
Chapter 1

Introduction and Background

Introduction and Background

The significance of information to society is becoming increasingly apparent with advances in technology. Most nations today realize that they must be able to obtain and efficiently process information vital to national life. They must especially be able to do this with respect to health information. The transfer of scientific information to researchers who require it in the conduct of their investigations and to practitioners of health care is essential for the health of the American people.

The National Library of Medicine (NLM, or the Library) is the Nation's—and in many respects the world's—principal resource for the collection, organization, and retrieval of scientific literature in the health and biomedical fields. Its efforts complement the Nation's investment in biomedical and other health-related research, and in medical education. NLM's purpose, as expressed in the National Library of Medicine Act (Public Law

84-941), is “to assist the advancement of medical and related sciences, and to aid the dissemination and exchange of scientific and other information important to the progress of medicine and to the public health.”

NLM is a complex organization that performs diverse and far-reaching activities, extending well beyond what is customarily considered as traditional librarianship. (See app. A for a description of NLM, including its organization, appropriations and staffing, and intramural and extramural activities.) It has funded grant programs that have enhanced the collections of academic health center and community hospital libraries, promoted library consortia, and supported the research and development of computer applications to medicine as well as to information science. The Library has also supported the development of a national system of regional medical libraries, which links NLM and local libraries.

HISTORY, OBJECTIVES, AND SCOPE OF THE STUDY

Two congressional requests to OTA prompted this study. While expressing different concerns, the two are complementary. The first request was part of a larger request from the House Interstate and Foreign Commerce Committee for a study addressing the techniques and methods available for assessing medical technologies. * The committee specifically asked OTA to examine the performance of MEDLARS, particularly the performance of its major biomedical data base MEDLINE, in disseminating health-related bibliographic information. ** The second request, from the Sen-

ate Labor and Human Resources Committee, resulted from interest in issues raised in an OTA staff paper on NLM and from hearings on reauthorization of the Medical Library Assistance Act held in April 1981. It asked that OTA explore the Government's role in the creation and the distribution of health-related information by means of computerized bibliographic retrieval systems.

Thus, this technical memorandum has two objectives. The major objective is to examine NLM's role in the creation and distribution of computerized health-related bibliographic information in light of the private sector's presence in this field and the public interest. NLM's role with respect to the transfer of bibliographic information is not a unique issue, but is part of a growing concern and discussion about the Government's role in the creation and the dissemination of all types and forms of information. *

*This issue is but one of many national and international policy issues. App. B provides a contextual setting for the specific information.

(continued on next page)

*See OTA's study *Strategies for Medical Technology Assessment* (117).

●*Since the inception of NLM's computerized system, the terms MEDLARS and MEDLINE have had a number of definitions. The most recent definitions and those used throughout this technical memorandum follow. MEDLARS (Medical Literature Analysis and Retrieval System), NLM's computerized retrieval and technical processing system, is a complex IBM multiprocessing system that maintains data files, provides on-line retrieval services and produces computer-photocomposed publications. MEDLINE (MEDLARS on-line) is the largest and most extensively used of NLM's data bases.

The second objective, which stems from the House Interstate and Foreign Commerce Committee request, is to examine MEDLARS' effectiveness in disseminating bibliographic health-related information. Part of the response to that request is contained in the OTA staff paper "The National Library of Medicine" (116) and in the OTA report *Strategies for Medical Technology Assessment* (117). This technical memorandum analyzes three

(continued from p. 3)

mation policy issues of interest by briefly reviewing the history, current standing, and future prospects of domestic information policy and describing the relationship between domestic and international information policy.

BACKGROUND FOR THE STUDY

One of the mandated functions of NLM is the collection and preservation of health-related primary library materials, such as books, periodicals, prints, films, and recordings. Another mandated function is to organize the primary literature and publish and make available indexes, catalogs, and bibliographies—i.e., secondary literature—in order to locate relevant primary literature.

Although both the public and private sectors have a long history of providing bibliographic biomedical information products and services—secondary literature—the Library of the Office of the Army's Surgeon General, the forerunner of NLM, began the first printed index to the biomedical literature in 1879 with *Index Medicus*. The index, which primarily contains references to published journal articles, was prepared manually until 1964, when MEDLARS mechanized the processing and printing functions. The index records were put into machine-readable form, thereby expediting the production of the printed *Index Medicus* and making these records usable as a machine-readable data base. Thus, in 1964, NLM started the first large-scale, computer-based, retrospective search service available to the general public. In 1966, as a result of research in both the public and private sectors, Lockheed Information Systems (now DIALOG Information Services, Inc.), a commercial firm, made available

system relevant issues: 1) the subject content of the literature cited in the MEDLARS data bases, in particular MEDLINE; 2) the coverage of the nonserial literature in the MEDLARS data bases, in particular MEDLINE; and 3) the evaluation of the methodology and statistics used in the literature cited in the MEDLARS data bases. OTA was not requested to, and has not attempted to, analyze the effectiveness of private sector computerized bibliographic health-related information products and services. Nonetheless, the issue of assessing the adequacy of methodological design and statistical analysis in cited literature pertains to all health-related data bases.

the first on-line search service on a regular production basis. *

Today, the data tape used for the printing of *Index Medicus* and other NLM printed publications is also used as the source of data for the data base MEDLINE. Subsets of the MEDLINE data base are incorporated into some of the other data bases that are now available on-line at NLM. In addition to producing data bases and data base products, NLM now provides direct on-line machine searching of the contents of the data bases to the information and health communities. The Library is connected by a complex telecommunications network to more than 1,800 terminals in institutions in the United States and abroad. It also leases tapes of its data bases to two commercial U.S. firms which disseminate the information on-line from their computer to their customers. Organizations in a number of foreign countries also lease some of the data bases.

Worldwide, *Index Medicus* has been the primary means for access to medical information for more than 100 years. In the past 17 years, MEDLARS has likewise achieved an impressive

* With on-line access, a person at a computer terminal can carry on a dialog with the computer and direct it to locate information, retrieve it, and provide it either at the terminal or in printed form for mailing to the requestor.

reputation. By providing on-line access to its bibliographic data bases, MEDLARS has made NLM's resources more accessible operationally and geographically to almost all segments of the biomedical community. Furthermore, NLM has sought to maintain a coverage of the biomedical literature that includes subjects of current interest, and the Library now serves a diverse constituency including many health science disciplines outside of traditional medicine.

There are, however, two current issues that challenge NLM and MEDLARS. One *issue is how to establish a suitable equilibrium between MEDLARS and the changing needs of its users.* Although the scope of MEDLARS data bases is limited, the quantity of knowledge in traditional biomedical subjects is increasing. Furthermore, much current biomedical research is interdisciplinary, and the present concept of health is broad and information pertaining to health can be found in journals in fields such as law and economics that are historically not the Library's province. Moreover, health-related information is often found in a form, such as technical reports and speeches, that is not cited in MEDLINE.

The expansion of the biomedical literature base has been accompanied by rapid progress in the application of computer and communications technologies to information systems, and a tremendous increase in the volume and ease with which information can be accessed. In order to assist the selection process for the health professional who uses the information that is retrieved and to provide a modicum of quality control, the scientific merit of literary material requires more and more scrutiny. One suggestion is that data base producers such as NLM assume some of the responsibility for assessing the evaluative methodology and statistical analysis used in the documents cited in bibliographic data bases.

The problem of accommodating user needs and MEDLARS' limitations is compounded because of the lack of comprehensive data on MEDLARS users. Although there is information available on the location of the terminals with access to MEDLARS by type of institution, information on the ultimate users of the information is sparse and insufficient for defining specific segments of the

user population. (System issues are discussed in ch. 3.)

The second and more pressing problem facing NLM is its role in the creation and distribution of computerized health-related bibliographic information through MEDLARS. Not only NLM but for-profit and not-for-profit organizations in the private sector create health-related bibliographic data bases; such organizations include BIOSIS (formerly Biosciences Information Service), E_x_c_e_r_p_t_a Medica, Information Retrieval, Ltd., and the Institute for Scientific Information. In addition, biomedical data bases produced by both the public and private sectors are vended by three commercial organizations—Bibliographic Retrieval Services, DIALOG Information Services, Inc., and System Development Corp. Nonetheless, health-related computerized bibliographic information is predominantly created and disseminated by NLM through MEDLARS. At issue is whether NLM's computerized bibliographic products and services and the products and services of the private sector substitute for or complement each other and whether NLM's leading portion in the biomedical information field is hindering the growth of the private information sector.

Some members of the information community, and some members of the commercial sector of the private information industry, have become increasingly concerned about NLM's dominant role in the field. Because the industry is heterogeneous and composed of a variety of firms, it does not have one position regarding all of NLM's computerized bibliographic activities. Individual firms have particular opinions about NLM's activities depending on their perspectives. Overall, the industry's concerns are with respect to NLM's preparation of computerized bibliographic health-related data bases, NLM's charges to commercial information services and foreign centers for leasing the data tapes, NLM's provision of direct on-line access to its data bases, and NLM's pricing of direct on-line access to the data bases. NLM's pricing of access appears to be the issue of paramount concern at this time. For the most part, prices are significantly lower than those charged by the private sector for access to health-related bibliographic information. (These issues are discussed in ch. 6.)

Historically, public policy has held that NLM's position in the creation and dissemination of health-related bibliographic information is in the best interests of the public's health and well-being. Policy concerning the Library has been closely tied to biomedical research policy. The political environment in this regard appears to be changing. Although Government-sponsored biomedical research still appears to be of major interest, recent announcements from the Office of Management and Budget stress the importance of increased private participation in Government information activities and the need for "full cost recovery" (undefined) of Government-sponsored information products and services. The Library's role in providing computerized health-related bibliographic information may be reduced if such suggestions are implemented.

This OTA analysis focuses on current NLM issues as they relate to national information policy. These issues have direct importance for institutional and individual users of health-related bibliographic data bases, NLM, private sector producers of health-related data bases, and private information retrieval services (vendors). The issues are also significant for foreign data base producers and information retrieval services, both governmental and private, and for foreign institutional and individual users of health-related bibliographic data bases. Because the issues are similar to those concerning other Federal information activities, their resolution also has implications for Federal and non-Federal organizations that create and distribute Government-sponsored information, private sector information enterprises, foreign information organizations, and all users of Government-sponsored information.

This report's emphasis on current issues of necessity gives inadequate attention to the future

effects of new and emerging technologies on biomedical bibliographic retrieval systems. Current issues result in part from technologies now in use. Evolving computer and communications technologies are still only in the early stages of development, and indications are that they will be much more powerful and varied in the future. With *expected dramatic changes in data base creation and access, some current issues may diminish in importance or disappear, and quite different ones may arise and require consideration.*

But it is not possible to know with any degree of certainty which technologies will be adopted and which issues will become significant. Social, political, and economic forces, as well as technological forces, are instrumental in determining the development and utilization of any innovation. Indeed, *the way decisions are made about current issues may affect which new and emerging technologies are implemented. This underscores the need for flexibility in public policy.*

There are indications that in the future there will be more distributive means of disseminating information than are utilized at present. * Indeed, NLM administrators think that the technologies will be available in the next 5 to 7 years; they are making long-range plans in line with this thinking. Currently, NLM is experimenting with making one of its data bases available for distributive searching in 1 or 2 years. Thus, future technologies and their possible effects on biomedical bibliographic retrieval systems are discussed briefly in this study (see app. H).

*A distributive data-processing system is one which uses multiple small computers to process all or portions of a data base. The small computers can be widely separated and may be linked by telecommunications lines to each other and to a large computer.

ORGANIZATION OF THE REPORT

Chapter 2 describes the development, current status and future plans for MEDLARS' data bases and on-line services. System issues related to the effectiveness of MEDLARS in disseminating bib-

liographic health-related information are discussed in chapter 3. Chapter 4 is a brief description of private sector health-related data bases and commercial information services. The next chapter,

chapter 5, discusses the considerations underlying the current debate on the appropriate role of the Government in information transfer. Finally, chapter 6 analyzes the domestic and international implications of changing the range and pricing structure of MEDLARS computerized prod-

ucts and services. That chapter also considers the effect of new technologies on present issues. There are 11 appendixes included—both for reference and, in some cases, for expanding ideas and issues contained within the body of this report.