
Evolution of Computerized Bibliographies

WHEN Billings began to develop the *Index-Catalogue* in the 1870's he unwittingly converted the Library into a publishing house, half or more of whose employees would spend their working hours year after year preparing annual bibliographies. The manual indexing, arranging, and editing for the *Catalogue* continued for seven decades, diverting the staff from other important library operations. Not until the 1940's did the situation change as directors and editors seized new techniques for processing and transmitting information.

DEVELOPMENT OF A MECHANIZED SYSTEM FOR PRODUCING THE LIBRARY'S PUBLICATIONS

The repetitive sorting, filing, and photographing that occurred every month during the production of *Current List of Medical Literature* suggested that much of the work could be done by machines. During the 1950's as the number of journal articles continually increased, straining the facilities of the Library to publish *Current List*, editor Seymour Taine had many discussions with Director Rogers about the possibility of mechanizing the operation. He examined equipment being used by business firms and government agencies to process data and concluded that the methods and machines could be applied in the Library. He drew up a plan to abandon the "shingling" procedure for producing the *List* in favor of punched cards which could be sorted by machines and photographed by an automatic high-speed camera to make a photo-offset negative. Since the system would be made up of data processing equipment, Rogers and Taine hoped that it could also be used for the selective retrieval of bibliographical data.¹

Rogers was unable to obtain appropriated funds to buy or rent machines, but at a meeting of the Board of Regents members suggested that he apply for a grant from the Council on Library Resources. Rogers did this, and on April 16, 1958, the council allotted \$73,800 for the Library to undertake the work.²

In order to understand more fully the potentialities of available equipment Taine and other staff members attended courses on data processing at the International Business Machines school in Endicott, New York. A room in the Library building was soundproofed, air-conditioned, and otherwise prepared to accommodate equipment. Consultants were engaged to assist. An advisory



Listomatic Camera photographing citations for Index Medicus being operated by Tyrone Ferguson, right. Pages being spliced together by Donald Dodson, left.

committee was appointed. Various tape operated typewriters, tabulating equipment, and other machines were evaluated for use, and the most satisfactory were ordered. The heart of the system was to be an Eastman Kodak Listomatic camera capable of photographing 230 punched cards a minute while adjusting its aperture to accommodate one, two, or three lines of text imprinted across the top of the cards.

Early in the experiment it was decided that the arrangement of citations in *Current List* would be changed. The grouping together of titles of all articles in a journal was almost a necessity for the "shingle" method of production, but here it would be preferable to place one complete citation on each punched card. The final publication would contain two sections; one would list citations under subject headings in alphabetical order, the other would list authors in alphabetical order.

There were a number of possible ways of arranging the flow of work, and much thought went into the development of the best system. Finally Taine settled on seven stages, starting with an indexer who scanned articles, translated foreign language titles, assigned subjects and subheadings, and typed this information on a form; an indexing assistant who added authors' names, other bibliographical information, and machine codes to the form; an input typist who turned out a proof copy and a coded punched paper tape; a key-puncher who punched subject and author cards; and an operator who ran tapes and cards through output typewriters to produce imprinted cards. The cards were sorted and interfiled by machines, then collated with cards bearing headings, subheadings, and cross-references. After corrections were made, the deck was

interfiled with a "program" deck containing the page numbers of the pages to be printed and other information. This complete deck was run through the camera, the film developed, cut into columns, the columns taped together into pages, and the pages sent to the printing firm.

Through the spring of 1959 Taine and his associates perfected the mechanized system, adjusting one part or another, replacing inferior components by better, removing bottlenecks and improving procedures. In May 1959 citations for the fourth volume of the *Bibliography of Medical Reviews* were typed on cards and filmed as a test of the new system. The published work, received from the printer in August, was quite satisfactory. The team then began to index for the first number of *Current List* to be published by the new system.³

While Rogers and Taine were happy with the success of the new publication system they were disappointed because it was not practical as a retrieval system. The fastest machine obtainable could sort 1,000 cards a minute. At that rate it would take 12 ½ hours to sort a 5-year accumulation of 750,000 cards, much too slow. Also, there were risks that the massive decks of cards might be mixed if they were disturbed before cumulation.

Before the Library had the opportunity to publish the first issue of *Current List* using the new system, the *List* metamorphosed into the *Index Medicus*. The transformation came about this way. Since 1950 Rogers had sought to find a way in which the Library and the American Medical Association could cooperate in publishing bibliographies. The Library's *Current List* and the AMA's *Quarterly Cumulative Index Medicus* together indexed about one-half of the world's medical literature, but they overlapped, so that about one-third of the citations in the *List* were duplicated in the *Index*. One of Rogers' suggestions to the AMA had been to divide the indexing of the world's literature with the Library, the AMA indexing all publications from the Western Hemisphere and the Library indexing everything from the rest of the world. Another suggestion was that the Library index all articles in foreign languages, the AMA all those in English. For various reasons none of the proposed methods of cooperation was acceptable.

After the mechanization experiment began Rogers saw another way by which the two organizations could cooperate. The Library could prepare the monthly index and the AMA could produce the annual cumulation. Every December the Library could alphabetize its accumulation of cards, photograph them, and send the film to Chicago, where the AMA could use it to publish the cumulation. The Library would benefit by not having to publish a cumulated volume, the AMA would benefit by not having to publish monthly volumes, libraries would benefit by having to purchase, handle, and shelve only one index instead of two, and readers would benefit by having to scan one index instead of two.

In June 1959 the AMA board of trustees and house of delegates endorsed the plan. The following month the Public Health Service and AMA signed an agreement under which the Library would publish a monthly bibliography, named *Index Medicus*, superseding the *Current List* and *Quarterly Cumulative*

Index Medicus, while the AMA would publish an annual volume titled *Cumulated Index Medicus*.⁴

The first number of the new *Index Medicus* appeared almost on schedule in late January 1960. Through the year the monthly issues became larger and more current as the team became more skillful in operating the mechanical equipment. By early January 1961 the cards from the monthly batches had been cumulated, interfiled, photographed, and the film shipped to Chicago. Some portions of the film were found to be defective and the work had to be done over, but even at that the three-volume *Cumulated Index Medicus* covering the year 1960 came from the AMA presses in April 1961.

MEDLARS

During the 3-year period in which Taine, Rogers, and their associates were using data processing equipment to prepare *Index Medicus*, they learned much about an alternative means of storing and retrieving information then being developed and installed in government agencies and business firms, namely the first generation of commercial electronic computers. Even before the mechanization project proved itself, Rogers was familiarizing himself with computers by reading, talking to experts, and attending courses in symbolic logic, computer operations, and related subjects at the Department of Agriculture Graduate School.

While computers held the promise of filling the Library's needs, they were expensive in relation to the Library's funds (NLM's appropriation was approximately \$1.5 million in 1960). But it so happened at this time that the Council of NIH's National Heart Institute had become concerned over the problem of bibliography in the field of cardiovascular diseases. A subcommittee of the Council studied the matter. There were vigorous debates in the Council about the methods of gaining control over the literature. James Watt, director of the Institute, and Rogers had several discussions about the possibility of using a computer to store and retrieve information. The Heart Council decided that computer retrieval of information was feasible and that it had the authority to support the development of a computerized bibliographic system in the Library. They offered to finance the work if, in return, NLM would give some priority to literature on cardiovascular diseases when the system was completed.

In November 1960 the Library engaged an analyst to draw up specifications for a system that Rogers and Taine named MEDLARS, the acronym for medical literature analysis and retrieval system.⁵ The preliminary specifications were so unsatisfactory that Rogers discarded them and, with Taine, rewrote them. The heart of MEDLARS was to be a digital computer. Information representing the indexing done by the staff was to be fed into the system, converted to magnetic tape, and manipulated in the computer. The processed magnetic tape would be used to activate a high-speed composing device capable of producing photographic masters for printing *Index-Medicus* and other publications.

EVOLUTION OF COMPUTERIZED BIBLIOGRAPHIES

NLM desired a system that would permit an increase in the number of journals indexed in *Index Medicus*, reduce the time required to prepare the monthly issues of *IM* from 22 to 5 days, produce for publication bibliographies similar to *IM* in format devoted to special fields, permit a search of the data base and retrieval of bibliographies for patrons upon request, include citations from books and other nonjournal sources, and reduce the need for duplicate literature screening operations at other libraries and information centers.

The development of MEDLARS was planned to take place in three phases, each under a separate contract. During phase I the contractor would make a preliminary design of the system, evaluate equipment available on the market, and select the equipment. In phase II all the engineering would be completed to ready the system for operation, major programs would be written for the computer, final specifications for equipment would be written, and operators would be trained. During phase III, which would overlap phase II, equipment would be ordered and installed. The entire system was to be ready for operation in the fall of 1963.

In February 1961 NLM sent invitations to bid to more than 45 firms that designed computer systems. Publicity about the project led other firms to ask for information, and Rogers extended the deadline. As replies from firms arrived, Rogers and Taine, later assisted by representatives of DHEW, Department of Defense, National Bureau of Standards, and Central Intelligence Agency evaluated them. Rogers hoped to have the contract for phase I of the project signed in June, but because of delays August arrived before it was signed, with General Electric Company.

A very competent General Electric team, composed on the average of six persons headed by Richard F. Garrard, began work on August 14. Taine acted as project director for NLM. For weeks thereafter the GE team, working with NLM employees assigned to the project, studied and amplified the MEDLARS requirements, determined possible system and subsystem configurations, and recommended various machine configurations. Important basic decisions were made, among them the decision to continue using the Library's medical subject heading list, MESH, instead of natural language or other indexing approaches; to index each article only once, for both publication and retrieval purposes; to train specialists to retrieve information for patrons; to use serial magnetic tape files for storing citations rather than random access devices; to segment computer programs into self-contained modules for ease of maintenance and system changes; and to develop a rapid, high-quality photocomposition device for preparing copy for *Index Medicus*.

They decided that only commercially available equipment would be used, with exception of the photocomposer. Eighteen different computers were considered. Seventeen were eliminated because they were too large, too small, too slow, or for other reasons, leaving the Minneapolis-Honeywell 800 as the survivor. A MH 800 was already in use in NIH, and the GE-NLM team considered the possibility of sharing this computer. But a study indicated that

sharing was not practical, and that the money saved would not be sufficient to justify the inconveniences and problems that would arise with joint usage.

During the design phase Rogers assigned to Deputy Director Scott Adams the task of considering secondary objectives for the system. Consulting with GE, Adams proposed several, one of which was the "decentralization" of MEDLARS. A national network of MEDLARS centers was to be set up, each with a search capability and duplicate copies of the master tapes. Searches could then be made locally instead of being made in Bethesda, and the number of searches that could be run daily would be increased tremendously. Other objectives were: the storage and retrieval of text images on microfilm, permitting copies of text to be provided to searchers, probably through a separate device linked to MEDLARS; the use of on-line and remote input and output facilities such as data inquiry and display stations; and the processing of internal library transactions, such as acquisition procurement control, interlibrary loans, inventory control, and other requirements of a similar nature.

The preliminary design phase of MEDLARS was completed on January 31, 1962. The contract for the next step, the detailed design phase, was not signed by GE until the summer, but work continued nevertheless. During phase II specifications were laid out for equipment, operators were trained, development was begun of composing equipment that would print *Index Medicus*, and the computer system was developed to the point where it would operate. An idea of the magnitude of the task may be visualized by noting the time required for programming, 30 man-years. The program was tested and "debugged" on a computer at Army Map Service until the Minneapolis-Honeywell 800 arrived at the Library in March 1963. The computer room had been carefully designed, and the computer was installed and tested without major problems. GE had provided on-the-job training for operators of the system. After the computer was installed the operators kept it busy an average of 12 hours a day debugging, modifying, and integrating program modules. The Library assumed responsibility for maintaining the system in February 1964.

In operation the new system began with indexers, each of whom was assigned a number of journals every day. Starting with the first article in a journal an indexer typed the author's name and other bibliographic data on a printed form. Because of the large number of articles scheduled for citation in *Index Medicus*, the indexers did not have time to read every word in an article but used a rapid scanning-reading method of deciding what each article was about and selecting the subject headings and subheadings to be cited, using terms on NLM's medical headings subject list. Approximately nine subject headings were assigned to the average article, editorial, letter, or other item. The indexer typed these terms on the form, placing an X along side of those destined for publication in *Index Medicus*, leaving the unmarked terms to be inserted into the computer for retrieval during searches. The indexer also checked on the form preprinted terms appropriate to the article, such as the age of a person discussed in the article. The form was attached to the article, and each suc-

EVOLUTION OF COMPUTERIZED BIBLIOGRAPHIES



The Honeywell computer, workhorse of MEDLARS I, was responsible for processing over 10 years of Index Medicus, Cumulated Index Medicus, Current Catalog, and numerous other bibliographies used throughout the world. It processed thousands of demand searches during the years before on-line searching took over. The system was also instrumental in providing data which led to the development of its own successor MEDLARS II.

ceeding article in the journal was indexed in the same manner. The journals were passed to a reviser who checked the form against the article quickly to confirm the indexing. After revision the information was typed on paper tape in machine-readable form, proofread, the tapes spliced into batches and fed into the computer. Through the computer's input program the information on the paper tape was recorded on reels of magnetic tape, edited, and incorporated into data files.

On the retrieval end of the system the process began with a patron's writing his request on a special form and mailing it to the Library. There a librarian, using the medical subject headings list vocabulary, coded the request into a form acceptable to the computer. In MEDLARS, as with most of the computer systems developed at the time, the entire file of magnetic tapes had to be searched in sequence in order to retrieve all of the citations on a given subject. In 1964 it took about 40 minutes for the computer to read all of the tapes. Therefore it was not practical to process only one search at a time. Instead a batch of requests was inserted into the computer which retrieved citations for all of them in one sweep and sorted them at the end. The retrieved list of citations was reviewed by the specialist who had coded the request and mailed to the patron.

As might be expected with large new systems, the development of MEDLARS did not proceed altogether smoothly. One problem was the revision of the medical subject headings list, MESH. The existing MESH list, developed during the previous decade for NLM publication, contained 4,500 headings and 67 topical subheadings. Winifred Sewell and her associates worked hard

integrating 1,400 new headings, eliminating subheadings and some old headings, and producing a list of 5,700 subjects. Inconsistencies and deficiencies were noted, and changes were made. Decisions had to be made as to which subject headings would be used in *Index Medicus*, and which would be placed in the computer only for searching.

Other troubles occurred during 1963 when the Library was switching from the old method of producing *Index Medicus* to the new, forcing indexers to process journal articles twice, once using the old MESH list to index articles for the current issues of *Index Medicus* and once using the new MESH list to index articles for a computer test tape. The double processing caused the production of citations to proceed slowly. By the end of 1963 the regular production of *Index Medicus* had been maintained, but only 45,000 citations had been stored in the test tape instead of the 150,000 that had been expected. This so-called "conversion" tape was useful in testing MEDLARS, but it contained so many inconsistencies that it was not used for searching after 1965.

It had been foreseen that the development of the output system that would print *Index Medicus* might present more problems than the development of the computer segments. The intention was to link the computer to a device that would compose the text of *Index Medicus*, perhaps a mechanized line-printer of the type commonly used with computers or a mechanical photocomposition device or a cathode ray tube device. A study of the existing printers showed that none was satisfactory; they were too slow or had a poor typographical appearance or possessed other drawbacks.

Seymour Taine had persuaded Rogers to include in the MEDLARS requirements one for a fast photocompositor. General Electric arranged a sub-contract with Photon Company to develop a new phototypesetter that would operate at high speed, using a magnetic tape input following computer transposition and editing. This device was first named GRAC (graphic arts composer) but this sounded somewhat like a monster in a science fiction movie and was changed to the felicitous GRACE (graphic arts composing equipment).

GRACE was scheduled for delivery in May 1963 to allow time for testing before being used to print *Index Medicus*. The engineering problems involved were difficult, and their solution took longer than expected. Several of the Library's knowledgeable visitors were very pessimistic about the slow progress of GRACE. Finally in the summer of 1963 Rogers was faced with a major decision, to delay the inauguration of MEDLARS or to find another way of printing *IM*. He decided to follow the schedule as closely as possible, and he ordered the system be placed in operation in January 1964.

Martin Cummings, who succeeded Rogers as Director in January 1964, closely monitored the construction of GRACE. He exerted pressure on the contractor and subcontractor to achieve the high quality they had promised, and he had everything in readiness for production when the machine arrived during the summer. In the meantime a conventional computer printer, loaned

by Documentation, Inc., composed the issues from January through June, and an IBM printer with a more attractive font turned out the July issue. GRACE began to produce *Index Medicus* in August.

GRACE accepted input from a magnetic tape that had been coded by the computer. The machine contained a matrix of 226 characters etched on a glass plate; these characters were in several fonts, with a complement of diacritical marks. Behind each character was a high-speed flash tube. The circuitry of GRACE timed the flashing of these lights. Between the matrix plate and a 9-inch wide roll of film was a mirror and reciprocating lens, constantly moving back and forth, photographing a line of characters, character by character, across the width of the page, at a speed of 1.7 seconds per sweep. The exposed film, in 100-foot-long rolls, was developed automatically in another machine, cut into page size strips and sent to the printer.

At the time GRACE was the fastest computer-driven photocomposer in the United States. It operated at a speed of 3,600 five-letter words a minute, more than 25 times the speed of previous phototypesetters. The August issue of *Index Medicus* contained 13,733 different citations and approximately 1.8 million five-letter words. GRACE needed only 16 hours to set and compose the type. GRACE continued to phototypeset publications until 1969 when, overtaken by age, it was replaced by an improved, faster commercial model Photon Zip 901, and sent to rest at the Smithsonian Institution. It was estimated that during its active life, from August 1964 until March 1969, GRACE composed 165,000 pages for *Index Medicus* and other bibliographies.

During the break-in period of MEDLARS, 1964, imperfections became evident in the analysis and retrieval system. There was need for improvement in the practices and techniques of indexing and for the recruitment and training of persons capable of carrying on the rigorous in-depth indexing necessary to prepare information for insertion into the computer. There was need for a method of evaluating the relevance of citations retrieved from the computer upon request and of estimating the percentage of citations missed during retrieval. There was need for computerization of the Library's technical activities, such as cataloging. There was need for further refinement and modification of the medical subject headings list. The latter was particularly important; the high quality of *Index Medicus*, of the proposed recurring bibliographies, and of individual searches depended upon MESH.

Director Cummings acted quickly to improve the list. He set up a MESH group headed by Peter Olch, a pathologist whom he brought from National Institutes of Health. Through the efforts of Olch's group MESH was expanded and developed further logically. Subheads were reinstated, the number of cross references was increased, and hierarchical or tree structures were developed for a number of categories. The MESH section received advice on special terminology from several professional groups, among them representatives of the epidemiology section of the American Public Health Association, the American

A HISTORY OF THE NATIONAL LIBRARY OF MEDICINE

Dental Association, *Journal of Medical Education*, *American Journal of Nursing*, and Chemical Abstracts Service. Cummings maintained a close watch on MESH, placing it under the care of Norman Shumway after Olch left.

Despite the problems of the inaugural year, MEDLARS performed well. Indexers, revisers, computer operators, search specialists, and others associated with MEDLARS gained the practical experience necessary to operate the system efficiently and to make full use of the capabilities of the system to store and retrieve information and to produce copy for printing bibliographies. Within a relatively short time they reached a high level of competence and produced work of whose quality they were proud.

The cost of development of MEDLARS, some \$3 million, was high but reasonable. At the time of the completion of MEDLARS there was no other

The first Photon ZIP Series 900 high speed computer phototypesetter, GRACE, installed in 1964, being operated by Donald Dodson. GRACE operated at a speed of 3600 words a minute, several times faster than previous phototypesetters.



EVOLUTION OF COMPUTERIZED BIBLIOGRAPHIES

publicly available fully operational electronic storage and retrieval system of its magnitude in existence. Not all of the objectives that Rogers and his staff had hoped for were attained, but on the whole the system was one of the largest and most successful library automation projects. It provided the medical profession of the United States, and later of other countries, with the most powerful bibliographic tool in the world. The original computer configuration served NLM well from the time it began to operate in January 1964 until its successor replaced it in January 1975. Its success marked a milestone in the evolution of modern libraries. And the development of the MEDLARS printer, GRACE, yielded a by-product, an advancement in the art of electronic typography.

RECURRING BIBLIOGRAPHIES

It was not intended that *Index Medicus* be the only bibliography produced by MEDLARS. Rogers, Adams, and Taine foresaw a stream of periodical bibliographies, ultimately perhaps as many as 50, issuing from the system. Once the citations to articles had been placed on the computer tapes, those in any field of medicine could be retrieved rapidly. Bibliographies covering the recent literature of any specialty could be produced without difficulty. There would be no necessity for any medical organization or agency to compile bibliographies for its members, it could be done by MEDLARS. Indeed, it would be a wasteful duplication of MEDLARS' mission to produce bibliographies by other means.

As the system approached completion Adams and Director-Elect Cummings discussed the policy that would be followed in producing these specialized lists of citations. They decided that the Library would not publish them itself, but only in cooperation with government agencies or nonprofit medical organizations. NLM would carry on the tradition of free service by supplying bibliographies without charge to the organizations, and the organizations would then have the bibliographies printed and distributed. The Board of Regents approved this policy in December 1963.

During 1964 under the direction of Taine and his successor Leonard Karel the Bibliographic Services Division formulated a number of experimental searches to obtain information on the construction of recurring bibliographies. By July 1964 the tests were completed. The first work printed from film composed by GRACE was *Cerebrovascular Bibliography*, under the sponsorship of the Joint Council Subcommittee on Cerebrovascular Disease of the National Institute of Neurological Diseases and Blindness and the National Heart Institute. The first recurring bibliography processed by MEDLARS from inception to retrieval was *Index to Rheumatology*, sponsored by the American Rheumatism Association with assistance from a grant provided by the National Institute of Arthritis and Metabolic Diseases. A pilot issue of the *Index* was printed in November 1964, and volume 1, number 1, was published on January 15, 1965.

Generally a recurring bibliography originated with a proposal from a health organization. If NLM concluded that the proposed bibliography would not duplicate any other publication and that it would help advance the subject, the

Library moved ahead. Representatives of NLM and the organization planned the format, frequency of publication, and specifications of contents. They defined the subject field using the terms in MESH, and then tested and developed the search strategy until the experts and searchers were satisfied. The Library produced a sample issue of the bibliography for review by experts in the field. Further modifications were made based on their criticism. Finally the Library produced a camera-ready copy for the organization to publish and distribute. Many months passed between the beginning of negotiations for a bibliography and publication of the first issue.

The Library and sponsoring organization benefitted mutually from the collaboration. The sponsor gained a current, thorough bibliography of primary importance to workers in its field while the library received expert advice on the literature needs of persons in the field and suggestions concerning terminology. On occasion NLM gained indexing assistance; the American Dental Association provided two indexers to the NLM staff to assist with the *Index to Dental Literature*, and the American Journal of Nursing Company added an indexer to help with the *International Nursing Index*.

Recurring bibliographies were quite successful. Within 6 months of the appearance of the first compilation, the Bibliographic Services Division was producing seven and was planning seven others. Organization after organization negotiated agreements with the Library, and each year two or three new titles appeared. By the end of a decade NLM was producing copy for 28 different recurring bibliographies.⁶

DEMAND SEARCHES

The second major type of bibliography produced by MEDLARS was not intended for publication but for the use of a person carrying out research. These lists of citations were retrieved from the system by a "demand search." While MEDLARS was nearing completion in 1963 trained search specialists began to make experimental demand searches to test the computer's ability to provide relevant bibliographies on specific subjects. As might be expected the operators found "bugs" in the system. They had difficulty preparing search formulations and procedures. This led them to evaluate items that influenced the search and the adequacy of terms used in formulations. They devised a new card input program to simplify search formulations, reduce processing time, and diminish the number of searches rejected for format error by the computer.⁷

Improving the search procedure took time. Finally in March 1964 the Library began to accept requests for searches. The early searches were made for physicians, researchers, health officials, and teachers who agreed to evaluate the bibliographies produced by MEDLARS for completeness and relevance so that the staff could improve the procedures. The service was limited at first because only a few operators were sufficiently familiar with the system to process searches, give demonstrations, and train associates.

The demand for MEDLARS bibliographies increased rapidly. By mid-1964

the operators had processed more than 600 searches. During fiscal year 1965 they successfully processed 1,623 out of 1,757 requests submitted to them. Most requests came from researchers and teachers. About 95 percent originated in the United States, the remainder in other countries.

During this period Director Cummings received a visit from Cyril Cleverdon, librarian of the College of Aeronautics, Cranfield, England. Cleverdon had evolved ideas for evaluating the efficiency and effectiveness of information systems by determining their recall and precision ratios. He explained his ideas to Cummings, who saw in them an opportunity for the Library to find out how satisfactory were the bibliographies produced by MEDLARS and to learn where improvements might be made.

Upon Cleverdon's recommendation the Director engaged F. Wilfrid Lancaster in December 1965 to evaluate MEDLARS. He retained Cleverdon as a consultant and appointed a committee of six knowledgeable computer specialists to review the test procedures and the analysis of results.

In making his test Lancaster sent forms requesting information to a selected group of persons who had asked for and received lists of citations retrieved from MEDLARS. Each person was asked his opinion of the usefulness, to him, of each of the articles on his list. From the replies returned by 302 persons Lancaster was able to tell, in each case, whether the system had retrieved informative articles, or relatively worthless articles, or some proportion of the two; and whether it cited sufficient articles or too few. He found failures attributable to lack of specificity in medical subject headings, variations in the exhaustivity of indexing, lack of specificity in indexing, failure of the requester to be specific in wording his request, and other causes. He recommended improvements in several areas, including the user-system interaction, index language, indexing, and search strategy. He suggested that NLM begin continuous quality control of MEDLARS searches.

Lancaster's study was useful in and out of the Library. Outside, the report was of interest to organizations with computers because of the diversity of subject areas that it covered and because it was the first large-scale evaluation of a major operating information system. In the Library Cummings ordered that Lancaster's recommendations be adopted, and he set up a quality control unit to check the effectiveness of every search requested by a patron. Later the designers of MEDLARS II benefited from this evaluation of MEDLARS.

Certain bibliographies produced by a demand search were, in the opinion of the staff, of interest to many persons and deserved a wide circulation. In June 1966 NLM began to reprint such bibliographies, christened "literature searches," announce them through various journals, and mail copies to patrons who requested them. The first literature search was titled "Anterior Pituitary Insufficiency due to Postpartum Necrosis, 1949-1965," and comprised 77 citations. By 1976 the Library had published 356 different searches and distributed between 30,000 and 40,000 copies.

DECENTRALIZATION OF MEDLARS

When MEDLARS was being designed the staff looked ahead to the time when the system would be completed, and they saw that search specialists would eventually be overwhelmed by requests for computer searches. The staff and the GE team considered several ways of preventing this. The Library could open branches in other sections of the country, each with its own computer facility and employees; it could open branches whose operators would send searches to, and receive bibliographies from, NLM via data communication equipment; it could make copies of MEDLARS tapes for other libraries which would then provide search service for users in their areas. The last option was selected because it was relatively inexpensive to duplicate tapes and because the establishment of search centers in other libraries would stimulate the growth of those libraries, relieving some of the pressure on NLM. The Library soon found that this decision appealed to the medical community; within a short time 35 institutions requested duplicate tapes in order to provide service to their clientele.⁸

The decentralization of MEDLARS began in 1964 when NLM awarded a contract to University of California in Los Angeles to serve as a search center. The UCLA and NLM computers were different, and one of the purposes of the agreement was to learn what difficulties libraries might face in reprogramming MEDLARS tapes. The reprogramming took longer than expected, and UCLA did not process tapes during the life of the contract.

The second search center was University of Colorado, awarded a contract in 1965. The University arranged with the Denver Federal Center for use of a computer identical to the NLM computer, and within a reasonable time it began to provide service to physicians in its area. The following year a committee of the Board of Regents considered applications from other institutions and recommended that contracts be awarded to University of Alabama, University of Michigan, and Harvard. Ohio State and Texas Medical Center obtained permission from the Library to establish MEDLARS centers using their own funds.

Further decentralization of MEDLARS in the United States came about through the Medical Library Assistance Act of 1965. This law authorized the granting of funds to medical libraries in various regions of the United States to enable them to provide services similar to, though on a smaller scale than, those provided by the National Library. Between 1967 and 1970 there were established 11 regional libraries, each of which became a MEDLARS center. Operators came from all MEDLARS centers to Bethesda to learn how the system operated and to formulate searches. After returning home the operators formulated search requests from local patrons and mailed the formulations to NLM for processing. Centers continued to send formulations to NLM until they became fully operational and were capable of processing searches themselves.

Institutions in Europe showed interest early in gaining access to MEDLARS so that they could provide service to physicians and libraries within their countries. Even before MEDLARS was perfected the Library received requests from abroad for services and data bases. In 1965 informal discussions between Director Cummings and physicians in Great Britain and Sweden led to an agreement: the Library would provide tapes to an institution in each of those countries and train operators for them if the institutions would evaluate the service and would index for NLM. Both countries complied and sent trainees to the Library. After these operators completed their training and returned home the British MEDLARS center began to operate in 1966 and the Swedish in 1967.

To guide the establishment of future MEDLARS centers abroad and to provide for cooperation between all foreign centers and NLM, Cummings laid down these policies. The Library would not select the institutions that would become centers, this would be done by governments or by organizations within the countries using standards suggested by NLM. The Library would provide access to MEDLARS and would train operators for the centers if in return the centers would index a reasonable number of articles monthly for MEDLARS. No funds would change hands, there would only be an exchange of service.

The success of the first European centers led the Office of Science and Technology and the State Department to suggest that NLM offer MEDLARS to the Organization for Economic Cooperation and Development. For 2 years Deputy Director Scott Adams endeavored to bring about a multilateral agreement with the organization leading to the setting up of a center to serve OECD countries, but the countries could not agree on a consortium. The Library finally decided to seek agreements with individual countries. Mary Corning, the Director's special assistant for international programs, negotiated agreements with institutions in several nations, among them Institut National de la Santé et de la Recherche Médicale in France, Deutsches Institut für Medizinische Dokumentation und Information in West Germany, The National Library of Australia, and Canada Institute for Scientific and Technical Information in 1970; Japan Information Center of Science and Technology, and World Health Organization in 1972, Iran in 1975; Mexico and South Africa in 1976; Italy in 1977; and Switzerland in 1980. The Library and its partners collaborated on policy and technical matters through an International MEDLARS Policy Advisory Group, which first met in 1972.

One disadvantage of the demand search service was its slowness. The time from the submission of the request, through the formulation of the request by an analyst at NLM and the processing in the computer, to the review and mailing of the bibliography to the patron, was usually 3 to 6 weeks. Therefore when MEDLINE, the Library's on-line retrieval service, became available in December 1971 patrons turned to it. They could obtain lists of citations within minutes. As the on-line system grew the centers sent fewer and fewer requests

A HISTORY OF THE NATIONAL LIBRARY OF MEDICINE

to NLM, and the Library discontinued the demand search service in January 1973.

Searches	FY 1965	'66	'67	'68	'69	'70	'71	'72
NLM	1,623	3,035	3,135	2,500	3,182	3,550	3,889	2,401
U.S. Centers			1,580	5,173	8,231	10,737	14,180	10,806
Foreign Centers			1,225	2,698	4,062	6,453	5,648	5,808
Totals			5,958	10,371	15,475	20,740	23,717	19,015

SELECTION OF JOURNALS FOR THE LIBRARY

The Library did not have sufficient funds to index every article in every medical and related journal in the world. Nor did it wish to cite articles that would have little value for most users. Its aim was to index as much of the world's substantive medical literature as possible while avoiding the indexing of periodicals of lesser value.⁹

The selection of journals for the Library's collection, and particularly the selection of journals whose articles were cited in *Index Medicus*, was very important. In the early days one person, Billings, had chosen the journals. After he built up the subscription list it was used year after year. New journals that came to the attention of Billings, Fletcher, or later editors of *Index-Catalogue*, were added for trial. Journals that went down hill were dropped. Wars killed journals and played havoc with the list. Business depressions forced journals out of business. The serials on the subscription list of 1960 were far different than those on the list of 1870, but they were still essentially the choice of a few persons within the institution. On the whole the editors had chosen wisely, and there was no criticism of their lists; still the volume of periodical literature kept increasing, the Library was broadening its scope, and Director Cummings felt it would be wise to seek advice from persons outside the Library.

In 1964, he appointed a Committee on Selection of Journals for *Index Medicus* which evolved into a less formally constituted group of consultants for the selection of literature for MEDLARS. These groups were composed primarily of non-NLM persons, including scientists, medical librarians, and physicians. The original committee, chaired by Leonard Karel, reviewed approximately 2,300 journals being indexed for, and others suggested for, *Index Medicus*, basing its decision on the quality of the journals in research, clinical applications, or education. Members found it easy to segregate the few hundred unquestionably superior periodicals out of the more than 15,000 titles acquired then (the number increased annually), but difficult to select 1,500 to 2,000 other journals of lesser importance, especially those in unfamiliar languages. Finally the committee, often with advice from specialists in certain subjects, recommended the addition of 466 titles to, and deletion of 324 from, the *Index Medicus* list. Each year thereafter the group's recommendations changed the character of *IM* slightly, with the aim of presenting to users an index to the most useful articles published throughout the world.

COMPUTER-AIDED CATALOGING

It had been hoped, during the early planning, that MEDLARS would be able to publish the Library's book catalog data in *Index Medicus*. The latter would then be a catalog of books, serials, and theses, as well as an index to articles. It would be reminiscent of the old *Index-Catalogue* in its universal coverage. In 1963 when MEDLARS was being programmed for the production of *Index Medicus*, General Electric and NLM made several attempts to satisfy the needs of the catalogers, but technical limitations in the programs prevented the publishing of cataloging data in acceptable form. Finally in October 1963 NLM decided to postpone the inclusion of book entries in *Index Medicus*.¹⁰

In February 1964 an analysis was begun of the procedures used by the Technical Services Division in selection, citation searching, acquisition, cataloging, and serial record keeping in order to obtain the information needed to design a computer system linking all of these functions. Several other libraries were consulted about their procedures, and librarians were asked for advice. While the analysis was going on NLM informed the medical library profession of its intention to publish a catalog with the help of MEDLARS and asked for preferences. On the basis of replies NLM decided to issue a biweekly serial that could be used by other libraries for cataloging and for selecting books for purchase.

The design of the system was completed in October 1965. Priority was given to the cataloging phase of the operation. Rather than write new programs the NLM staff modified several programs, adapting the MEDLARS input module to the specific requirement for displaying cataloging data. In January 1966 NLM began publication of the biweekly *National Library of Medicine Current Catalog*, one of the first regularly recurring, completely automated book catalogs in the world. The first issues listed by author and title the books and serials that had been acquired by the Library from December 1965 to January 15, 1966. Cumulations were issued quarterly and annually. After 1966 biweekly issues included only material published during the current and preceding 2 years. NLM also produced catalog cards, containing approximately the same information as the printed catalog. The Library estimated that American medical libraries could save a total of \$4 million a year if they adopted the *Current Catalog* for use in procurement and cataloging.

After a few years' experience NLM concluded that a monthly catalog would better serve the profession than a biweekly. It discontinued the biweekly issue at the end of 1969. Some libraries, however, still felt the need for biweekly lists. Beginning in July 1970 the Medical Library Association arranged to receive proof sheets twice a week from NLM, print copies, and distribute them to subscribers. Later the semiweekly gave way to weekly proof sheets.

The Library began the monthly catalog in January 1970. The monthly was short-lived owing to the development of an on-line bibliographic retrieval system named AIM-TWX which led to MEDLINE and to CATLINE. The latter

A HISTORY OF THE NATIONAL LIBRARY OF MEDICINE

possessed a base of catalog data that could be searched by libraries through computer terminals, starting in 1973. The monthly issue of the *Catalog* was then superfluous and it was discontinued at the end of 1973, leaving the quarterly and annual issues.

INDEXING BY CONTRACTORS

When indexers began to prepare citations for MEDLARS, they assigned many more subject headings than for the old *Index Medicus*. As a result they indexed fewer articles per hour, and a backlog of unindexed articles began to pile up. The Library hired additional employees, it asked indexers to work overtime, and it modified the procedure. These steps increased production, but still the backlog grew. Although the number of articles indexed rose from 144,057 in fiscal year 1964 to 168,310 in FY 1967, the backlog jumped from approximately 12,000 to about 70,000.¹¹

Indexers had to have an excellent knowledge of science, be very intelligent, be nearly fluent in foreign languages, and have good judgment. It was difficult for the Library to recruit persons with such unusual abilities because salaries were too low. NLM had set up a training program for indexers, but this did not fill vacancies.

The Library, therefore, considered the possibility of having some of the indexing done under contract. A pilot study in 1966 showed that the idea was practical. The following year NLM contracted with Keio University in Japan to index articles in Japanese, the editor of the *Israel Journal of Medical Sciences* to handle articles in Slavic languages, and the Parkinson Information Center at Columbia to index certain domestic journals. To make sure that the work was done properly the Library trained the contractors' employees, reviewed the indexes, and when necessary, revised them. During 1967 12,300 indexed citations, less than 8 percent of the total, were produced outside of NLM, but the following year commercial contractors at home and MEDLARS partners abroad began to contribute a large proportion. The backlog shrunk and disappeared, and soon all articles were being indexed currently. In 1969 the number of articles indexed by U. S. contractors and foreign MEDLARS centers outnumbered the articles indexed at NLM. By 1976 only 15 percent of all articles, or 38,400 out of 255,000, were indexed in the Library. Thereafter, on orders from the Director, 25 percent of all indexing was done at home.

In 1976 about 100 indexers and revisers, most of whom were employed by contractors or foreign partners, produced citations for the MEDLARS data base. This was more than 10 times the number of persons who indexed for the

Indexed articles	FY 1968	'69	'70	'72	'74	'76
U. S. contractors	59,000	76,200	75,500	120,700	74,000	100,000
MEDLARS centers abroad	22,000	34,900	43,600	72,800	106,200	116,600
NLM	112,000	99,400	90,900	39,900	44,100	38,400
Totals	193,000	210,600	210,000	233,400	224,300	255,000

old *Index-Catalogue* and *Index Medicus* from the 1880's to the 1920's. One monthly issue of *IM* in 1976 contained more citations (approximately 25,000) than the entire first volume in 1879 (approximately 18,000).

THE LIBRARY AS PUBLISHER OF *Cumulated* AND *Abridged Index Medicus*

In 1959 the Library and the American Medical Association had begun to publish the *Cumulated Index Medicus* as a joint venture. The Library, at the time, was using data processing equipment and an automatic camera to produce copy for printing the journal. This publication system was as primitive, compared with a computerized system, as the horse and buggy was to the automobile. When NLM ascended out of the mechanized stage into the computer stage, it had the potential to publish many more bibliographies each year.¹²

Among the ideas for new bibliographies was the thought that the Library could easily relieve the American Medical Association of the burden of completing each volume of the *Cumulated Index Medicus*. Director Cummings wondered if the AMA would not prefer to divest itself of its share of the work rather than expend part of its energies in completing, handling, and selling the volume. Finally in the summer of 1965 he met with Hugh H. Hussey, Jr., director of the division of scientific activities of the AMA and a Regent of NLM, and volunteered to assume responsibility for the entire production of *CIM*. Hussey agreed tentatively. Shortly thereafter Morris Fishbein, associated with the AMA publications for many years, visited the Library and completed the arrangements. Beginning with volume 6, 1965, NLM became sole publisher of the massive reference work. The four-volume set of 5,697 pages was processed in 120 hours over a period of 2 weeks utilizing the computer and GRACE.

The concept of an abridgement of the *Index Medicus* was another of the ideas that arose. Scott Adams recommended an abridgement containing references to articles on clinical medicine of interest to practicing physicians. The Library offered to publish an *Abridged Index Medicus* jointly with the AMA. In January 1965 Leonard Karel and his staff in the Bibliographic Services Division compiled a sample issue of the proposed journal containing 166 subject pages and 88 author pages, listing 3,660 citations from 216 journals. The association examined the sample, gave suggestions for improvement, and agreed to appoint a committee to draw up specifications for an abridged *IM*.

But by this time MEDLARS was operating, the Library was having difficulty maintaining its schedule, indexing was falling behind, operators were being trained, and the computer group had more than enough work turning out *Index Medicus*, *Cumulated Index Medicus*, and recurring bibliographies. The staff had to suspend work on the abridged *IM*.

In 1969 Clifford Bachrach, successor to Karel, and later editor of *Index Medicus*, returned to the abridgement. After much study to determine which high-quality periodicals would be most useful to practicing physicians, the staff and consultants selected 100 English-language journals from 2,300 journals being indexed in *Index Medicus*. Using MEDLARS they produced a pilot issue

in August 1969. The American Medical Association conducted a market survey and concluded that there would not be sufficient subscribers to warrant publication. But Bachrach was more optimistic and went ahead to prepare the first regular monthly issue of *Abridged Index Medicus* in January 1970. The AIM attracted sufficient subscribers to justify continued publication. Each year approximately 33,000 articles were listed in AIM, about 13 percent of the number in *Index Medicus*.

MEDLINE AND OTHER LINES

In 1967 Director Cummings engaged the nucleus of a new group whose mission was to study the ways by which modern methods of communication could be applied in the Library. The group shortly became the Lister Hill National Center for Biomedical Communications. Ralph Simmons of the center set up facilities for experimenting with computer on-line retrieval systems. This led to the development by a contractor, System Development Corporation, of a practical on-line bibliographic system named AIM-TWX, from *Abridged Index Medicus*, used as a data base, and Teletypewriter Exchange Network, the communication system. AIM-TWX was opened to a select group of users across the country in June 1970. During a trial period of several months users became enthusiastic over the speed with which it supplied bibliographic information.

Guided by experience gained during the test, Davis McCarn of the center and the contractor planned an on-line system that would accommodate 10 times as many searches as MEDLARS each year at one-tenth the cost. For the data base they selected citations to articles in 1,200 of the approximately 2,200 journals that were covered by *Index Medicus*. These citations amounted to about 60 percent of the total number, they included those most frequently sought, and provided a manageable base.

The operation of AIM-TWX had indicated that the cost of communication between the terminal and computer might be twice as much as the cost of the computer search, and that the communication cost would increase as the distance to the terminal increased. The staff were concerned that the expense might prevent many potential customers from using the new system. In order to make the data base as accessible to as many libraries as possible, NLM decided to subsidize the basic communication network. With assistance from the National Bureau of Standards, the staff sought the cheapest means of communication. For the first few months of operation they depended upon the Western Union Datacom System supplemented by phone lines leased from American Telephone and Telegraph Company. They then contracted with Tymshare for use of that firm's commercial communications network of high speed transmission lines connecting more than 50 cities in the United States and Europe. Libraries paid only the cost of telephone service to the nearest Tymshare city, from which they were connected to the contractor's computer in Santa Monica, California.

The new service named MEDLINE, from MEDLARS onLINE, began trial

runs in the Library on October 18, 1971, and was opened to a selected group of institutions in December. Among the original users were NLM's regional medical libraries and large medical school libraries. Later other schools, research institutions, hospitals, clinics, and independent medical libraries were admitted to the system. Libraries did not have to pay for use of the new bibliographic service, but in return for free access to MEDLINE they had to agree to provide service to persons who were not among their usual clientele.

As had been anticipated the new on-line system was used much more extensively than the MEDLARS demand search service. It provided bibliographies within a few minutes, in contrast to the demand search which took 3 or more weeks. The MEDLINE terminal showed the patron his citations, and thus gave him the opportunity to modify his request if he wished to obtain references more pertinent to his need. A year after MEDLINE began 150 institutions were connected to the system. Twenty-five libraries, on the average, were using it simultaneously, making 10,605 searches a month, or approximately 140,000 a year. About two-thirds of the patrons were satisfied with the bibliographies that appeared on their terminals; the remainder desired long bibliographies (more than 100 citations) which were printed off-line at NLM and mailed to them.

In February 1973 the Library stopped using the SDC computer in Santa Monica and provided MEDLINE service from Bethesda. The following month NLM arranged with the State University of New York at Syracuse, SUNY, to provide MEDLINE service through the Tymshare network. The SUNY computer could handle 40 searches simultaneously, nearly doubling the capacity of MEDLINE and assuring that the service would be maintained if the NLM computer shut down for some unforeseen reason.

Initially the MEDLINE data base was a selected bibliography designed for the majority of users, and it omitted a large proportion of the references in *Index Medicus*. To satisfy researchers who wanted every citation on a subject, NLM placed the *Index Medicus* references that had been left out of MEDLINE into a new data base called COMPFIL, from COMPLEMENT FILE. Into COMPFIL were also inserted citations from *Index to Dental Literature* and *International Nursing Index*. COMPFIL was made available to searchers in February 1973. By that time the on-line system was receiving so much use that the Library had to restrict access to COMPFIL to 2 days each week. COMPFIL was eliminated in 1974 when all the citations in *Index Medicus* were placed in MEDLINE.

To keep the MEDLINE file relatively current, containing only articles published within 2 or 3 years, NLM periodically removed older citations and placed them in BACK files, as BACK 66, holding the references from the period January 1966–December 1968. The back files were accessible for searching off-line.

The Library found that some users of MEDLINE were interested in only the most up-to-date articles in their field. Lister Hill Center therefore devel-

A HISTORY OF THE NATIONAL LIBRARY OF MEDICINE

oped a data base containing citations from the forthcoming monthly issue of *Index Medicus*. Each month users could receive citations weeks before the issue was printed and distributed by the Government Printing Office. This data base named SDILINE, from Selective Dissemination of Information onLINE, was made available in September 1972.

The on-line retrieval system was so successful that the Library staff incorporated all manner of useful data into computer files. By 1976 almost 3.5 million citations were in the data bases.

Name	Scope
AVLINE	Citations to audiovisual teaching packages used in health sciences education from 1974 on.
CANCERLIT	References and abstracts of articles on cancer from 1963 on. Originally called CANCERLINE.
CANCERPROJ	Descriptions of projects in cancer research for the present and preceding 2 years.
CATLINE	Reference to books and serials cataloged at NLM from 1965 on.
CHEMLINE	Chemical dictionary containing names and information on hundreds of thousands of substances.
EPILEPSYLINE	Citations and abstracts of articles on epilepsy in <i>Excerpta Medica</i> , from 1947 on.
MEDLEARN	A computer-assisted instruction program used in teaching persons to operate NLM's on-line system.
MEDLINE	Citations to articles and selected books from January 1978.
BACK 75	MEDLINE citations from January 1975 to December 1977.
BACK 72	MEDLINE citations from January 1972 to December 1974.
BACK 69	MEDLINE citations from January 1969 to December 1971.
BACK 66	MEDLINE citations from January 1966 to December 1968.
MESH VOCABULARY	Medical subject headings file, used for indexing, cataloging, and searching.
NAME AUTHORITY (MEDNAM)	Authority list of names and companies, used by NLM catalogers.
RTECS	Registry of toxic effects of chemical substances, an on-line quarterly compilation of data starting in 1976. The hard copy version was prepared by the National Institute for Occupational Safety and Health.
SDILINE	MEDLINE citations for the current month.
SERLINE	Bibliographic information on all serials ever cataloged by NLM.
TOXLINE	Citations and abstracts of published studies relating to toxicology, from 1977 on.
TOXBACK 74	TOXLINE citations from 1974 through 1976.
TOXBACK 65	TOXLINE citations from 1965 through 1973 plus citations from the private file of W. J. Hayes.

Patrons referred to the MEDLINE file far more frequently than any other file, communicating with it approximately 71,000 out of 77,000 connect hours the system was in use during 1976. CATLINE was the second most frequently

used file, but most of its usage occurred within NLM. TOXLINE was second only to MEDLINE in outside usage.

As more and more libraries hooked onto the bibliographic network, the Library purchased faster and larger computers. The IBM 360/50 used initially was replaced by a 370/155, this by a 370/158, and followed in 1977 by a twin 370/158 multiprocessor. These improvements permitted NLM to give the best service possible, to make additional files available, and to allow more libraries to search simultaneously. Paralleling improvements in the computer system was the upgrading of the retrieval program.

Extracting citations from the MEDLINE and other data bases was far different from obtaining them from the old *Index-Catalogue*, *Index Medicus* and *Current List*. The person sitting at the terminal and desiring citations from MEDLINE needed to know the procedure, which could be learned in a short time, and the strategy of locating citations indexed under the Library's medical subject headings list, which required, for proficiency, many hours of training and experience. From the opening of the MEDLINE system, NLM offered a course of 3 weeks duration for librarians. The course provided intensive training in the content of the data bases, the use of MESH, indexing practices, cataloging practices, and the use of the name authority file. The student spent approximately one-third of the time searching the bases gaining practical experience. The course for TOXLINE operators was shorter, generally of 3 days length.

In 1976 NLM developed a brief self-instructional course named MEDLEARN for beginning operators to study before coming to Bethesda for training. MEDLEARN required only half a day or less for completion, yet instructors found that it greatly increased the effectiveness of the main course. By 1977 the Library had trained approximately 750 operators to use MEDLINE, and almost 600 to use TOXLINE, CHEMLINE, and related files.

As in earlier days the Library had shared with institutions in other countries its printed *Index-Catalogue*, and in recent times its MEDLARS tapes, so now it invited its international partners to utilize the on-line system. By 1980 Canada and Mexico in North America, France and Italy in Europe, Iran in the Near East, and South Africa far away in the Southern Hemisphere linked themselves to the NLM computer in Bethesda through commercial communication lines. France extended the on-line service into Spain, Belgium, and Switzerland. Japan and Australia placed MEDLARS data bases on their own computers and provided on-line service to institutions in their own countries. Sweden, Germany, and United Kingdom also used their own computers and stretched the service into the Scandinavian countries, Poland, Netherlands, East Germany, Austria, and Belgium.

An important difference between cooperation earlier and in the 1970's was speed. In Billings' day at least 1 or 2 years elapsed between the time an article was indexed at the Library, the citation published in *Index-Catalogue* and the latter shipped to a European agent and distributed to users. If the citation had to wait in the file, because of the place of its subject heading in the alphabet,

15 or 20 years could elapse before it reached a user. Now the average length of time between the arrival of a journal and the appearance of the citation in the data base ready for users was only 80 days.

From the beginning of its existence the Library had given its services as freely as its resources would allow. It developed the MEDLARS demand search service and the MEDLINE system, and made them available to qualified users without charge. But the Library's funds were limited, and it could not continue to pay all the expenses indefinitely. After discussions in meetings of the Board of Regents about funding, Director Cummings ordered that users be billed for a portion of the communication costs. Beginning in August 1973, users paid \$6 per terminal connect hour, and 10¢ a page for off-line prints. The cost was raised to \$8 an hour in February 1975. On July 1, 1975, the Regents changed the rate structure to \$15 an hour for service between 10 a.m. to 5 p.m., and \$8 an hour from 5 to 8 p.m., and 3 to 10 a.m.

Physicians, nurses, researchers, other health workers, and students learned quickly of the great convenience, usefulness, and speed of NLM's on-line service after it became generally available in 1972. The number of searches increased from an estimated 22,000 in fiscal year 1972 to approximately 165,000 in FY73, 278,000 in FY74, 402,000 in FY75, and 579,000 in FY76. By the end of 1976 approximately 550 institutions were linked to the Library's computer. In 1978 NLM provided a million searches from all data bases, more than half of all the searches made in the United States in all fields of science and education. MEDLINE was the first large-scale successful on-line library-based bibliographic system, and the first international telecommunications-based science information network.¹³

MEDLARS II

Computer firms were continually improving their products. By the time MEDLARS went into operation it was evident that it would be obsolete within a few years. In the summer of 1966 Director Cummings contracted with Auerbach Corp. to draw up specifications for a new system, MEDLARS II, that would outperform MEDLARS, and in addition would accommodate elements of the cataloging process and the keeping of serial records, permit on-line retrieval of citations, include a drug information module, and store and retrieve graphic images. Cummings appointed a task force composed of NLM, NIH, and other government agency employees to assist in determining the Library's needs.¹⁴

In 1967 the Library requested proposals from industry. The Library's specification described what the new system should do. Firms were to state how the system would be developed, recommend a computer, and estimate the cost of developing the system. Seven firms delivered proposals.

On June 11, 1968, Cummings contracted with Computer Sciences Corp. to design, develop, and support the programming of MEDLARS II for \$2,037,505 (this did not include the cost of an IBM computer that NLM planned to pur-

chase). Development of the system was to take place in three phases and be completed by December 1971. The heart of the new system was to be a set of interrelated computer programs called COSMIS, computer systems for medical information services.

As the months passed by the contractor was unable to keep up to the schedule. Furthermore costs escalated. In the spring of 1969 the Board of Regents Subcommittee on Research and Development, whose function was to advise Cummings, met with the contractor's staff to try to locate the problem. Receiving the subcommittee's report, Cummings persuaded the contractor to change the MEDLARS team. Ralph Simmons, who had been overseeing the development of NLM's on-line retrieval system, spent weeks working out the provisions of a new contract which Cummings signed with CSC on June 20, 1969. The new contract stated explicitly the roles and interactions of the NLM and contractor teams, it gave new cost estimates, and contained penalties for cost overruns. Cummings set up a small unit to monitor the project and placed Simmons at its head, with authority to report directly to him.

Still the contractor fell behind the planned schedules. Simmons became pessimistic about the contractor's ability to meet deadlines. Regents Alfred R. Zipf, of the Bank of America, and Bruno W. Augenstein, of Rand Corp. received critical reports from members of their firms' computer staffs, whom they asked to check on the project. Cummings met with top management of the contractor to persuade them to hasten the work. But MEDLARS II proceeded slowly and on April 19, 1971, Cummings cancelled the contract.

In the meantime the System Development Corp. under contract had developed the Library's experimental on-line bibliographic retrieval system, AIM-TWX. A test of the system was to be started shortly. On June 9, 1971, Cummings contracted with System Development Corp. to complete MEDLARS II. The plan was to design the system around the structure and logic of AIM-TWX, adding an improved file generation and maintenance system and a new set of programs for photocomposition. The scope of some of the earlier objectives was reduced.

Over a span of 3½ years the contractor designed and developed the system. When completed in 1974 MEDLARS II contained a seven-level vocabulary (MESH) file, a journal file, current citation file, MEDLINE file, and other files. Davis McCarn and the computer staff tested the system during the latter months of 1974 and accepted it on behalf of the Library on January 3, 1975. The new system possessed all the capabilities of the original MEDLARS, it was faster, it was cheaper per unit cost of processing, it permitted higher standards for data, and it was more responsive to interactive searching and retrieval. New files with different record formats could be designed and implemented without interfering with other activities. The scope and variety of data bases and publications could be amplified more readily, and components of the system could be installed easily in other libraries.

In the summer of 1979 Director Cummings appointed Joseph Leiter to lead

a team whose task it would be to prepare specifications for a new computerized system, MEDLARS III, capable of managing and delivering a wider range of bibliographic information, data, and documents, faster and more rapidly than MEDLARS II. Eventually MEDLARS II would follow the original MEDLARS into oblivion, after years of service as the backbone of medical communication in the United States.

Notes

¹ In 1947 Scott Adams, then The Acting The Librarian, organized a meeting of librarians at the Pentagon to discuss the conflict in subject headings in the *Quarterly Cumulative Index Medicus*, *Current List*, and *Index-Catalogue*. During lunch, Director McNinch mentioned to the group that the Statistical Department of the Surgeon General's Office was using IBM machines, and he wondered if it was possible for such equipment to be used in medical indexing. He took the group to see the keypunch, card sort, and other machines in operation. McNinch's suggestion and the tour of the Statistical Division may have been one of the roots of the mechanization of *Current List*.

The conclusions of the Welch Medical Library Indexing Project, 1948 to 1953, also suggested that some of the operations could be done using business machines.

² The Council on Library Resources, Inc., was organized in September 1956 as a nonprofit body. The establishment of the council was made possible by a grant from the Ford Foundation of \$5 million to be expended over a 5-year period for "aiding in the solution of the problem of libraries generally, and of research libraries in particular." A copy of the Library's Proposal is in MS/C/295.

³ Information on the mechanization project may be found in: The National Library of Medicine Index Mechanization Project, July 1, 1958–June 30, 1959, *Bull Med Lib Ass* 49 Part 2 of number 1, Jan 1961, 1–96, records of the Board of Regents, Sept., 1957–Apr 1961, Outline of a Proposal Made by the National Library of Medicine to the Council on Library Resources, Inc., Feb., 1958 MS/C/47, annual reports of the Library, 1957–1961, Frank B Rogers, Tape-recorded comments on Index Committee May 24, 1979 NLM.

⁴ Correspondence between Rogers and officials of the AMA, a copy of the agreement between the PHS and AMA, and other pertinent documents are in MS/C/295. See also annual reports of the Library, 1961–1963.

⁵ A medlars is a fruit that resembles a crab apple and may be used in preserves. In earlier

times it was used in medicine, according to Thomas Cogan's *Haven of Health* (London, 1584) "Medlars are cold and drie in the seconde degree, they streine or binde the stomacke, and therefore they are good after meales, especially for such as bee over laxative, being much eaten they engender melancohe, and bee rather meat than medicine, as Galen saith. Yet of the stoness or kernelles of Medlars, may be made a very good medicine for the stone, as Matthio writeth. The stoness of medlars made in powder, driveth out the stone of the reynes, if you take a spoonfull thereof in white wine wherein the rootes of perseele have bene boyled."

Information on the development of MEDLARS may be found in many sources, among them: General Electric Co., Final technical report for MEDLARS preliminary design, Jan 31, 1962, Archival Coll NLM *The Principles of MEDLARS* (NLM, 1970) *The MEDLARS Story at the National Library of Medicine* (NLM, 1963) Charles J Austin, *MEDLARS, 1963–1967* (NLM, 1968) *NLM News* records of the Board of Regents articles, including F B Rogers "The National Library of Medicine's Role in Improving Medical Communications" *Amer J Med Electronics* 1 230–41 (July–Sept 1962), S I Taine, "The Medical Literature Analyses and Retrieval System," *Bull Med Lib Ass* 51 157–67 (1963), Winifred Sewell, "Medical Subject Headings in MEDLARS," *Bull Med Lib Ass* 52 164–70 (1964), S Adams and S I Taine, "Searching the Medical Literature," *JAMA* 188 251–4 (April 20, 1964), L Karel, C J Austin, M M Cummings, "Computerized Bibliographic Services for Biomedicine," *Science* 148 766–772 (1965) Tape recorded reminiscences of Frank B Rogers, May 24, 1979, and of Winifred Sewell, Mar 12, 1979, NLM. Information about the financing of MEDLARS by the National Heart Institute was obtained from James Watt.

⁶ A list of recurring bibliographies, with names of sponsoring organizations, may be found in each issue of *Index Medicus* and *Monthly Bibliography of Medical Reviews*. Information on the bibliographies may be found in *NLM News*, annual reports of the Library, and records of the

EVOLUTION OF COMPUTERIZED BIBLIOGRAPHIES

meetings of the Board of Regents Charlotte Kenton, "The Recurring Bibliographies Program of MEDLARS," *Bull Med Libr Ass* 54 135-7 (1966) S Adams, "MEDLARS and the Library's role as publisher," *National Library of Medicine Programs and Services*, F Y 1976, pp 5-10 Information was also obtained from Scott Adams and Clifford Bachrach

⁷ Information on the demand search service may be obtained from writings by the NLM staff, including F Wilfrid Lancaster, *Evaluation of the MEDLARS Demand Search Service* (1968), "Evaluating the Performance of a Large Computerized Information System," *JAMA* 207 114-120 (1969), and, with Grace Jenkins, "Quality Control Applied to the Operation of a Large Information System," *J Amer Soc Information Sci* 21 370-71 (1970) Data was also obtained from annual reports of the Library, and persons associated with MEDLARS Lists of subjects of some early demand searches are in *NLM News*, July through November 1964 Titles of literature searches are listed in the monthly issues of *Index Medicus*, 1967 onward Policy is in the records of the Board of Regents

⁸ Information on decentralization of MEDLARS may be found in records of the Board of Regents, *NLM News*, and annual reports of the Library

⁹ Statistics on journal acquisition and selection may be found in annual reports of the Library Leonard Karel, "Selection of Journals for Index Medicus," *Bull Med Lib Ass* 55 259-78 (1967), contains many references to published and unpublished sources Information was also obtained from Clifford Bachrach, editor of *Index Medicus* Policy is in the records of the Board of Regents

¹⁰ Statistics on cataloging may be found in annual reports of the Library Narrative information may be found in articles by NLM catalogers, *NLM News*, and Emilie Wiggins' unpublished manuscript, "Cataloging at the National Library of Medicine" Information was also pro-

vided by Carolyn Davis See also records of the Board of Regents

¹¹ Information on difficulties in indexing and on indexing under contract was obtained from Stanley Jablonski and Lloyd Wommack, both of whom were project officers on contracts Statistics were obtained from annual reports of the Library Policy statements are in the minutes of the Board of Regents

¹² Information on the *CIM* and *AIM* was obtained from Martin Cummings, Scott Adams, and Clifford Bachrach

¹³ For the development of AIM-TWX see the chapter on the Lister Hill Center Information on MEDLINE and other data bases may be found in articles by members of the staff, records of the Board of Regents, *NLM News*, on-line services reference manuals, and Library network/MEDLARS technical bulletins Statistics may be found in annual reports of the Library Information was also obtained from Grace McCarn, Scott Adams, Lloyd Wommack, Donald Hummel, George Cosmides

¹⁴ Information on MEDLARS II may be found in Auerbach Corp, Functional system specifications for the National Library of Medicine, July 1, 1967 (copy in archival collection), records of the Board of Regents *NLM News*, annual reports of the Library, and articles, including Robert V Katter, K M Pearson, Jr, "MEDLARS II, A Third Generation Bibliographic Production System" *J Libr Autom* 8 87-97 (1975), Davis B McCarn, J Leiter, On-line Services in Medicine and Beyond, *Science* 181 318-24 (1973), D B McCarn, 'National Library of Medicine-MEDLARS and MEDLINE,' in Belzer, Holzman, and Kent, *Encyclopedia of Computer Science and Technology*, v 11, pp 116-52 (N Y 1978) The on-line services reference manuals and Library network/MEDLARS technical bulletins contain much detail Information was also obtained from Martin Cummings, Ralph Simmons, and Joseph Leiter