Appendix D.— Development of Computerized Biomedical Bibliographic Retrieval Systems

In *every* half-century since 1750, the number of scientific journals increased tenfold, reaching well over 100,000 by 1950, and straining the abilities of readers to stay abreast of the literature and retrieve earlier published information (l). Following World War II, science shifted away from strictly academically based, disciplinary research and toward mission-oriented, multidisciplinary approaches to solving social and technical problems. Thus, scientists had to draw needed information from a body of knowledge that was not only expanding at unprecedented rates, but compiled in increasingly diverse (and seemingly remote) sources.

This dilemma was quite pronounced in biomedicine: 19,000 biomedical journals were printed in ***1950,*** well over ***22, ooo*** in **1970,** and the boundaries that traditionally defined biomedicine as a field of practice and research were continually expanding to encompass new disciplines. Today, the field is so broad it defies attempts at delineation. In **1950,** it was clear that traditional scientific publications were exhausting their effectiveness in communicating important advances and keeping readers abreast of the literature (1). Science turned to computers to accelerate the processing, storage, and retrieval of information, and thereby revolutionized all information-seeking activities.

Information Products and Services

The single most important source of biomedical information are primary publications—books, journals, technical reports, Government studies, patents, etc,—though the dramatic rise in their number often makes access to this literature difficult and confusing. Secondary publications—bibliographies, abstracts, catalogs, and indexes which facilitate access to the primary literature—are thus serving as an increasingly important element in the transfer of information.

Secondary services trace their historical roots to the learned and professional societies that flourished in the late 19th century, and which sought to fulfill the discipline-based information needs of their members in the basic sciences. In general, such services acquire and analyze primary publications, usually provide a bibliographic description of each document, indexing each according to an established system of content analyses, and, periodically, produce a product which provides subject access to the references for the documents indexed during a specific time period (112). Secondary services guide users to the primary source of needed data, rather than directly provide the information. By 1950, 3,ooO abstract journals for the scientific literature were available in print form (1).

Until World War II, the need for interdisciplinary communication was met through a few broad-based professional organizations (e.g., the American Association for the Advancement of Science), the various academies (e.g., New York Academy of Sciences), and multidisciplinary publications (e. g., *Science* and *Nature).* When new scientific or technical areas arose, they were usually cross-disciplinary, like biochemistry. Information access needs could be met by establishing a new section of an abstracting and indexing service: witness the issuing of *Chemical Abstracts* by sub-area of chemistry, today including a section on biochemistry, and one on toxicology.

After World War II, the rise of Government-sponsored, mission-oriented research programs created a need for broader, multidisciplinary coverage of the scientific literature within a single secondary service. While professional (nonprofit) societies addressed this need through subspecialization, commercial organizations began offering selective, focused services relating to particular program areas that cut across disparate subject disciplines (e. g., the environment).

Technological Developments and Growth of On-Line Services

The secondary service adaptations after World War II were hard-pressed to meet the growing needs of the scientific community. In the late ***1950's,*** producers of printed abstracts and bibliographic indexes turned to the new computer technologies to reduce costs by mechanizing the construction and publication of their products. The sophisticated application of computers to the processing of biomedical bibliographic information was pioneered by the National Library of Medicine (NLM), and by **1964** the first computerized biomedical bibliographic retrieval system, MEDLARS, was in use.

Most of the other machine-readable data bases in health-related fields, as well as in other subject matter fields, were originally built when the production of traditional indexing and abstracting journals switched to computer-driven photocomposition. It soon became clear that machine-readable data bases developed in the 1960's for the production of print copy could also be searched via computer: this development became the basis for on-line information retrieval services, These and subsequent developments

were sponsored by both public and private enterprises. Government investment enhanced the potential of communication networks to attract other customers. The increased volume of business then lowered costs sufficiently to allow the industry to be cost effective and to grow.

All libraries have been affected by the electronic revolution, but none so much as those which serve the scientific, technological, and medical communities. Though these libraries have provided and will continue to provide traditional services and print products, the use of computerized services to support and enhance their services is rapidly growing. The growth of computerized services has depended on, and been driven by, the technical development and coordination of time-sharing computers, machine-readable data bases, fast-access disk storage devices, interactive retrieval programs, and low-cost terminals and telecommunications networks. Although this study focuses on data bases and their on-line services, other technological developments were necessary to the growth which we now see.

Machine-readable data bases, which began and continue as a more efficient method of preparing printed indexes, enhancing the value of primary publications, have grown in size and scope to compete with their printed counterparts, and form the basis of information transfer activities. A 1979 directory of data bases available on-line lists **528** separate entries, up from 301 cited 3 years earlier; 90 of these were relevant to biomedicine and health care (166). The number of requests for on-line searches grew from 700,000 in 1974 to over ***4*** million in **1979** and to 6 million in 1981. Over one-quarter of these were conducted on MEDLARS alone by users in **40** countries (1). Specific data on the services provided by commercial firms is considered proprietary and is thus unavailable for publication, though the industry as a whole is said to be enjoying a 20-percent annual growth rate **(168).**

As data bases have grown to be at least as, if not more, important than printed products as a source of bibliographic information, they have diverged from the publications they were developed to prepare. For example, a number of journals in the special interest areas of dentistry, nursing, and population science are indexed for NLM's major data base MEDLINE. However, these citations do not appear in Index Medicus. On-line products are beginning to contain more information than their printed counterparts, clearly the result of the economics of the process of information transfer to secondary sources. It is simply less expensive to add information to a data tape than to a printed product (112).