's need.

But there is certainly much to be said for private enterprise as a means by which products are manufactured and distributed. It is private enterprise that put all sorts of appliances and products in our homes, and it is generally believed that profit encourages competition . . . which tends to lower prices and encourage the widest possible dissemination through marketing. Obviously this competitive element is going to apply to products and services for which there is demand by those who can afford to pay, in other words products and services which meet economic criteria.