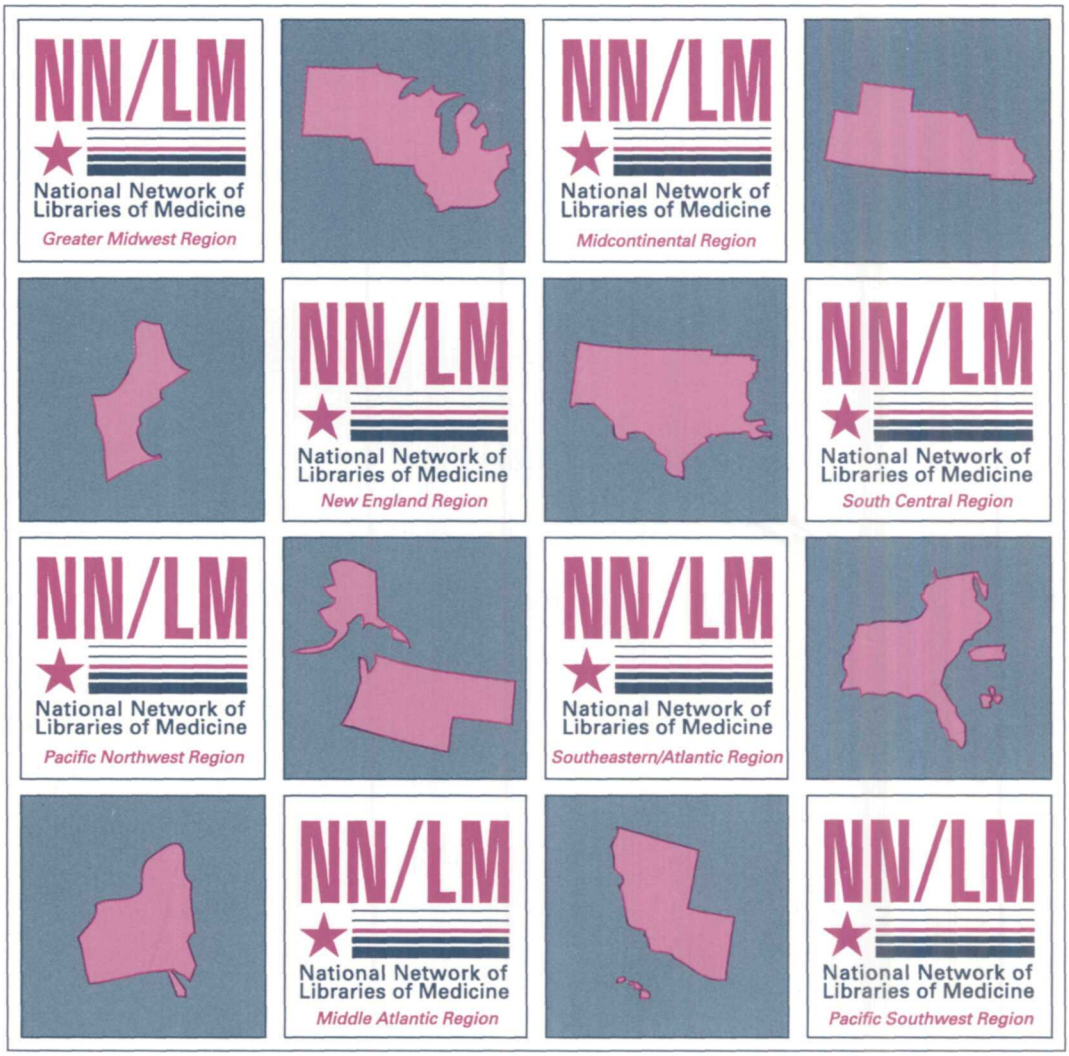
NATIONAL
INSTITUTES
OF HEALTH

NATIONAL LIBRARY OF MEDICINE

PROGRAMS & SERVICES
FISCAL YEAR 1992



U. S. DEPARTMENT OF HEALTH AND HUMAN SERVICES • Public Health Service

Further information about the programs described in this administrative report is available from:

Office of Public Information
National Library of Medicine
8600 Rockville Pike
Bethesda, MD 20894
(301) 496-6308

NATIONAL
INSTITUTES
OF HEALTH

NATIONAL LIBRARY OF MEDICINE

PROGRAMS & SERVICES
FISCAL YEAR 1992

U. S. DEPARTMENT OF HEALTH AND HUMAN SERVICES · Public Health Service

National Library of Medicine Catalog in Publication

Z
675.M4
U56an

National Library of Medicine (U.S.)
National Library of Medicine programs and services.--
1977- -- Bethesda, Md The Library, [1978-
v : ill., ports.
Report covers fiscal year
Continues. National Library of Medicine (U.S.). Programs and services Vols. for 1977-78 issued as DHEW publication , no (NIH) 78-256, etc.; for 1979-80 as NIH publication ; no. 80-256, etc
Vols for 1981-available from the National Technical Information Service, Springfield, Va
ISSN 0163-4569 = National Library of Medicine programs and services

1. Information Services - United States - periodicals 2. Libraries, Medical - United States - periodicals I. Title II. Series: DHEW publication , no. 80-256, etc

PREFACE

It is gratifying to report that the National Library of Medicine's collections, and the information tools that provide access to them, continue to grow and be widely used. More than 400,000 articles were indexed for MEDLINE in Fiscal Year 1992; the catalog reached 700,000 records; there are now almost 60,000 MEDLARS access codes.

We placed particular emphasis on providing assistance to the member libraries of the National Network of Libraries of Medicine in 1992. NLM:

Funded a new round of outreach grants aimed primarily at hospitals,

Instituted a new fellowship in applied informatics for librarians,

Revised the resource grant program to encourage hospitals of all sizes to apply,

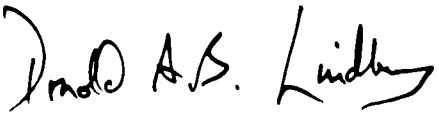
Began a new program to fund "connection" grants to Internet for which all public and private nonprofit health science organizations are eligible to compete, and

Began a three-year outreach program at the Marine Biological Laboratories (Woods Hole, Massachusetts) for which librarians and others are eligible to apply (see description under "Special Initiatives).

In June 1992, Grateful Med Version 6.0, with its access to additional databases, enhanced capabilities in networking (Internet, LAN, TOXNET), and built-in Loansome Doc connection, made its long-awaited appearance. Under a successful experimental arrangement between the American College of Physicians and the NLM, flat-rate access to MEDLARS is being provided to the College's members.

Many of the Library's R&D projects will feel the impact of the High Performance Computing and Communications (HPCC) initiative. NLM is a lead medical organization in this program, which is described in the "Special Initiatives" section of this report. The HPCC program has even more relevance to the Library, since the National Coordination Office for High Performance Computing and Communications was established at NLM on September 1, 1992.

Improving health communications is an inherently collaborative undertaking. On behalf of all the staff, I would like to thank our many friends in the National Network of Libraries of Medicine and the advisors around the country who unselfishly give of their time and talent to serve the NLM and further our mutual aims.



Donald A. B. Lindberg, M.D.
Director

CONTENTS

	Page
Preface	iii
Calendar of Events	vi
Special Initiatives	1
Outreach	1
Health Services Research Information	3
High Performance Computing and Communications	3
Library Operations	6
Planning and Management	6
Collection Development	6
Bibliographic Control	7
Network Services	8
Special Onsite Programs	11
NLM Associate Program	12
Specialized Information Services	18
Databases under ELHILL	18
AIDS	19
TOXNET and Its Files	20
Other Programs	21
Lister Hill Center	23
Computer Science Branch	23
Information Technology Branch	28
Communications Engineering Branch	29
Educational Technology Branch	33
Audiovisual Program Development Branch	35
National Center for Biotechnology Information	37
Database Building and Enhancement	37
Basic Research	39
Communication	40
Extramural Programs	40
Biotechnology Information in the Future	40
Extramural Programs	41
Overview	41
Training	42
Research Grants	43
Conferences	44
Bioethics	44
Resource Grants	44
Publication Grants	45
IAIMS	45
HPCC and IAIMS	46
Minority Support Activities	46
Committee Activities	46
Plans for FY 1993	47
Office of Computer and Communications Systems	48
Development Branch	48
Applications Services Branch	49
Systems Support Branch	50
Computer Services Branch	50

International Programs 52
Collaboration with Individual Countries 52
International MEDLARS Agreements 52
Collaboration with the W.H.O. 52
Special Foreign Currency Program 53
International Meetings and Visitors 53

Administration 55
Financial Resources 55
Personnel 55
Awards 55
Equal Employment Opportunity 56

Appendixes 58
1. Acronymns, Abbreviations, and Initialisms 58
2. Staff Bibliography 62
3. Extramural Programs-Supported Publications 67
4. Regional Medical Libraries 73
5. Board of Regents 74
6. Board of Scientific Counselors/LHC 75
7. Board of Scientific Counselors/NCBI 76
8. Biomedical Library Review Committee 77
9. Literature Selection Technical Review Committee 79

TABLES

Table 1. Growth of Collections 13
Table 2. Acquisitions Statistics 13
Table 3. Cataloging Statistics 14
Table 4. Bibliographic Statistics 14
Table 5. Circulation Statistics 14
Table 6. Online Searches 15
Table 7. Offline Searches 16
Table 8. Reference Services 16
Table 9. History of Medicine Activities 17
Table 10. Extramural Grants and Contracts Program 47
Table 11. International MEDLARS Centers 53
Table 12. Financial Resources and Allocations 55
Table 13. Staff 56

CALENDAR OF EVENTS—FISCAL YEAR 1992

1991
Oct 1 Visit and interview with Sister Anne Brooks
Oct 17 Literature Selection Technical Review Committee
Oct 25 Hearing on permanent paper use
Nov 7 LHC Board of Scientific Counselors
Nov 21 Biomedical Library Review Committee
Nov 26 NLM honor awards ceremony
Dec 6. Conference on the Regional Medical Programs initiative

1992
Jan 9 NCBI Board of Scientific Counselors
Jan 29 Rep Louis Stokes address honors Martin Luther King
Jan 30 Board of Regents
Feb-May History of Ophthalmology (exhibit)
Feb 5· Lectures by Dr Vanessa Gamble and Dr Todd Savitt in observance of African-American History Month
Feb 6 Literature Selection Technical Review Committee
Mar 4 Biomedical Library Review Committee
Mar 20 Address by Senator Nancy Kassebaum in observance of Women's History Month
Apr 1 Interlibrary loan charge increased to $8
Apr 1 NLM/American College of Physicians experimental flat rate agreement effective
May Office of Health Information Program Development established
May-Jul· Images of Hospital Pharmacy in America (exhibit)
May 28 Board of Regents
May 14 LHC Board of Scientific Counselors
May 28 Donna J Cox delivers 1992 Leiter Lecture
May 31 NLM-sponsored 1-week medical informatics fellowship program at Woods Hole
Jun Grateful Med Version 6 0 released
Jun 1. History of Medicine Division marks 50th anniversary
Jun 11 Literature Selection Technical Review Committee
Jun 24 Biomedical Library Review Committee
Aug-Dec Mind and Body Rene Descartes to William James (exhibit)
Aug Regents Naomi C Booker, Robert J Joynt, Carol M Newton appointed
Sep 1 NLM Director named to head HPCC National Coordination Office
Sep 24 700,000th record added to CATLINE, NLM's online catalog

SPECIAL INITIATIVES

Outreach

Outreach has been identified as NLM's highest priority in the past three annual reports, and it is highlighted again in this report. Outreach is more than an individual program: the mandate to improve health professionals' access to information is inherent in the Library's fundamental mission and cuts across NLM organizational lines. Outreach as a special initiative provides a focal point for a broad range of activities that are intended to ensure that health professionals are aware of and have access to the latest scientific findings in an easy-to-use form.

A new office was created within the Office of the Director during FY 1992. The Office of Health Information Programs Development, under the direction of Elliot R. Siegel, Ph.D., contains three units: the Office of Outreach Development, the Office of Planning and Analysis, and the Office of International Programs. The Office of Outreach Development will plan, develop, and evaluate NLM's outreach programs. It will serve as a catalyst, addressing new opportunities for outreach activities throughout the Library. To this end, an interdivisional NLM Outreach Coordinating Committee has also been established.

The NLM outreach program is a cooperative effort with the member institutions of the National Network of Libraries of Medicine (NN/LM). NLM has initiated more than 100 outreach projects since the publication of the DeBakey report in 1989[1]. They involve extensive efforts to train physicians and other health professionals to use Grateful Med, through special projects at the Regional Medical Libraries and awards to individual small-to-medium sized libraries in the network. There is an emphasis on libraries in rural and inner city areas.

With several years' experience in outreach behind us, we are now seeking to discover what can be learned from the completed projects, and what changes, if any, should be made in the future. A recent NLM staff evaluation of 30 Grateful Med outreach projects conducted between 1990 and 1992 will provide guidance for refining and improving this type of outreach activity. Other approaches to outreach have been evaluated, including the use of a circuit librarian in South Texas and the Meharry approach described below. Systematic evaluation will continue to be an important component of outreach, to identify and document those strategies that have been effective and those that have been found wanting, and to share these experiences with current and future collaborators who will benefit from this knowledge in the course of undertaking new outreach initiatives.

Some examples of specific outreach projects:

- 15 additional outreach projects were funded to introduce Grateful Med to a wide range of unaffiliated health professionals in rural and inner-city communities, including frontier counties in Colorado, inner-city Philadelphia, Hispanic inter-urban areas of El Paso, Texas, and 16 Northeast Louisiania parishes, among many others.
- Several outreach projects covering areas of Georgia, Maryland, New Jersey, North Carolina, North Dakota, Western Pennsylvania, Western New York, and Virginia were funded through the Regional Medical Libraries. A subcontract in Virginia will develop a network system called VAMIS, the Virginia Medical Information System. The North Carolina project will develop prototypes in information services at remote ambulatory care sites where residents, interns, and medical students are trained for family care medicine. Mercer University in Georgia will combine basic computer skills training with Grateful Med and online training for health professionals. The basic objectives of the other subcontracts are to: (1) identify and determine the number of unserved health professionals in the relevant areas, (2) determine their access to library services, (3) exhibit and demonstrate Grateful Med, and (4) provide training, support and library services.
- An earlier project completed in FY 1992 and conducted by Baptist Medical Center-Princeton in Birmingham, Alabama, provided Grateful Med training at seven rural hospital sites. The project trained 158 unaffiliated health professionals and 120 affiliated health professionals at one of the training/demonstration sessions conducted.
- NLM has targeted the Lower Mississippi Delta (a 214-county, seven-state region characterized as the most impoverished area of the nation) because of

[1] National Library of Medicine, Long range plan; Improving Health Professionals' Access to Information; report of the NLM Board of Regents. Bethesda, MD.: National Library of Medicine, August 1989.

its pressing health needs, and the expectation that provision of high quality health information can ultimately be demonstrated to have an observable impact on the quality of health care delivery. A comprehensive outreach program is planned for this region. Currently, a pilot project is under way now at the rural clinic of Dr. Anne Brooks, a practitioner serving a predominately minority population in Tutwiler, Mississippi. Dr. Brooks is providing NLM with valuable feedback on the usefulness of Grateful Med access for patient care in this setting.

- The Information Access Project at Meharry Medical College has trained more than 300 faculty, residents, students, and practicing health professionals to use Grateful Med and MEDLINE. Off campus, workstations have been established in the offices of eight health professionals (Meharry preceptors) practicing in rural and inner-city locations in Tennessee; these will be expanded to 23 by 1993. This project is giving us significant insights into impediments or barriers that prevent health professionals from accessing online information sources as part of their daily routine. More than merely identifying the barriers, the project seeks to find ways to turn the barriers into opportunities for change.
- An Undergraduate Research Study Program to stimulate undergraduate medical informatics research programs in Historically Black Colleges and Universities (HBCUs) has been established for electrical engineering and computer science students. Four electrical engineering students from three institutions—Morgan State University, Southern University at Baton Rouge, and North Carolina A & T State University—comprise the first class. These students will complete two school-year R&D assignments and two summer internships at the Lister Hill Center.
- A new initiative in the Toxicology Information Program is strengthening the capacity of nine HBCUs to train medical and other health professionals in the use of toxicological, environmental, occupational, and hazardous waste databases at NLM. Pilot strategies are being developed to utilize these HBCUs as vital information channels to reach minority and underserved populations who are disproportionately exposed to hazardous materials at home and in the workplace. Training workstations have been established at the participating HBCUs, including state-of-the-art hardware and software. Representatives from participating institutions have been trained in database searching. In cooperation with the Agency for Toxic Substances and Disease Registry, training and online access have been provided for 13 additional HBCUs with programs and concerns in the environmental and hazardous waste areas. A Toxicology Information Outreach Panel has been instituted to oversee this program.
- Grants under the Medical Library Assistance Act are available for health science librarians to help get the word out to their clientele. At hospitals, universities, and schools of the health professions, medical librarians are holding Grateful Med workshops and training sessions and they serve as a source of advice and consultation for the medical staffs and students. Their outreach has been especially valuable in small communities and rural areas, where the need is greatest.
- New publicity efforts have been designed to inform health professionals of the resources available to them, including exhibits at meetings, publications, and a logo and certificates for the NN/LM.

A new outreach program was begun in FY 1992 in collaboration with the Marine Biological Laboratory in Woods Hole, Massachusetts. This is a fellowship program, funded by the NLM, and aimed at medical educators, medical librarians, medical administrators, and young faculty who are not knowledgeable about medical informatics but who have the potential to become change agents in their institutions. The program consists of three annual one-week medical informatics training programs held at the Woods Hole facility in late spring. Thirty fellows are accepted to the program each year.

The first course was held during the week of May 31 - June 6, 1992. The course director was Homer Warner, M.D., Ph.D., professor and chair of the Department of Medical Informatics at the University of Utah School of Medicine. He was assisted by other instructors, among them Donald A. B. Lindberg, M.D., Daniel R. Masys, M.D., and David J. Lipman, M.D., all of the National Library of Medicine.

The objective of the course is to train individuals in the applications of computer and information science in medicine. The training consists of computer-assisted learning, retrieving and organizing information from computerized databases, the application of medical informatics tools to the critical appraisal of literature and associated statistical software packages, hospital- and office-based information systems, and electronic communications.

Students build and use a knowledge base for an expert system, and work with software for analysis of biological sequence data. Students also access the broad range of published scientific literature using NLM's Grateful Med. Lectures alternate with hands-on experience. The hardware and software set up for the one-week course is available during the rest of the year for demonstrations to and use by visiting groups.

E.R. Siegel

Health Services Research Information

The 1989 legislation that created the Agency for Health Care Policy and Research (AHCPR) assigned to NLM the responsibility to work with AHCPR to improve the organization and dissemination of health services research information. This mandate reflected congressional concern that information about the results of health services research, including practice guidelines and technology assessments, was not readily available in a format useful to health care practitioners, administrators, and third party payers. It also addressed the need for health services researchers to have better access to information that would facilitate their work. Using funds provided by AHCPR under interagency agreement, NLM established a new Office of Health Services Research Information, initiated improvements in the coverage of health services research in its existing services, and began preliminary work on a number of new information products and services. Some of these new initiatives will require additional resources for full implementation.

During this past year NLM developed a 5-year plan for its Health Services Research Information Program that responds to recommendations made by the Institute of Medicine in its 1991 report, "Improving Information Services for Health Services Researchers: A Report to the National Library of Medicine." The plan outlines the steps NLM should take to enhance the national information infrastructure that supports health services research and to expand outreach to the producers and users of health services research and to the librarians who serve both groups.

Related FY 1992 enhancements to NLM's current information services include expanded acquisition, cataloging, and indexing of health services research literature, improvements in MeSH coverage of the field, and publication of a guide to information resources of interest to health services researchers. There was also progress toward the development of new databases and services related to health services research and on extensions to the Unified Medical Language System (UMLS) Knowledge Sources that will be useful to health services researchers. Details about these activities and about NLM's support for AHCPR's practice guideline panels appear in the Library Operations and Lister Hill Center chapters of this report and in the UMLS subsection of the HPCC Special Initiative report.

B. L. Humphreys

High Performance Computing and Communications

The federal High Performance Computing and Communications (HPCC) Program is a research and development effort with the goals of creating computers a thousand times faster than today's supercomputers and a National Research and Education Network (NREN) that transmits information at speeds one hundred times faster than today's wide area computer networks. The HPCC program was begun in 1992 as a Presidential Initiative in science and technology, and received added impetus from the passage of a multiyear authorization in the High Performance Computing Act of 1991 (Public Law 102-192).

On the occasion of signing this legislation in December 1991, President Bush stated, "The development of high performance computing and communications technology offers the potential to transform radically the way in which all Americans will work, learn, and communicate in the future. It holds the promise of changing society as much as the other great inventions of the 20th century, including the telephone, air travel and radio and TV."

The focus of HPCC technology development is the solution of the "Grand Challenges" of science, a group of research problems that are in theory amenable to high performance computing but whose size and complexity currently make them computationally intractable for even our fastest supercomputers. Biomedical applications of computers, such as prediction of molecular structure from genetic sequence, analysis of human genome data, and structure-based drug design figure prominently as life science Grand Challenges.

There are four major components of the HPCC program. In three of the four, the NLM has significant related activities:

- High Performance Computing Systems (this is concerned essentially with hardware—NLM has no involvement)
- Advanced Software Technology and Algorithms (pertinent NLM programs—IAIMS, digital imaging, and biotechnology research programs);
- National Research and Education Network (pertinent NLM programs—UMLS, connections, and gateway projects); and
- Basic Research and Human Resources (pertinent NLM program—medical informatics training and fellowships).

The NLM was the first institution within the Department of Health and Human Services to create an HPCC program. The Library is now concentrating on the emerging need for applications. In 1992, NLM's HPCC program supported advanced technology and systems in the following several areas: enhancement of the Unified Medical Language System (UMLS), a prototype project of three-dimensional anatomic imaging called the Visible Human, and a grants program to link medical centers to the National Research and Educational Network. Each is briefly described below.

The UMLS. The UMLS strategy involves the development of three new machine-readable Knowledge Sources that can be used by a range of different interface programs

to interpret user inquiries, to locate information sources relevant to these inquiries, to formulate searches in the language that is appropriate to these sources, and to carry out the searches and organize the information retrieved in a helpful manner. The three UMLS Knowledge Sources are: a Metathesaurus of biomedical concepts and terms derived from many vocabularies and classifications, a Semantic Network that defines the "semantic types" or broad categories to which all Metathesaurus concepts are assigned and identifies sensible relationships between these categories, and an Information Sources Map that describes the content and location of biomedical information sources as well as how to connect to and actually search them. To encourage broad experimentation and feedback on potential enhancements, the Knowledge Sources are available free of charge under the terms of an experimental agreement.

In 1992, NLM issued the third experimental edition of the UMLS Knowledge Sources. The 1992 edition includes a number of additions in content and changes in format made in response to feedback received from those who experimented with the earlier editions. Among the improvements to the Metathesaurus are a major simplification of its logical structure and relational format, the inclusion of a word index, the incorporation of French language terms, and the addition of terminology from the Classification of Nursing Diagnoses; the Nursing Interventions Classification, AI/Rheum, NLM's rheumatology expert system; COSTART, the FDA's Thesaurus of adverse reaction terms; CRISP, the USPHS thesaurus for indexing scientific projects; UMDNS, ECRI's Universal Medical Device Nomenclature System; and DxPLAIN, Massachusetts General Hospital's expert diagnostic system. The Semantic Network has been enhanced by the revision of its hierarchical structure, the addition of 11 new allowable relationships, and the establishment of many additional non-hierarchical links between specific pairs of semantic types. The Information Sources Map has been expanded to include entries for additional environmental, toxicological, and molecular biology databases.

As of September 1992, more than 250 institutions and individuals had received copies of the UMLS Knowledge Sources and were applying them to a number of information problems in library and information service, patient care, medical education, and research. NLM itself is experimenting with the use of the Knowledge Sources in the Coach expert search assistant software, in an expert indexing assistant program, and in natural language processing. NLM's UMLS contractors are also developing a range of applications, including links between patient record systems and MEDLINE and a prototype of the Information Sources Map on a campus-wide network. In the summer of 1992, NLM awarded five small purchase orders to support the incorporation of the UMLS Knowledge Sources into existing applications in medical informatics training, curriculum analysis, indexing images, user friendly interface language, and clinical information systems.

The UMLS project depends on the technology being fostered by the HPCC program. Developing the UMLS Knowledge Sources, particularly the Metathesaurus and the Information Sources Map, requires high-speed telecommunications among the various sites involved. The 1992 version of the UMLS Metathesaurus could not have been built without rapid transfer of large data files and programs between NLM's Bethesda location and Alameda, California. As procedures for distributed maintenance of the Knowledge Sources evolve, the requirement for rapid transfer of large quantities of data and new versions of programs and for remote execution of validation routines will also increase. In addition, certain tasks associated with Metathesaurus construction, such as the processing of large information sources to derive statistics about the co-occurrence of concepts within them, are computationally intensive and can benefit from the use of high speed computers.

Even more critical than the use of HPCC technology in building the UMLS Knowledge Sources, however, is its enabling role in the achievement of the UMLS goal of integrating information from disparate machine-readable sources. Efficient access to biomedical information distributed among computer systems around the world (or across the state) will require reliable high speed and high capacity data transfer.

B. L. Humphreys

The Visible Human Project. The Long Range Planning effort convened by the Board of Regents in 1985 recognized the growing importance of electronic images as a source of biomedical knowledge. The report of the planning panel on Health Professions Education specifically recommended that "NLM investigate the feasibility of building digital image libraries."

In 1988 NLM convened a meeting of developers of computer systems capable of rendering 3-dimensional reconstruction of anatomic images; this group observed that NLM might facilitate progress in this area by building a digital image library representing the x-y-z numerical coordinates of an entire human body, a project too large to be undertaken in any single university center. The usefulness and feasibility of such a project, and the overall role that NLM might play in digital imaging for the life sciences was considered by an Electronic Imaging Planning Panel whose final report was reviewed by the Regents in 1989 and approved for incorporation into NLM's Long Range Plan.

Phase I of the "Visible Human Project" is now under way via a contract with the University of Colorado Health Sciences Center. Under the guidance of principal investigators, Dr. David G. Whitlock and Dr. Victor M. Spitzer, the project will acquire image data from photographic, computed tomography, and magnetic resonance imaging

of representative normal male and female cadavers. These data are expected to be available in 1994 as a set of digital images which NLM will make available over the National Research and Education network.

Subsequent phases of the project will address the intriguing and conceptually challenging task of identifying the objects of biological interest within those images, such as tissues, organs, and body systems, and linking the representation of those objects to the knowledge represented in textual databases such as MEDLINE and the NLM's varied factual databases. Current manual methods for identifying structures within anatomical cross sections require the expert skills of a trained anatomist, as well as extraordinary patience and attention to detail; the sheer volume of images in the Visible Human project will require the development of computer-assisted methods of object identification and classification. NLM is supporting research to develop such automated methods as part of its extramural grants program.

A far-reaching goal is to support the searching and retrieval of pictures using both language and visual indices, and the corresponding retrieval from textual sources by pointing to relevant structures within anatomic images (i.e., using words to find pictures and pictures to find words). Once retrieved, the 3-D anatomic coordinates will form the basis of "living images" rendered by computer graphics, which can be rotated, viewed from any angle, and reversibly "dissected." Instantly rendering these moving images with photorealistic quality is itself a computational Grand Challenge which is currently beyond the capacity of even supercomputers. The HPCC program will create new computer designs capable of meeting the challenge. The Visible Human project promises to provide an unprecedented national image resource for education of health professionals, and will foster development of computer graphics tools which will support computer-aided surgical planning and patient education.

Medical Connections Program. Among the thousands of research and education institutions which connect to federally supported computer networks, medical institutions are only sparsely represented. As part of its 1992 HPCC activity, NLM created a Medical Connections Program to provide "jump start" grants to medical centers wishing to connect to the National Research and Education Network (NREN). The program was conducted in collaboration with the National Science Foundation, as a supplement to a highly successful NSF Connections program for universities and colleges. In 1992 ten Medical Connections awards were made.

A variety of activities were supported by these awards. Examples include:

- statewide and multi-state outreach to community-based affiliated teaching hospitals to provide dial-up access to existing NREN computers at university centers;
- a regional set of network connections tying together a tertiary referral center with 45 physician practice sites;
- a metropolitan area network initiative to link five major teaching hospitals to a university NREN center;
- connections to multiple centers participating in a study of digital angiography for the diagnosis of coronary disease, for purposes of transferring digital image files;
- NREN access for a biomedical research institute to provide network-based searching of nucleic acid and protein sequence databases, and exchange of magnetic resonance images;
- extension of local computer center connectivity to the NREN to provide "desktop connectivity" for 120 medical students and 100 medical school faculty; and
- support for microwave regional links to an existing NREN site for distribution of computer-assisted instruction software.

The Medical Connections program is a step towards the creation of a nationwide and worldwide electronic biomedical community. It recognizes that we are in an era when the volume of medical knowledge far outstrips the ability of a single practitioner or even a single group of health professionals to remain current with the information relevant to all patient care and biomedical research problems. The program seeks to demonstrate in medicine a phenomenon which is well established in other scientific disciplines: that high-speed digital communications can pool and amplify the expertise of persons in widely separated locations.

Network connections support personal and group electronic mail, international newsgroups ("electronic bulletin boards") on special topics, access to biomedical databases, and provide a medium for exchange of large biomedical computer data files such as CT, MRI, and digital radiography images. Future enhancements to the NREN will support workstation video, teleconferencing, and other methods for distance-independent collaboration.

In an associated project, the University of Washington was awarded an enhancement to its Regional Medical Library contract to explore the efficacy of Internet connections and resources in community hospitals. There are three components to the project: extending Internet connectivity to seven community hospital libraries in the Pacific Northwest, developing network-based connection to anatomical images in support of clinical and educational needs, and providing DNA diagnostic laboratory directory information to clinicians in community practice.

The transformation of the desktop workstation from personal computer to "window-on-the-world" is a central theme of NLM's Medical Connections program, and a harbinger of a global biomedical community.

D. R. Masys

LIBRARY OPERATIONS

Lois Ann Colaianni
Associate Director

The Library Operations Division (LO) is responsible for carrying out NLM's basic services These include acquiring and preserving the world's biomedical literature, organizing this literature through indexing and cataloging, distributing NLM's authoritative bibliographic data online, in machine-readable products, and in publications, providing document delivery as a backup to the service available from other U S health sciences libraries, offering reference and research assistance to health professionals and researchers, and coordinating the National Network of Libraries of Medicine (NN/LM) LO also conducts research and evaluation related to these basic services and maintains an active research program in the history of medicine

The largest of NLM's Divisions, LO employs a staff of more than 260 librarians, library technicians, technical information specialists, subject matter experts, health professionals, and administrative support personnel LO has four main components Bibliographic Services, Public Services, Technical Services, and History of Medicine, two smaller units the Medical Subject Headings (MeSH) Section and the National Network of Libraries of Medicine Office, and a small administrative staff in the Office of the Associate Director

Planning and Management

In FY 1992, LO prepared a 5-year plan for NLM's Health Services Research Information Program (see Special Initiatives Section) LO activities in support of health services research are described throughout this chapter

During the year, the LO senior staff reviewed existing programs and products to identify potential areas for cost savings, with a goal of ensuring the continuation of essential basic services and of permitting progress on new initiatives The review was prompted by the expectation of increasingly restricted budgets throughout the federal government A number of program adjustments were made in FY 1992 or are planned for FY 1993 to respond to the changing resource picture Some of these changes are described in this report

On a positive note, a congressional recommendation to reduce the categories of costs recovered through fees for online services and for machine-readable products led to a major restructuring of NLM's MEDLARS pricing which will reduce NLM's charges to all categories of users of MEDLARS-related services

Collection Development

NLM's comprehensive collection of biomedical literature provides the essential foundation for many of its services The development of the collection requires establishing and revising literature selection policy, acquiring relevant biomedical literature in all formats and languages, receiving and processing the materials acquired, and maintaining and preserving the materials added to the NLM collection As of September 30, 1992, NLM holds 2,044,901 printed books, journal volumes, theses and pamphlets and 2,879,603 nonprint items, including audiovisuals, computer software, microforms, prints, photographs, and manuscripts (Table 1)

Selection

LO staff members select materials for the Library's collection according to guidelines published in the *Collection Development Manual of the National Library of Medicine* Although specific sections of the Manual are revised and updated as needed, NLM also undertakes a complete review and revision of its selection policy and guidelines every 5 to 8 years This systematic review involves staff members throughout the Library as well as outside subject matter and collection development experts In FY 1992, LO completed a major revision of the Manual to be published in FY 1993 following final review by the NLM Board of Regents and approval by the NLM Director The goals of this revision were to reflect ongoing developments in biomedicine and the biomedical literature, to focus the Library's collecting effort more sharply on biomedical subjects thereby reducing overlap with other national libraries, to clarify selection guidelines in problem areas, and to improve the Manual's format

Acquisitions

NLM acquired 44,390 volumes and 59,083 other items (e g , audiovisuals, microforms, software, pictures, manuscripts) for its collection in FY 1992 (Table 1) The LO staff received and processed 183,635 modern books, serial issues, audiovisuals, and software packages (Table 2) The Library's historical collection was enriched by the addition of a number of significant works Noteworthy examples include *Cirugia* (Venice 1491), the third edition of surgical treatise originally entitled *La Ciroxia Vulgarmente Eata* written by Gulielmus de Saliceto about 1275 and considered the most important work on surgery of the 13th century, *Regimen Sanitatis Pariense* (Leipzig, 1499) a Latin and German edition of advice (in verse form) on exercise, sleep, diet, and personal hygiene, the first edition of John Locke's *An Essay Concerning Humane Understanding* (London, 1690) which laid the foundation of modern psychology, the first edition of Florence Nightingale's *Notes on Matters Affecting the Health, Efficiency, and Hospital Administration of the British*

Army, Founded Chiefly on the Experience of the Late War (London, 1858), one of her most significant contributions and an important document on military medicine and hospital administration; and the personal papers of C. Everett Koop, M.D. from the period of his tenure as Surgeon General of the Public Health Service.

In FY 1992, NLM negotiated with the Library of Congress (LC) to ensure more rapid transfer to NLM of copyright deposit copies of a number of biomedical journals. In addition, an increasing number of publishers are agreeing to provide NLM with free or low-cost second copies of journal titles indexed in *Index Medicus*. Both of these arrangements allowed NLM to cancel some second copy subscriptions. The Library continues to purchase one full-price subscription for each commercially published *Index Medicus* title. Continued improvements to the systems providing automated support for serials processing and monograph acquisitions both reduced data creation time and improved the quality of information on literature acquisition and processing.

Collection Preservation and Maintenance

Binding, microfilming brittle volumes, conserving rare and unique items, promoting the use of permanent paper in new biomedical publications, and exploring new methods for preserving library materials are the major components of NLM's program for preserving and maintaining the biomedical literature.

In FY 1992, NLM microfilmed 1.2 million brittle pages and preserved 401 items from the special collections, including the last segment of a collection of 175 films produced by the former National Medical Audiovisual Center which were transferred to 1" masters. Substantial staff resources were expended on working with the new microfilming contractors to bring them up to appropriate production and quality levels. NLM's Preservation Section completed a thorough analysis of NLM's collection of commercially produced microfilms of U.S. *Index Medicus* titles and found that 92 percent were complete and of acceptable technical quality. Arrangements are being made to replace or re-film the remaining 8 percent. NLM's preservation copies of historical films were moved to offsite storage in better environmental conditions. To improve access to and control of NLM's microfilm service copies, the service copy collection was consolidated into a single location within the Library. To increase onsite shelf space for older serial volumes, NLM purchased additional compact shelving. When the new shelving is installed in early FY 1993, the Library will be able to accommodate projected growth of its serials collection through 1998.

Following several rounds of paper testing and the circulation of three preliminary drafts, the revision of the American National Standard NISO/ANSI Z39.48 (Permanence of Paper) was completed in August 1992. NLM staff directed the National Information Standards Organization Committee that prepared the revision. The revised standard is expected to be approved by the American National Standards Institute in early FY 1993. The revision includes provisions for coated, as well as uncoated, paper.

Bibliographic Control

NLM's authoritative cataloging and indexing records provide the basis for effective access to biomedical literature. NLM's Medical Subject Headings (MeSH) is the vocabulary used to describe and to search the subject content of items cataloged and indexed by NLM and by other health-related institutions around the world.

Thesaurus

MeSH is a hierarchical thesaurus that contains 17,007 subject headings and 48 publication types. MeSH's supplementary chemical file contains about 68,000 records for substances. In FY 1992, 415 new MeSH headings and 1,593 new entry terms were added to the vocabulary to keep pace with developments in biomedicine and changes in the usage of biomedical terminology. The additions included new terms for specific phases of clinical trials and new and updated terminology related to AIDS, molecular biology, endogenous receptors, nematodes, and countries in Eastern Europe. Substantial revisions were made to the MeSH hierarchies for health services research, dermatology, some types of chemicals and drugs, enzymes, neoplasms, and geographic areas. The revisions to the health services research hierarchies were based on recommendations from a joint NLM/Agency for Health Care Policy and Research Task Force. New publication types were added for meta-analyses and specific phases of clinical trials to facilitate retrieval of articles that are particularly useful in the development of practice guidelines.

In addition to enhancing NLM's thesaurus, members of the MeSH Section staff are also heavily involved in editing additions to the UMLS Metathesaurus and in modifying the structure of MeSH to facilitate its incorporation into the Metathesaurus.

Cataloging

LO catalogs works acquired for the NLM collection both to describe what is available here and to provide authoritative cataloging records for use by other health sciences libraries throughout the country and the world. To achieve these objectives LO also creates and maintains the Library's automated files of cataloging and name authority records, contributes NLM's cataloging data to national cooperative bibliographic databases, and maintains the NLM classification scheme. In FY 1992, the Library cataloged 21,221 modern books, serials, nonprint items, and Cataloging-in-Publication galleys, using a combination of

inhouse staff, contracts, an interagency agreement with the Library of Congress, and assistance from the MEDLARS Center in China. The cataloging production level was the highest in a decade and an increase of 11 percent from FY 1991. Because of the large number of receipts, the inventory of modern works received but not cataloged still increased by 3,051 items during the year.

During September 1992, the nineteenth anniversary of the CATLINE (Catalog Online) file, the 700,000th record was added to the database. NLM continues to make progress on increasing CATLINE's coverage of its special historical collections. In FY 1992, the History of Medicine Division began to add records for NLM's early Arabic and Persian manuscripts to the file, based on the descriptions of these works prepared by Dr. Emilie Savage-Smith, an authority on the history of Arabic medicine. More than 6,000 abbreviated records for items in NLM's picture collection were upgraded to fully cataloged status in FY 1992. A test version of a special integrated retrieval system for these records and the videodisc images of NLM's historical picture collection was completed and is currently being enhanced.

During FY 1992, NLM's online cataloging system was upgraded to provide immediate online name and subject validation to catalogers. Improvements in remote access to the system benefitted offsite cataloging contractors and staff who work from home. The Cataloging Section made substantial progress on revising and updating the NLM Classification.

Indexing

NLM indexes articles from selected journals to ensure rapid access to recent developments in biomedicine. The Index Section also provides special indexing of the gene sequences published in articles for incorporation into the National Center for Biotechnology Information's backbone database (see NCBI chapter). If previously published articles are retracted, corrected, or challenged in subsequently published commentaries, the indexed citations to these articles are updated and annotated accordingly.

The Literature Selection Technical Review Committee (LSTRC) (Appendix 9) advises NLM on the journals that should be indexed in MEDLINE, *Index Medicus*, and other NLM databases. In FY 1992, the LSTRC reviewed 362 journals and rated 51 sufficiently high for NLM to begin indexing them. After considering recommendations prepared by professional societies in the fields of rheumatology and neurology, the LSTRC advised NLM to begin indexing 6 titles and to cease indexing 13 titles in these areas. During FY 1992, the Library began to index AHCPR-sponsored practice guidelines. These are in addition to the many practice guidelines that are published in journals already indexed for MEDLINE.

MEDLARS indexing is done by NLM staff, commercial contractors, some international MEDLARS centers, and cooperating organizations such as the American Hospital Association, the American Journal of Nursing Co., and the American Dental Association.

In FY 1992, NLM added 401,000 indexed citations to MEDLINE, a 10 percent increase from the previous year and a new record for citations added in a single year. Seventy-six percent of the citations added included English-language abstracts. Previously indexed citations in the MEDLARS databases were updated to reflect information about 27 retractions, 2,787 published error notices, and 20,894 substantive commentaries. In a special project, the new publication type "Practice Guideline" was added to a number of retrospective MEDLINE and HEALTH citations identified by AHCPR.

The Library continued its phased acquisition of new PC workstations for inhouse indexers. The new workstations support more effective online access to MeSH, the Indexing Manual, and other indexing tools. The Kennedy Institute for Bioethics at Georgetown University began use of a new PC-based indexing system that accommodates creation of non-journal citations for BIOETHICSLINE. A modification of this system, which was developed by NLM's Office of Computer and Communications Systems, is being tested for maintenance of existing citations in MEDLINE and in other MEDLARS files.

Network Services

NLM's network services promote easy access to its comprehensive collection and to its authoritative bibliographic data. These services include: distribution of indexing and cataloging data in publications, machine-readable formats, and an online retrieval service; reference assistance to onsite and remote users; document delivery to onsite users and to remote requesters who need items from the NLM collection to supplement the resources of other U.S. libraries; and overall direction of the National Network of Libraries of Medicine (NN/LM). The major thrust of NLM's outreach initiative is to link more health professionals to these services.

Publications

In FY 1992, NLM produced more than 100 individual issues of some 24 recurring indexes, catalogs, and bibliographies. For the first time this year, NLM generated photocomposed copy of the *Bibliography of Bioethics* for the Kennedy Institute of Bioethics. As part of NLM's expanded health services research emphasis, the Office of Health Services Research Information and the Reference Section compiled and issued *Health Services: Sources of Information for Research*. The Reference Section also produced a number of

Current Bibliographies in Medicine, often in conjunction with NIH Consensus Development Conferences. The subjects covered included adolescent alcoholism, methods for voluntary weight loss, and silicone implants.

Despite the wide availability of NLM's data in machine-readable formats and online services, a number of NLM's publications continue to be "best sellers." Because the prices charged by the U.S Government Printing Office and the National Technical Information Service do not recover NLM's production costs, LO reviewed some publications with declining numbers of subscribers to identify any that could be distributed via alternative means or eliminated without significant negative impact on users. As a result of this review, NLM will cease publishing the *National Library of Medicine Current Catalog* and the *National Library of Medicine Audiovisuals Catalog* after the release of their 1993 cumulations. NLM's cataloging data are readily available and more easily searched in a variety of online bibliographic databases. By the end of FY 1993, they will also be available on the Internet through NLM's new Online Public Access Catalog interface. NLM has also decided to reduce the number of *Current Bibliographies in Medicine* produced each year, again due to declining subscriptions In FY 1993, NLM will experiment with distribution of the *AIDS Bibliography* via the Department of Health and Human Services electronic AIDS Bulletin Board in addition to its printed format.

Machine-readable Databases

As part of its effort to provide the widest possible access to its authoritative data, NLM leases its databases in machine-readable form to commercial database vendors, international MEDLARS centers, universities, and other organizations that make the data available online or in CD-ROM products. In FY 1992, NLM distributed more than 9,000 tapes of various databases to more than 100 different licensees. At the end of the year, 12 licensees were producing CD-ROM products containing data from 5 different MEDLARS databases. During the year, tape licensees were asked to complete a Quality Assurance Self-Evaluation Form to verify that they were storing, updating, and retrieving MEDLARS data accurately Early in FY 1992, NLM implemented a revised pricing structure for tape licensees and data distributed on CD-ROM. In response to a Congressional request, a revised pricing structure has been developed for 1993 that eliminates all use charges, i.e., tape licensees will pay to lease the data, but there will be no additional charges based on how much the data they lease are actually used. A foreign surcharge will remain in effect.

Online Services

NLM offers direct online access to 46 databases. In FY 1992, online users performed 5 6 million searches of these databases during 305,000 online connect hours (tables 6 and 7). These figures exclude online searching performed on the computer systems of organizations that lease data from NLM.

Special outreach efforts by NLM and health science libraries throughout the NN/LM have resulted in the number of users of NLM's online system growing rapidly. At the end of FY 1992, there were 58,335 active codes, an increase of 24 percent from FY 1991 Most of the new codes were assigned to individuals rather than to institutions. Essentially all of the individuals who received codes in FY 1992 indicated an intention to use Grateful Med, NLM's user-friendly microcomputer search software, to access the NLM databases. Since Grateful Med first appeared in 1986, the National Technical Information Service has distributed 50,003 copies (42,530 IBM PC version; 7,473 Macintosh) Purchasers receive new versions of the software at no extra charge

In FY 1992, NLM issued version 6 0 of the IBM PC Grateful Med. New features include the ability to access NLM's databases via the Internet, an expanded MeSH display that includes a hierarchical arrangement of terms, assistance in selecting publication types and journal title abbreviations to use in searches, tailored search screens for BIOETHICSLINE and the TOXNET databases, and the ability to search current MEDLINE and retrospective MEDLINE data in a single search transaction.

A test HSTAR (Health Services and Technology Assessment Research) database was developed from relevant citations in MEDLINE, HEALTH PLANNING & ADMINISTRATION, and additional citations indexed expressly to expand NLM's coverage of health services research literature Initial inhouse testing has been completed. A broader test involving health services researchers, users of the results of health services research, and librarians who serve these groups is planned for FY 1993. LO staff also provided technical advice to the Foundation for Health Services Research on the development of a prototype Health Services Research-in-Progress database that may be mounted on the MEDLARS system in the future

NLM entered into an experimental flat-rate pricing arrangement with the American College of Physicians to gather information that will help in establishing a general flat-rate policy. To provide better support to its growing population of online users, the Library is using contractor assistance to extend the online service desk hours to 12 midnight Eastern time on weekdays and 8.30 a.m. to 5:00 p.m. on Saturdays. In addition to responding to questions posted on the Grateful Med Bulletin Board, the MEDLARS Management Section also established two Internet addresses to which users may send questions regarding Grateful Med and other MEDLARS services.

A total of 1,843 librarians, other search intermediaries, and individual users received online search training at NLM or from staff at the three RML online training centers in Regions 1, 4, and 7. In FY 1992, a new specialized training module was developed for AIDS-related databases. NLM staff also taught a special course in searching biotechnology information at the 1992 annual meeting of the Medical Library Association.

Reference Services

NLM provides reference service and assistance to onsite users and to remote requesters as a backup to the service available from U.S. health sciences libraries. In FY 1992, NLM's Reference Section responded to 72,548 requests for reference assistance: 68 percent from onsite requestors, 31 percent in telephone calls, and 1 percent in letters or telefacimile transmissions (table 8). The Library now monitors several Internet discussion lists and responds to reference questions posted there when appropriate. A large number of additional inquiries regarding hours of service, directions to the library, etc., are handled by an automatic telephone answering system.

In FY 1992, staff in the Office of Health Services Research Information and the Reference Section continued to provide special literature search support to a number of panels established by the AHCPR to develop clinical practice guidelines for specific conditions. NLM staff members also analyzed the bibliographic references actually used as evidence in the first three AHCPR-sponsored guidelines as background for developing recommendations for streamlining the literature search process.

Following an analysis of patterns of onsite use, NLM eliminated Monday evening and Saturday afternoon service hours. There was little use of NLM's onsite services by health professionals or biomedical researchers during these hours. The Library's regular hours of operation are now 8:30 a.m. to 5:00 p.m. on Monday and Friday, 8:30 a.m. to 9:00 p.m. Tuesday through Thursday, and 8:30 a.m. to 12:30 p.m. on Saturday. NLM continues to be closed in the evenings between Memorial Day and Labor Day.

Fourteen different indexes, catalogs, and full-text sources are now available on NLM's CD-ROM network. A new Journal Information System was installed in test workstations in the main Reading Room in FY 1992. The system, developed by the Lister Hill Center from SERLINE data, assists onsite users in obtaining information needed to locate journals in the Reading Room or to request journal issues from NLM's closed stacks. When more fully deployed throughout the Reading Room, it will replace the multi-part listings now used by onsite patrons to identify journal call numbers.

Substantial progress was also made on a joint LO/Office of Computer and Communications Systems project to develop an Online Public Access Catalog Interface (OPAC) to NLM's CATLINE and AVLINE databases.

Document Delivery

NLM provides document delivery service to remote requesters to supplement the service available from other libraries in the NN/LM and to onsite users who wish to use items from NLM's closed stacks (table 5). NLM's Collection Access Section received a total of 528,288 interlibrary loan and onsite requests for post-1913 documents in FY 1992. The History of Medicine Division filled 17,434 requests for items from the historical collections, including orders for copies of pictures. NLM received 302,271 modern interlibrary loan requests and filled 73 percent of them. If requests for which the requester was unwilling to pay are excluded, NLM's fill rate was 78 percent. Sixty-two percent of all filled requests were processed within a single day of receipt. Eighty-seven percent of the interlibrary loan requests received by NLM were submitted via DOCLINE, NLM's automated document request and routing system. Four percent or 11,750 requests were received via telefacimile transfer; of these, 3,418 were needed for clinical emergencies and were processed within two hours of receipt.

The test of the System for Automated Interlibrary Loan (SAIL) continued throughout FY 1992. Scanned bit-mapped page images of recent issues of 65 journals newly selected for indexing in MEDLINE and *Index Medicus* from 1988 to 1992 were used to fill requests referred to NLM via DOCLINE. Requested articles are printed out for mailing or converted to faxable form and sent without human intervention. During FY 1992, SAIL filled 10,444 document requests or about 5 percent of the interlibrary loan requests filled by NLM.

The number of libraries using DOCLINE increased to 2,385 in FY 1992. Among the new users are BIREME in Sao Paulo, Brazil and the Korean MEDLARS Center in Seoul. Participating libraries entered almost 2.4 million document requests into DOCLINE and filled 85 percent of them. NLM filled an additional 8 percent for an overall fill rate of 93 percent. During FY 1992, DOCLINE service hours were extended to 10:00 p.m. Eastern Time on weekdays and to Saturdays from 7:00 a.m. to 5:00 p.m. Recent enhancements made to the system include: access via the Internet; notification to the requesting library when documents are sent directly to a health professional; an increase in the number of libraries that can be included in a routing table; the ability to control the routing of requests by designating start and stop routing cells; and new statistical reports.

Individual Grateful Med users may employ the Loansome Doc feature to send requests for documents identified in MEDLINE to a network library which has agreed to handle their requests. The requests are routed by DOCLINE if that library is unable to fill them. In FY 1992, Grateful Med users initiated Loansome Doc requests for 51,253 documents.

DOCLINE routes requests to appropriate libraries based on approximately 1.26 million SERHOLD records, which represent the holdings of 3,093 NN/LM member

libraries. The procedures and programs for updating the SERHOLD were improved in FY 1992. SERHOLD can now accept data submitted in the MARC holdings format. Validation routines have been enhanced to detect duplicates and nonstandard holdings data. When journal titles change, holdings data from the old title are now copied to the new title, pending the next opportunity for participating libraries to update their holdings statements.

A study of use of NLM's document delivery service by other federal libraries revealed that these libraries account for an increasing percentage of interlibrary loan requests directed to NLM, including many requests for titles that are commonly held throughout the NN/LM. Federal libraries have not in the past been charged for document delivery service from NLM, although NLM incurs substantial contract costs, as well as inhouse staff costs, to respond to requests from these libraries. After reviewing the study data and discussing their implications, the NLM Board of Regents reaffirmed that it is the Library's policy to charge for interlibrary loan service and recommended that NLM staff work with segments of the federal library community to reduce inappropriate use of NLM's collection and to phase in charges if necessary. Work is under way to increase federal library participation in DOCLINE and use of other NN/LM resources. High volume federal library requesters will be charged for interlibrary loan service in 1993.

In FY 1992, onsite users requested 226,017 documents from NLM's closed stacks. Sixty-three percent of the requests were for regular day-time stack service; 29 percent were for the fee-based overnight photocopy service; and 7 percent for special Information Broker Stack Service. Eighty percent of the requests were filled; 95 percent of the requests filled during the day-time were delivered to users within 30 minutes.

National Network of Libraries of Medicine Program

The goal of the NN/LM is to make up-to-date biomedical information readily accessible to U.S. health professionals and researchers, irrespective of their geographic locations. There are 3,600 Network members including health sciences libraries of every size and type located in all parts of the country. LO's NN/LM Office oversees and coordinates Network programs which are administered by the eight Regional Medical Libraries (RMLs). Table 4 lists the regional libraries. The NN/LM Office and the RMLs communicate regularly via E-mail and teleconference to ensure new services are introduced smoothly throughout the country and to identify needed enhancements to existing services, such as DOCLINE and Loansome Doc.

The NN/LM program is a critical component of NLM's outreach initiative. The RMLs, as well as many individual network members, carry out specific outreach projects to underserved rural and inner-city health professionals. (See Special Initiatives Section). In FY 1992, participants in the NLM Associate Program conducted projects to evaluate the impact of the first 30 individual library Grateful Med outreach projects. The findings identified both effective and ineffective outreach strategies and should be useful to those undertaking new outreach projects. The RMLs now perform most of the exhibiting and demonstration of NLM's products and services at health professional meetings around the country. NLM's MEDLARS Management staff exhibit at the meetings in the Washington, D.C. area. In FY 1992, a new NN/LM logo, brochure, and other Network "identity" materials were developed for use by the RMLs and other network members in publicizing NLM and network services.

Another related focus of the NN/LM program is technology transfer. In FY 1992, the Region 6 RML began a special HPCC-related project to link health care institutions in the northwest to the Internet (See Special Initiatives Section). With assistance from the Friends of the National Library of Medicine, some of the RMLs are also organizing conferences on technology-related topics. On June 24, 1992, the Louise Darling Biomedical Library at UCLA held a Technology Awareness/Transfer Conference entitled "The Electronic Health Sciences Journal: A New Information Mechanism."

Special Onsite Programs

In addition to the reference and document delivery services provided to onsite users, NLM offers a variety of special programs and services to those who visit the Library in Bethesda, including guided tours, briefings on NLM's operations and services, and historical exhibits and symposia. NLM also has a visiting Historical Scholar Program and a one-year post-master's training program for librarians with potential for substantial contributions to health sciences information services.

Public Tours and Briefing

NLM is a popular attraction for domestic and international visitors with an interest in any and all facets of biomedical communication, medical librarianship, and information technology. In FY 1992, LO staff members conducted 144 regular daily tours for a total of 434 visitors. The Office of Public Information (Office of the Director) arranged 129 special tours and orientation programs for groups (1,686 visitors). NLM staff members also arranged special briefings on library programs and services for many individual visitors. A list of the countries from which these visitors came is in the International Programs chapter.

Historical Programs

In FY 1992, NLM organized a special conference on the history and effects of the Regional Medical Programs (RMPs). The conference brought together many of those who planned and directed these innovative attempts to improve health services during the 1960s. In conjunction with

this program, the Library published an extensive bibliography with a legislative history. LO is continuing its effort to expand NLM's collection of manuscript materials related to RMI's and is collaborating with the Office of the Director and the Lister Hill Center to scan and index RMP documents. Stephen Strickland, Ph.D, currently a visiting scholar at NLM, is writing a history of the RMI's.

In observance of African-American History Month, the History of Medicine Division and the NLM EEO Office organized a program on "Medical Institutions for Blacks: An Historical Perspective." Dr. Todd Savitt, an NLM visiting historical scholar, spoke on "The Rise and Decline of Medical School for Blacks, 1865-1920" and Dr. Vanessa Gamble of the University of Wisconsin, Madison spoke on "The Black Hospital: A Vanishing Institution." The program was opened with remarks from NIH Director Dr. Bernadine Healy and a taped message from the Secretary for Health and Human Services, Dr. Louis Sullivan.

Other historical conferences and lectures held during FY 1992 included: a symposium cosponsored by the Louise Darling Biomedical Library, UCLA, on "Medical Aspects of the European-American Encounter" in conjunction with the Columbus Quincentennary and a lecture by Dr. Roger O Egeberg on "A Doctor Goes to War: New Guinea and MacArthur in World War II"

The Library prepared several historical exhibits including "Highlights in the History of Ophthalmology," "Images of American Hospital Pharmacy," in conjunction with the 50th anniversary of the American Society of Hospital Pharmacists, and "Mind and Body: Rene Descartes to William James", in conjunction with the centennial of the American Psychological Association.

Members of the NLM's History of Medicine Division continued their research using NLM's collections. Staff research results appeared in many publications and were presented at professional meetings and invited lectures throughout the year.

NLM Associate Program

The NLM Associate Program is a one-year competitive program that allows library school graduates to become familiar with NLM's operations, to gain an understanding of key issues facing health sciences libraries, to use new information technologies, and to develop their skills by conducting special projects. In addition to the outreach evaluation projects mentioned previously, projects carried out by Associates in FY 1992 included a study of structured abstracts in MEDLINE, development of a manual and alpha-testing procedures for the Coach expert search assistant program, and analyses of semantic relationships between MeSH headings in MEDLINE citations as possible input to expanding the UMLS Semantic Network. Associates also have an opportunity to visit the other national libraries and various types of health sciences libraries and information centers and to attend professional meetings

Four Associates completed the 1991/1992 program, two accepted jobs in the Lister Hill Center; one went to work at the National Academy of Sciences; and one is working on outreach projects in Kentucky Five new Associates began the program in September, including the first international Associate who is from China.

Table 1
Growth of Collections

Collection	*Previous Total (9/30/91)*	*FY 1992*	*New Total*
Book Materials			
Monographs:			
Before 1500	571	4	575
1501-1600	5,750	10	5,760
1601-1700	10,061	9	10,070
1701-1800	24,363	20	24,383
1801-1870	39,926	25	39,951
Americana	2,341	0	2,341
1870-Present	551,827	14,149	565,976
Theses (historical)	281,794	0	281,794
Pamphlets	172,021	0	172,021
Bound serial volumes	947,006	30,173	977,179
Volumes withdrawn	(35,149)	0	(35,149)
Total volumes	2,000,511	44,390	2,044,901
Nonbook Materials			
Microforms:			
Reels of microfilm	(56,333)	(420)	(56,753)
Number of microfiche	(253,312)	(12,568)	(265,880)
Total microforms	309,645	12,988	322,633
Audiovisuals	52,117	1,883	54,000
Computer software	541	37	578
Pictures	56,600	0	56,600
Manuscripts	2,401,617	44,175	2,445,792
Total nonbook	2,820,520	59,083	2,879,603
Total book and nonbook	4,821,031	103,473	4,924,504

Table 2
Acquisition Statistics

Acquisitions	*FY 1990*	*FY 1991*	*FY 1992*
Serial titles received	21,557	21,181	21,863
Publications processed:			
Serial pieces	144,356	158,939	157,882
Other	21,068	23,344	25,753
Total	165,424	182,283	183,635
Obligations for:			
Publications	$3,632,746	$3,943,338	$4,358,439
Included for rare books	($203,559)	($184,742)	($193,193)

Table 3
Cataloging Statistics

Item	*FY 1990*	*FY 1991*	*FY 1992*
Completed Cataloging			
Full	12,060	12,707	15,278
Limited	7,309	6,480	5,943
Total	19,369	19,187	21,221

Table 4
Bibliographic Services

Services	*FY 1990*	*FY 1991*	*FY 1992*
Citations published in MEDLINE	391,172	363,344	401,562
For *Index Medicus*	363,890	341,874	380,485
Recurring bibliographies	28	23	23
Journals indexed for *Index Medicus*	2,973	3,020	3,048
Abstracts entered	275,000	281,644	295,803

Table 5
Circulation Statistics

Activity	*FY 1990*	*FY 1991*	*FY 1992*
Requests Received:	456,904	494,515	528,288
Interlibrary Loan	258,421	281,606	302,271
Readers	198,483	212,909	226,017
Requests Filled:	349,999	385,405	401,565
Interlibrary Loan	183,950	207,670	219,831
Photocopy	170,605	193,855	207,685
Original	12,054	12,606	10,726
Audiovisual	1,291	1,209	1,420
Readers	166,049	177,735	181,734
Requests Unfilled:	106,905	109,090	126,723
Interlibrary Loan	74,471	73,936	82,440
Referred	3,431	2,050	3,249
Returned	71,040	71,886	79,191
Reader Service			
Returned as unavailable	32,434	35,154	44,283

Table 6
Online Searches

DATABASES	*FY 1990*	*FY 1991*	*FY 1992*
AIDSDRUGS	247	310	389
AIDSLINE	24,525	36,904	38,165
AIDSTRIALS	768	646	925
AVLINE	12,879	15,760	15,435
BIOETHICS	8,505	11,221	13,536
BIOTECHSEEK	—	—	738
CANCERLIT	63,898	79,511	79,562
CATLINE	158,293	213,376	211,834
CCRIS	2,983	4,862	4,038
CHEMID	3,497	7,939	9,961
CHEMLINE	22,683	26,878	20,928
CLINPROT	2,432	1,117	—
DART	1,244	4,632	4,168
DBIR	1,942	2,241	1,302
DENTALPROJ	262	279	205
DIRLINE	7,120	9,482	10,072
DOCUSER	3,309	9,475	10,982
EMICBACK	1,517	2,107	3,794
ETICBACK	1,627	1,776	1,264
GENETOX	—	919	1,778
HEALTH	136,616	175,285	172,124
HISTLINE	4,643	5,918	5,697
HSDB	34,939	42,479	36,934
INFORM	127	306	1,202
IRIS	4,959	12,133	20,710
LOAN STATUS	211	737	3,821
MEDLINE	2,058,301	2,731,557	2,788,562
MED86	602,394	700,010	580,815
MED83	402,341	402,914	327,954
MED80	225,340	247,919	209,780
MED77	132,976	151,620	134,417
MED72	91,601	106,882	97,449
MED66	68,373	84,586	76,500
MESH VOCABULARY FILE	20,448	27,079	35,936
NAME AUTHORITY FILE	3,179	3,507	3,644
PDQ	69,684	44,194	25,748
POPLINE	19,140	17,260	18,696
REFLINE	41,902	51,393	47,347
RTECS	16,303	17,684	16,291
SDILINE	41,185	55,077	52,666
SERLINE	55,038	75,309	81,232
STORED SEARCH	107	154	127
TOXLINE	68,911	80,191	68,634
TOXLINE65	9,962	10,411	7,138
TOXLIT	15,516	17,653	14,047
TOXLIT65	5,116	5,634	3,867
TRI	30,625	35,674	30,207
TRIFACTS	—	—	392
YEAR86	655	18	4
Total	4,478,323	5,533,019	5,291,017

Table 7
Offline Searches

DATABASES	*FY 1990*	*FY 1991*	*FY 1992*
AIDSLINE	1,294	1,807	1,998
AVLINE	103	129	157
BIOETHICS	25	30	23
CANCERLIT	3,654	3,586	3,408
CATLINE	536	555	565
CHEMLINE	1	1	2
DIRLINE	4	0	0
HEALTH	10,983	10,654	10,983
HISTLINE	9	2	7
MEDLINE	4,630	5,364	4,419
MED86	5,993	4,751	3,726
MED83	5,211	3,415	2,318
MED80	3,557	2,513	1,667
MED77	2,073	1,401	876
MED72	1,452	993	616
MED66	838	610	368
MESH VOCABULARY FILE	1	1	0
POPLINE	5,107	3,778	4,627
SDILINE	229,625	226,397	227,103
SERLINE	10	8	4
TOXLINE	6,608	5,421	5,610
TOXLINE65	76	24	2
TOXLIT	5,497	4,441	4,191
TOXLIT65	103	22	3
Total	287,390	275,903	272,673

Table 8
Reference Services

Activity	*FY 1990*	*FY 1991*	*FY 1992*
Reference Section:			
Requests by telephone	19,222	19,889	22,437
Requests by mail	585	487	600
In-person requests	40,823	46,140	49,511
Total	60,630	66,516	72,548

Table 9
History of Medicine Activities

Activity	*FY 1990*	*FY 1991*	*FY 1992*
Acquisitions			
Books	360	66	152
Modern manuscripts	128,088	73,388	538,125
Prints and photographs	642	0	0
Processing			
Books cataloged	232	330	355
Modern manuscripts processed	112,541	129,000	44,175
Pictures cataloged	0	0	0
Citations indexed	5,136	5,888	5,777
Pages microfilmed	66,581	88,524	39,654
Public Services			
Reference questions answered	13,982	12,184	9,528
ILL and pay orders filled	3,506	3,477	3,369
Reader requests filled	9,358	5,992	6,447
Pictures supplied	5,872	4,683	7,618

SPECIALIZED INFORMATION SERVICES

Michael J. Ackerman, Ph.D.
Acting Associate Director

In June of 1992, as the eyes of visitors to Rio de Janeiro were drawn to the dramatic outcropping of that city's Sugarloaf Mountain, so the attention of the world was focused on Rio's Earth Summit. Officially designated the United Nations Conference on Environment and Development, this summit welcomed heads of state of 178 nations to discuss plans to protect the global environment within the context of sustainable development. Among the many goals addressed were several dealing with hazardous chemicals, pollution, and human health, subjects with which NLM's Specialized Information Services Division (SIS), through its Toxicology Information Program, has long been intimately acquainted. Indeed, SIS was already six years old in 1972 when the first UN Conference on the Environment was held in Stockholm.

The Earth Summit's look at the future of the environment was paralleled by an examination, undertaken by a Long Range Planning Panel on Toxicology and Environmental Health, on the future of SIS. The panel's report outlined the types of core toxicology and environmental health information resources and services that should be strengthened within NLM itself, went on to urge NLM to facilitate access to related resources outside the Library and worldwide, and recommended the development of an information infrastructure for future scientific discovery.

Operationally, 1992 was a boom year for the Division. TOXNET (Toxicology Data Network) continued to grow and evolve with the addition of two new files: TRI90 and TRIFACTS. TRI90 is the 1990 edition and fourth year of the popular Toxic Chemical Release Inventory (TRI) series of files. This series presents data, collected annually by the Environmental Protection Agency, on the estimated releases of toxic chemicals to the environment, and their transfer to waste sites. TRIFACTS (Toxic Chemical Release Inventory Facts) is a companion file to the TRI Series and contains health effects and ecological data, presented in nontechnical language, on the TRI chemicals.

Of all TOXNET's files, IRIS (Integrated Risk Information System), has experienced the most rapid increase in usage, attesting to the strong interest in chemical risk assessment. HSDB (Hazardous Substances Data Bank), with its broad scope and peer-reviewed data, continues to be TOXNET's flagship and most heavily used file.

TOXNET's conversion from minicomputers to a microprocessor-based network has proven highly successful, allowing for more efficient searching at a reduction in cost to NLM. In 1992, the system received another boost in processing efficiency as the thirteen 386 microprocessor machines were replaced with 486 computers.

Important strides have also been made by SIS in the area of outreach through the establishment of a Toxicology Information Outreach Panel. This panel is helping to advise SIS on programs to strengthen the capacity of Historically Black Colleges and Universities (HBCU) to train health professionals in the use of toxicological and environmental health information resources developed at NLM.

SIS continues to look outward in space and forward in time for ways to improve the collection, organization, and dissemination of information in toxicology and environmental health. At its inception in 1966, SIS responded to society's growing concern about hazardous chemicals. It was on the forefront of toxicology information in 1972, the year of the Stockholm Conference and has continued to maintain that role through the 70s, 80s and the 1992 Rio Earth Summit. Increasingly, SIS is not merely responding to emerging information needs in toxicology, but anticipating them, in part by developing state-of-the-art computer systems, such as TOXNET, to better manage the flow of information. SIS expects to continue its leadership role and to develop new mechanisms to handle rapidly changing information needs in toxicology and environmental health. The 1992 Earth Summit may have passed into memory but Sugarloaf Mountain and TOXNET have a long way to go.

Databases under ELHILL

ChemID (Chemical Identification File) is an online chemical dictionary that covers primarily chemicals of biomedical and regulatory importance. Available at normal MEDLARS rates without the royalty charges required by CHEMLINE, ChemID allows users to search by a variety of chemical and biological identifiers; to formulate strategies for searching other MEDLARS files; and to locate other files on the ELHILL and TOXNET systems, or external data, which contain more information about the chemical in question.

ChemID contains an important set of regulatory data, collectively known as SUPERLIST. Over 8,600 records are augmented with the name and an indication of source for chemicals mentioned in one or more of 19 lists of regulatory or biomedical importance. Coverage includes such lists as the Department of Transportation Hazardous Materials List, the Hazardous Substances Reportable Quantities (RQ) List, and the ATSDR Priority List. These data allow users to determine if a certain chemical is mentioned on a given list and under what name; to search for chemical classes on these lists; and to show coverage overlap between lists. In 1992, several additional lists were prepared for inclusion in SUPERLIST; they are the GRAS (Generally Recognized as Safe) list from FDA, and HAP (Hazardous Air Pollutants) list from EPA. ChemID contains about 275,000 records.

CHEMLINE (Chemical Dictionary Online) is an online chemical dictionary and directory file which allows users to identify chemical substances via nomenclature and other identifiers, and to formulate optimum search strategies for other NLM files. Each chemical record has pointers to other files on the ELHILL and TOXNET systems which contain information about that chemical substance. CHEMLINE is updated every two months and regenerated annually. Most of CHEMLINE's data are supplied by the Chemical Abstracts Service (CAS) from its Registry System, but this is augmented by NLM; users must pay CAS royalty fees to use CHEMLINE. Originally made available in January 1974 with about 59,000 records, the file now contains over 1,100,000 records of chemical substances.

During FY 1992, the scope of coverage of the CHEMLINE file continued to increase. CHEMLINE now covers substances in the AIDSTRIAL file on ELHILL, as well as the TRI series of files on TOXNET. For seekers of regulatory information, the data from the Environmental Protection Agency's Toxic Substances Control Act (TSCA) Inventory was updated, as well as that for the European Inventory of Existing Commercial Chemical Substances (EINECS). Data from all sources contributing to CHEMLINE were updated and enhanced during the file regeneration of 1991.

In 1992, a new Component Registry Number (RC) field was added to CHEMLINE. This field uses data provided by CAS for over 20,000 records of chemical mixtures and provides the CAS Registry Numbers of the records which define the parent chemical components in that mixture. A user may now find all mixtures containing the pesticide DDT by qualifying the RN of that substance with an RC qualifier, as in "(RC) 50-29-3."

For the first time, CAS provided locator values in CHEMLINE records using data sent by NLM. Previously this task was done at NLM after receipt of data from CAS. The new method eases the workload on NLM staff.

TOXLINE (Toxicology Information Online) is an online bibliographic retrieval service, produced by merging "toxicology" subsets from some seventeen secondary sources. TOXLINE and its backfile, TOXLINE65, contain data from sources that do not require royalty charges based on usage.

Information from Chemical Abstracts Service, which requires usage royalties, is used for two other online bibliographic files, TOXLIT and TOXLIT65. The four databases in the TOXLINE family of services now contain nearly 3,500,000 records.

During FY 1992, the TOXLINE files were regenerated to add current MeSH indexing vocabulary to the subset derived from MEDLINE and to the Biological Abstracts subfile records. A new subfile, DART (Developmental and Reproductive Toxicology), was added to TOXLINE in May. This subfile, also available as a separate file in TOXNET, continues the Environmental Teratology Information Center subfile, providing an expanded scope of coverage in this area. New initiatives for TOXLINE include the development of a subfile on toxicology research projects from an annual survey done by the National Toxicology Program.

DIRLINE (Directory of Information Resources Online) is an online directory of information resources including organizations, databases, electronic bulletin boards, as well as programs and projects with special biomedical subject expertise. These resources provide information and assistance which may not be available from NLM's bibliographic databases.

In the past DIRLINE has been compiled from data collected by other Federally funded organizations. However, NLM has initiated a maintenance program for the general biomedical component of the DIRLINE database. Much of this data was previously provided by the Library of Congress' National Referral Center Database which is now defunct. The maintenance program gives NLM more control over this major component of the database, as well as an opportunity to expand into areas which previously were not adequately covered.

AIDS

NLM, one of four Public Health Service agencies cooperating to provide the AIDS Clinical Trials Information Service, continued to improve its AIDS information resources including those mandated by the Health Omnibus Programs Extension Act.

The AIDSTRIALS and AIDSDRUGS databases, both products of this PHS cooperative project, have continued to grow. AIDSTRIALS, which includes trials actively adding new patients, those which have completed their accrual, as well as trials which have been completed, is an effective mechanism for health professionals to identify suitable trials to which they may refer patients. It is also a useful tool for clinical researchers to review trial design and locate colleagues working in areas of interest. The Division has started working with the Canadian HIV Trials Network and other international groups to add non-U.S. AIDS-related clinical trials to the AIDSTRIALS database.

AIDSDRUGS contains information about the agents being tested in the clinical trials included in AIDSTRIALS. As AIDSTRIALS has expanded, more agents have been described in AIDSDRUGS. The information about each agent includes pharmacology, interactions, adverse effects and chemical/physical properties. Bibliographic citations are included for those who wish to locate more detailed information.

NLM is the lead agency for planning an NIH AIDS/HIV Information Services Conference to be held in 1993. This conference will provide NIH with guidance in the development of information products and services that will be of most use to the biomedical community.

TOXNET and Its FILES

The new TOXNET computer system in its networked microprocessor configuration continued to evolve over the last year to accommodate new technologies and enhancements. TOXNET's mirrored architecture was upgraded from 386 microprocessors to 486 machines with over 23 gigabytes of online disk storage and 224 available user ports on four terminal servers. One of these terminal servers handles TCP/IP protocol allowing TOXNET to handle INTERNET traffic. System performance over the last year has been excellent at a much lower cost-to-performance ratio than the previous minicomputer implementation.

During FY 1992 two new files were added to the TOXNET system, bringing the total to 14: TRI90 (the Toxic Chemical Release Inventory's 1990 reporting year data), and TRIFACTS, a companion file to TRI, which supplements the environmental release data in TRI with information related to health and ecological effects, as well as to safety and the handling of these chemicals.

Some of the major system-wide enhancements to TOXNET include: (1) Grateful Med access to TOXNET via form screens for simplified menu searching of the following files: Hazardous Substances Data Bank (HSDB), Registry of Toxic Effects of Chemical Substances (RTECS), Chemical Carcinogenesis Research Information System (CCRIS), and the Toxic Chemical Release Inventory (TRI) series. (2) The HSDB file also became searchable by means of special CONCEPT menus. These lead the user through a series of steps without the need to know standard TOXNET command-line searching. The CONCEPT menus are ideal for the novice or occasional user. (3) Another major enhancement has been the development of SHARED PRINTGROUPS that can be used by more than one user code. HAZMAT (Hazardous Materials Response) teams scattered throughout a county, but under a central administration, for example, may wish to have their user codes pooled for SHARE GROUP purposes. (4) Two error correcting file transfer protocols have been developed for downloading TRI data. KERMIT and XMODEM protocols are now available for downloading with eight specially designed TRI report formats.

During FY 1992, several new TOXNET documentation packages were prepared including a Brief Guide for searching files, a Reference Guide and Quick Reference Guide for TOXNET, as well as Reference Guides for HSDB, IRIS, CCRIS, RTECS, and TRI.

Hazardous Substances Data Bank (HSDB) continues to be the most highly used data bank on the TOXNET system, averaging over 600 hours of online access each month. The file building activities for HSDB continue to be supported by the Agency for Toxic Substances and Disease Registry (ATSDR). During this period, nine chemical records were added, 402 records were peer-reviewed by the Scientific Review Panel, and 3,592 records went through Public System Updates. Hazard summaries were prepared for 61 peer-reviewed records and 104 toxicity summaries were written by SIS staff.

Toxic Chemical Release Inventory (TRI) series of files, including TRI87, TRI88, TRI89, and TRI90, remain an important information resource with continued high usage on TOXNET. Mandated by the Emergency Planning and Community Right-to-Know Act (Title III of the Superfund Amendments and Reauthorization Act of 1986), these EPA-sponsored databases contain environment release data to air, water, and soil for about 330 EPA-specified chemicals. During FY 1992, several additional submissions were included to the annual updates of each TRI file. The TRI90 file includes 83,407 records. TRI advanced menus, introduced a year ago, continue to have a significant impact on how the TRI files are accessed.

TRIFACTS is a new database that supplements the environmental release data on chemicals in the TRI series of files with information related to the health and ecological effects, and safety and handling of these chemicals. This data is especially useful to workers, employers, community residents, and health professionals. TRIFACTS is based largely upon New Jersey's Hazardous Substances Fact Sheets, together with EPA-generated ecological data.

Chemical Carcinogenesis Research Information System (CCRIS) continues to be built and maintained directly on TOXNET by the National Cancer Institute. The data bank contains results of research in carcinogenicity for 3,686 chemicals. The chemical-specific data falls into the areas of carcinogenicity, tumor promotion, mutagenicity, and tumor inhibition.

Integrated Risk Information System (IRIS), EPA's health risk assessment file, continued to make significant growth in usage during the past year. The size of the file has increased to 606 chemical records. IRIS is now TOXNET's third most heavily used file. Extension of Concept Menu to IRIS will increase its usefulness even more for novice and occasional users.

Developmental and Reproductive Toxicology (DART) database continues to be accessible through TOXNET. DART contains over 13,000 citations from literature published since 1989 on agents that may cause birth defects. Records in DART contain bibliographic citations, abstracts (when available), Medical Subject Headings (MeSH), and the names and Chemical Abstracts Services (CAS) Registry Numbers (RN) for all chemicals mentioned in the publications. Over half of the records are derived from MEDLINE and supplemented with additional chemical index terms. Records not found in MEDLINE, such as citations to meeting abstracts, articles from journals not indexed for MEDLINE, books, and technical reports, make

up the remainder of the database. Citations from DART have also been used to create a new TOXLINE subfile and will be added to TOXLINE on a quarterly basis.

DART is a continuation of the Environmental Teratology Information Center Backfile (ETICBACK) database on TOXNET. ETICBACK, produced by the Department of Energy's Oak Ridge National Laboratory, contains more than 50,000 citations to literature published from 1950 to 1989. ETICBACK citations are also found in TOXLINE.

Environmental Mutagen Information Center (EMIC) database is produced by ORNL and is managed by SIS. It contains citations to literature on agents that have been tested for genotoxic activity. A new EMIC database, with citations to literature published since 1992, is now being built and maintained on the TOXNET system. Most of the citations are derived from MEDLINE and supplemented with EMIC special keywords, chemical names and CAS RNs. ORNL is also responsible for locating and adding citations to relevant publications not found in MEDLINE. If NLM specification are met, EMIC may be publicly available on TOXNET late in 1992.

A backfile for EMIC (EMICBACK) has been publicly available through TOXNET since June 1989. EMICBACK contains over 70,000 citations to literature published from 1950-1991. Records contain bibliographic citations, EMIC special keywords, and the names and CAS RNs for all chemicals tested.

These four bibliographic databases on TOXNET are funded by the Environmental Protection Agency and the National Institute of Environmental Health Sciences and are operated by NLM.

Registry of Toxic Effects of Chemical Substances (RTECS) is a data bank based upon a National Institute for Occupational Safety and Health (NIOSH) file by the same name which NLM has restructured and made available for online searching on TOXNET. SIS continues to add new data to this file as NIOSH makes them available. This year, an Other Registry Number field was added to indicate previously used CAS Registry Numbers for RTECS compounds. In addition, SIS has contributed to RTECS by locating CAS Registry Numbers for records that lack these important identifiers. RTECS now contains 116,000 records.

Other Programs

Microcomputer Workstation for Chemical Emergency Response

SIS, with the ATSDR, has built a portable, microcomputer-based workstation that provides information assistance to emergency response teams working on accidents involving hazardous chemicals. The prototype, known as ANSWER (an acronym for ATSDR/NLM's Workstation for Emergency Response), has been described in previous reports. ANSWER consists of software modules designed to facilitate easy access to information useful to response teams during emergencies.

Relational Technology

In the Relational Toxicology (RelTox) Project, SIS is investigating the use of relational database management system (RDBMS) technology for building and operating its chemical and toxicological files. The first phase of the project involved developing a relational model of the data elements in relevant files on TOXNET and ELHILL. SIS is now investigating the possibility of using relational technology to maintain and deliver AIDS information to users. The AIDSTRIALS, AIDSDRUGS, and AIDSLINE databases represent a wide variety of information/data types that could serve as a test case for future library systems. This pilot study presents the challenge of integrating large blocks of text, such as abstracts, into a relational database.

Outreach

The Division continued to support the Toxicology Information Outreach Project. It is focused on strengthening the capacity of Historically and Predominantly Black Colleges and Universities (HBCUs) to train medical and other health professionals in the use of toxicological, environmental, and occupational information resources developed at NLM. During the past Fiscal Year, PC-based workstations were installed at the nine participating HBCUs with programs in medicine, pharmacy and nursing. These workstations included state-of-the-art hardware and software (including Grateful Med) for accessing NLM online databases; microcomputer-based tutorials for the CHEMLINE, TOXLINE, and MEDLINE databases as well as the ELHILL software; and a multimedia demonstration for the Toxicology Information Program Files.

Representatives from each participating institution attended a 3-day training class held at the Oak Ridge Associated Universities in Oak Ridge, Tennessee, to learn to effectively search the NLM databases. Subsequently, they also attended a 1-week "train-the-trainers" class to prepare them for conducting classes at their respective institutions. In order to successfully implement this pilot project, each participating institution had free access to NLM's online databases during the start-up period.

This outreach initiative is overseen by the Toxicology Information Outreach Panel. The Panel met at the NLM in spring 1992 to discuss the implementation activities at each institution. These activities ranged from conducting seminars and training classes to community outreach efforts.

As a joint effort with the ATSDR, the Library has extended its HBCU outreach to 13 additional institutions with programs or concerns in the environmental and occupa-

tional health and hazardous waste areas. Representatives from these schools attended a 3-day training class at ORAU and will have free online access to the MEDLARS databases for one year. The first issue of the *NLM HBCU Connection* was published during the summer.

User Support Services

The DOS versions of MEDTUTOR and TOXLEARN, microcomputer-based tutorials for the MEDLINE and TOXLINE databases, respectively, were updated to reflect 1992 MeSH as well as any changes to the MEDLINE and TOXLINE databases. Perhaps the most important design feature of these tutorials is their systematic practice and diagnostic feedback. These programs provide an alternative or augmentation to formal classroom training.

Additionally, the Library began to develop multimedia, microcomputer-based tutorials covering the TOXLINE, CHEMLINE, and MEDLINE databases as well as the ELHILL software. These tutorials are being developed for delivery in a WINDOWS environment and will take full advantage of the WINDOWS graphical user interface while incorporating color graphics, photography, animation, and audio. They will also include mini-databases for practice searching.

Alternatives to Animal Testing

SIS staff have taken complete responsibility for continuing the quarterly annotated bibliography "Alternatives to the Use of Live Vertebrates in Biomedical Research and Toxicological Testing." These bibliographies continue to be distributed free to requestors and are announced as available in the *ILAR NEWS*, a publication of the National Research Council. The NRC's Institute of Laboratory Animal Research (ILAR) under an agreement with the National Library of Medicine also publishes a concatenated annual version of these bibliographies.

Directory of Biotechnology Information Resources

The **Directory of Biotechnology Information Resources** (DBIR) has been available online since January 1989, both as a separate file in TOXNET and as a subset of DIRLINE in ELHILL. It currently contains over 1,900 records, describing databases and other information services, organizations, collections and repositories, publications, and sanctioned nomenclature committees, all related to biotechnology and molecular biology. There are plans to consolidate DBIR within DIRLINE, and discontinue its maintenance and distribution on TOXNET.

Planning Panel on Toxicology and Environmental Health

The Division has been cooperating with the Office of Planning and Evaluation in formulating a summary of conclusions reached by the NLM Long Range Planning Panel on Toxicology and Environmental Health at its several meetings (October 1991, January and March 1992).

Information Services to Other Agencies

SIS, as described, continued providing support for building, maintaining, deploying, and training in the use of information resources as required by ATSDR, NIOSH, NCI, NIESH and EPA. The Division also provided leadership and senior science staff support for the Subcommittee on Information Coordination (SIC) of the DHHS Committee to Coordinate Environmental Health and Related Programs (CCEHRP). Activities of the SIC included determining the data and information needs of risk assessors and epidemiologists, as well as developing record structures for a Directory of DHHS Risk Assessment Projects and a Directory of DHHS Supported or Conducted Epidemiological Projects. Such Directories would focus on the risk posed or effect that may be induced on human health by chemical, biological, or physical agents. SIS continued to represent NLM on CCEHRP's Subcommittees on Environmental Health Risk Assessment, and Testing and Test Method Validation.

SIS also represented NLM during development of the DHHS report titled "Healthy People 2000: Health Promotion and Disease Prevention Objectives" and also represented the Library in development of a "Directory of Exposure Data and Information Resources" as sponsored by the Congressionally mandated Interagency Task Force on Environmental Cancer, Heart and Lung Diseases. Staff participated on the NIEHS Interagency Working Group responsible for the legally mandated Annual Report on Carcinogens. The Division also supports development of the National Toxicology Programs Annual Plan including its review of current DHHS, DOE, and EPA research relevant to toxicology. SIS has made progress in making the full annual plan accessible through the Library's MEDLARS System.

LISTER HILL NATIONAL CENTER FOR BIOMEDICAL COMMUNICATIONS

Daniel R. Masys, M.D.
Director

The Lister Hill National Center for Biomedical Communications (LHNCBC) was established by a joint resolution of Congress in 1968. The Center serves as an intramural research and development division of the NLM. LHNCBC research programs apply state-of-the-art computer and communications technologies to the management of biomedical knowledge. Such knowledge can take the form of procedural rules found in expert systems, information in bibliographic and factual databases, as well as images, electronic signals, and sounds. LHNCBC programs create innovative methods for acquiring, storing, retrieving, analyzing, communicating, and presenting the knowledge of the life sciences represented in computerized form.

A Board of Scientific Counselors meets to review the quality and contents of the intramural research programs within the Lister Hill Center. The Board is composed of scientific and technical experts (see Appendix 6 for a list of members) who are prominent leaders in the fields of medicine, computer science, engineering, and health professions education.

The Center is organized in five component branches:

- Computer Science Branch
- Information Technology Branch
- Communications Engineering Branch
- Educational Technology Branch
- Audiovisual Program Development Branch

The research and development programs of the LHNCBC fall into three categories:

- Computer and information science as applied to the problems of the Library, of biomedical research, and health care delivery;
- Biomedical image engineering, including image acquisition, processing, storage, retrieval, and communications; and
- Use of computer and image technologies for health professions education.

FY 1992 marked the completion of a successful first year of the Undergraduate Research Study Program that provides 2-year scholarships and research experience for sophomore students majoring in electrical engineering, computer science, computer engineering, or physics at participating historically black colleges and universities. The program takes advantage of the staff and facilities of the Center's Communications Engineering Branch to provide experience in biomedical image processing and computer visualization. In 1992 a 5-year collaboration was begun with three educational institutions: Morgan State University (Maryland), Southern University (Louisiana), and North Carolina Central University. Participating students completed their second summer internships at the Lister Hill Center and have now begun the second academic year assignments under the guidance of their preceptors.

Computer Science Branch

Research projects of the Computer Science Branch (CSB) concentrate on the application of artificial intelligence techniques to problems in the representation, retrieval and manipulation of biomedical knowledge. CSB projects involve both basic and applied research in such areas as expert systems, natural language systems, machine learning, and machine-assisted indexing for information classification and retrieval. The research addresses issues in knowledge representation, knowledge base structure, knowledge acquisition, the validation of automated consultant systems, and the human-machine interface for complex systems. Important components of the research include multimedia knowledge-based systems with interactive video capability, and embedded intelligence systems which combine local reasoning with access to large-scale mainframe databanks.

Branch staff members participate in individual and team research projects within the branch. Several are principals in the development of the Metathesaurus, the Semantic Network, and the Information Sources Map of NLM's Unified Medical Language System initiative. They participate actively in the medical informatics and information science research communities and professional specialty societies. And recognizing the importance of addressing the future of medical informatics by helping to train new researchers, Branch Chief Dr. Lawrence Kingsland directs the 8-week NIH "Medical Informatics" elective for third-year and fourth-year medical students each spring.

Expert Systems Program

Expert systems are computer programs that combine knowledge of a particular subject area with inferencing mechanisms which enable them to use this knowledge in problem-solving situations. An artificial intelligence research program concentrating in expert systems was established at LHNCBC in 1984. The objective of the Expert Systems Program is to facilitate computer-assisted access to knowledge. The great variety and creativity of human endeavor is such that this knowledge may reside in different forms, in different places, on different media, with different structures and naming conventions.

The primary research projects of the Expert Systems Program for FY 1992 are the AI/RHEUM consultant system in rheumatology, the COACH expert searcher system, the CTX "criteria engine" shell and its family of auto-

mated testing and validation tools, the medical expert systems evaluation project, and the Rheumatology Image Library videodisc.

In preparation for placement of the system in clinical settings, the AI/RHEUM knowledge base was updated and nearly doubled in scope in 1991-92. The system's data entry process was streamlined and a Case Data Editor module added. In a milestone step for AI/RHEUM and the expert systems evaluation project, the system was tested in 1992 with collaborating clinicians in family practice settings in five sites in Missouri and in internal medicine practices in five sites in Utah. To help non-rheumatologist clinicians make accurate observations, the AI/RHEUM system offers online access to 468 text definitions, to more than 6,800 still images and 23 minutes of brief motion sequences on the Rheumatology Image Library videodisc, to 136 automated MEDLINE searches using the Grateful Med Standalone Search Engine, and to the 44 disease criteria tables which are the heart of its knowledge base.

AI/RHEUM is the best known of a series of knowledge-based medical consultant systems using the criteria table form of knowledge representation pioneered by NLM researchers. CTX, a multimedia expert system shell written at NLM for the development of criteria-based reasoning systems, is in beta-test with projects in several subject domains. One use of the shell is to serve as a building block for other complex projects needing decision-support components. The CTX shell, with its explicit and very unusual multimedia links to knowledge sources in different forms in different places, even on different machines, is one focus of the overall Expert Systems Program goal of providing users with access to knowledge. Several software tools written as adjuncts to CTX provide utilities assisting the developer in manipulating multi-thousand-frame videodisc image banks and in automating the performance evaluation of CTX-based consultant systems against benchmark sets of test cases.

The most recent of the Expert Systems Program projects is the Coach expert searcher system. Coach uses the UMLS Metathesaurus and other knowledge sources to assist Grateful Med users seeking help in improving retrieval from databases on NLM's mainframe. Initial work has concentrated on MEDLINE and its backfiles, and in particular on the problem of null retrieval. Subsequent functions will address searching problems relevant to other MEDLARS files. The system works interactively with the user, with Grateful Med and with ELHILL.

Coach emulates a number of the actions of an expert human searcher in diagnosing user search problems and determining which of a series of functions to invoke for their solution. It has access to multiple knowledge sources built to help augment or replace the user's query terms or to map to new terms in helping users get more retrieval or better focused retrieval. Coach's primary knowledge source, the UMLS Metathesaurus, is an extremely rich resource for which the program offers a browser having both Boolean and non-Boolean search capabilities. The browser can display synonyms, lexical variants, related terms, child and sibling terms, semantic types, Metathesaurus concept definitions and other information, helping users to visualize search terms in MeSH tree context and to choose appropriate additional terms to incorporate into their queries.

The program uses MeSH and its own special knowledge sources to detect occupational specialty headings and synonyms of topical subheadings and map them to better terms or to conceptual clusters of subheadings. Coach knows which MeSH terms can be "exploded" to include their child terms and which are available pre-exploded by ELHILL. It spots appropriate opportunities to offer MeSH "consider also" terms and "forward see related" cross reference terms to improve a search. Coach can guide the user in applying subheading qualifiers for focusing retrieval, displaying dynamically created subheading "pick lists" specific to the MeSH heading in question. The usage of each subheading qualifier displayed in the pick list is explained. Like AI/RHEUM, the Coach project also has reached a major milestone this year: it has entered the alpha-test phase, with use by dozens of colleagues inside NLM.

Dr. Kingsland of the Expert Systems Program served again in FY 1992 as coordinator for the 8-week NIH "Medical Informatics" elective for third-year and fourth-year medical students. Five students from medical schools across the U.S. completed the elective. The course included a seminar series of nearly three dozen 90-minute lectures, independent research projects under the direction of NIH preceptors, and oral and written presentations of research results. Some of these extremely bright, highly motivated students have themselves made important contributions to Expert Systems Program projects.

Natural Language Systems Program

A significant aspect of the interaction between computers and humans involves questions of language. The Natural Language Systems (NLS) Program is concerned with investigating the contributions that Natural Language Processing (NLP) techniques can make to the complex task of mediating between the language of users and the language of the databases they attempt to access. The focus of the NLS program is the development of SPECIALIST, an experimental NLP system for the biomedical domain. The system includes a broad-coverage parser supported by a large lexicon, modules that access the extensive UMLS knowledge sources, and a retrieval module for experimentation in information access and retrieval. SPECIALIST runs on Sun SPARCstations. It is implemented in Quintus Prolog, with some support modules written in C.

The SPECIALIST system includes several modules based on the major components of natural language. The morphological component is concerned with the structure of words and the rules of word formation. The syn-

tactic component treats the constituent structure of phrases and sentences. The semantic component is concerned with the meaning of words, sentences, and discourses. All three rely heavily on the lexical component, which encodes the information specific to the lexical items in the language.

The lexicon currently contains over 58,000 lexical items with over 115,000 lexical forms. It includes both general English lexical items and items specific to the domain of biomedicine. Recent work has focused on the incorporation of semantic information into the lexicon. This involves identifying the various senses a single lexical item may have and assigning semantic roles and a semantic structure to its complements. It also involves developing semantic rules to interpret the lexical structures that have been built.

The morphological component includes rules of inflectional and derivational morphology. Inflectional morphology deals with the different forms of a given lexical item. In English, this is used to mark nouns for number, verbs for tense, and adjectives and adverbs for their comparative and superlative forms. Derivational morphology relates alternates of lexical items that are grammatically related by affixation, but that generally do not share the same word class. Since derivational morphology is highly idiosyncratic in English, it is preferable to store these alternations directly. However, in the case where a particular alternate is missing from the database, rules of morphology are used heuristically to identify the grammatical relationship between pairs of lexical items.

As part of the process of natural language interpretation, the SPECIALIST parser assigns syntactic structures to sentences exhibiting a wide range of linguistic phenomena. The parser reliably determines the syntactic structure which supports the semantic interpretation of a particular sentence. However, it typically also produces numerous additional parses which are syntactically correct yet do not contribute to the final interpretation. Current NLS research is being conducted on a pilot project which has the goal of developing methods for eliminating these unwanted parses.

NLS staff is involved in the UMLS project. Staff members have contributed to the design and development of the UMLS knowledge sources; NLS research activities serve as an application area for the UMLS. The continually evolving Metathesaurus is a rich source of biomedical vocabulary, with the 1992 release including over 130,000 concepts. NLS staff have developed a Unix-based retrieval system called Meta for browsing, navigating, and extracting information from the Metathesaurus. Meta can be used both for batch processing and interactively, to answer individual user questions.

The 1992 release of the UMLS Semantic Network includes 134 semantic types and 47 relationships. The Semantic Network covers the scope of biomedicine, albeit at a fairly coarse level of granularity. It is a testable hypothesis that even this level of granularity can be useful for a variety of fairly sophisticated applications. Initial investigation of use of the Network as a domain model for natural language processing indicates that the semantic types can be profitably used to establish conceptual structures for phrases and sentences in biomedical texts.

An important aspect of the retrieval of information from textual databases involves mapping one conceptual structure to another conceptual structure. This is rarely a one-to-one correspondence, and generally requires a variety of inferences. Recent NLS work has involved identifying and analyzing the types of inferences needed to effect a map between the language of queries and the language of relevant documents. When concepts do not map directly to each other, it is often the case that various types of relations between them are the key to a successful mapping. NLS staff continue to explore these phenomena. Approaches to handling some of them have already been developed.

NLS program staff participate actively in the medical informatics community. Dr. Alexa McCray, head of the NLS program, was co-editor of the first *Yearbook of Medical Informatics* with Dr. Jan van Bemmel of The Netherlands.

Machine Learning Project

A research project began at LHNCBC in 1989 to investigate the subfield of artificial intelligence known as machine learning. The field encompasses a variety of mechanisms for creating computer programs that improve their performance with use. The objective of this project is to develop and apply methods by which programs can automatically acquire knowledge and put it to work.

The underlying motivation for this work arises from the explosion of available biomedical information and the less well acknowledged explosion of the analytical tools and techniques applied to that information. The NLM has long recognized the need for automated assistance to help researchers and clinicians gain access to this extremely valuable corpus of knowledge, and has supplied the community with a wide variety of databases. However, to take full advantage of the anticipated exponential growth of biomedical data and of the increasingly evident interrelationships among previously disparate information sources, dramatic improvements in automated knowledge manipulation, analysis and inference will be necessary.

Programs like expert systems have already moved from the manipulation of information toward the manipulation of knowledge. The Machine Learning Project creates computer programs that not only manipulate knowledge, but also can acquire it themselves. Ideally, a researcher or clinician with a question should be able to have a machine learning program identify where to find relevant information; retrieve that information (possibly from multiple data sources); and analyze and assemble the information into a complete, accurate and comprehensible representation of the desired knowledge.

Machine learning research may also help transcend the traditional computer interaction: a user issuing commands

and a program responding. In a world of rapidly advancing knowledge, programs will have to do more than retrieve information when asked; they will have to manage information retrieval and inference over time. Once a user has specified a question of interest, a machine learning program should be able continuously and intelligently to track evolving knowledge sources for potentially relevant information. When the program finds relevant information, it should automatically assemble, analyze and send that information to the questioner.

These visions are the driving force behind the LHNCBC Machine Learning Project. Currently, machine learning technologies focus primarily on inducing concept definitions from externally specified datasets. To pursue the vision, the Project endeavors significantly to advance the state of the art in machine learning, creating a computationally tractable theory of how to use diverse sources of knowledge and deploy diverse (and complex) analytical tools in pursuit of explicitly stated goals. This approach, called knowledge acquisition planning, is in an early stage of development. Although achieving the vision described above is clearly a long-term goal that will require fundamental advances in basic computer science, the process of developing the theory and implementing prototypes has already produced some useful results.

The primary testbed for research in knowledge acquisition planning at LHNCBC is a program that selects and manages the use of computerized analytical tools and database searches to achieve specific goals. This experimental program, INVESTIGATOR, operates by selecting other programs, such as statistical analysis packages and database search engines, which can be applied to achieve its human-provided knowledge acquisition goals. Each program INVESTIGATOR deploys must have an internal representation describing the preconditions for executing the program: the data formats it requires, its expected outputs, how long the program takes to run, and so forth. From this information, INVESTIGATOR's planning mechanism can select the appropriate tools and databases to accomplish a variety of tasks.

INVESTIGATOR has been programmed to use several important computerized analytical tools and to plan to acquire knowledge in several different domains. The analytical tools include inductive category formation, heuristic Bayesian classification, marker passing intersection search, analysis of variance, random sampling and backpropagation trained artificial neural networks. The databases INVESTIGATOR has accessed include MEDLINE, GenInfo, the Protein Information Resource, Brookhaven National Laboratory's Protein Databank of crystallographic structure information, and others. Results of INVESTIGATOR-managed knowledge acquisition plans have addressed questions in diverse domains including early Eukaryotic evolution, classification of the sequence-level and structural building block of proteins, and prediction of protein structure from sequence.

Our research in knowledge acquisition planning has the potential for high payoff in improving the capabilities of machine learning systems generally. Current machine learning technology is fundamentally limited by the computational complexity of exploring the space of hypotheses compatible with a set of data. The knowledge acquisition planning approach provides important constraints on the space of hypotheses searched by using the specific goals of the learning system to focus computation on the available methods and data most likely to lead to answers to its questions. This method of constraining the search space on the basis of the content of desired knowledge has the potential to result in significant improvement in the performance of machine learning systems.

The general problem of selecting and coordinating diverse and complex sources of knowledge touches on many open questions in cognitive science. The only available models for designing a system that might accomplish these tasks are human beings. A significant component of the Project's research, therefore, is the analysis of human subjects as they acquire knowledge. Machine Learning Project staff work with computer-sophisticated biomedical researchers to gather data on how people manage knowledge acquisition tasks. Protocols of researchers using computer tools and devising retrieval and analysis strategies have been gathered and analyzed to provide insight into this complex cognitive process. Results from these experiments have led to the identification of connections between attentional phenomena in cognitive and social psychology and computational complexity considerations in the design of machine learning systems. Potential implications of this research for the understanding of human cognitive phenomena are also being pursued. The machine learning research program places a strong emphasis on the use of cognitive models in the design of artificial intelligence systems.

Another product of the Machine Learning Project resulted from the need to manage expert protocol data. Building artificial intelligence systems generally requires gathering extensive data about people doing the task to be mechanized. This information comes in the form of interviews, think-aloud protocols, interactive problem-solving sessions and other verbal interactions. Traditionally, the in-formation gathered is painstakingly transcribed and the transcripts are then analyzed by a knowledge engineer. To eliminate the expensive and error-prone transcription step, the staff of the Machine Learning Project designed a computer tool for digitally recording audio onto an optical disk and building an extensive set of indices into the information. Points or passages of the digital audio recording can be marked and annotated, and the annotations can be grouped or organized into a hierarchy. Simply by pointing to an annotation with a mouse pointing device, the user can hear the original audio segment.

By retaining the original audio in easily accessible form, this tool both obviates the need for transcription and preserves potentially valuable nonverbal information such as pauses or "hmmm" sounds. A prototype of this tool, called the AKAT (for Audio Knowledge Acquisition Tool) has been developed for the Apple Macintosh. AKAT is in the beta-testing stage, with versions in use both within NLM and at several external sites.

MedIndEx Project

The MedIndEx Project develops and tests interactive knowledge-based systems for computer-assisted indexing of medical literature currently indexed in the MEDLINE database using terms from the Medical Subject Headings (MeSH) thesaurus. The main objective of MedIndEx is to facilitate expert indexing that goes into the MEDLINE product. Another focus of this research has been developing intelligent retrieval systems utilizing the same representations and environment of the indexing system.

The MedIndEx (Medical Indexing Expert) System is written in a frame language, a type of object-oriented language where objects, known as frames, are used for representing concepts. In a frame, a concept (the frame name) is described as a list of pairs of slots and values, where a slot is a relation, and a value is another frame name that completes the relationship, for example "Heart LOCATION Thorax." Frame descriptions contain not only this detailed factual knowledge, but also procedural knowledge. Specifically, slots contain not only values but also other data, such as defaults, as well as executable procedures that enable the system to assist indexers interactively, as outlined below. Just as slots link frame concepts with values (other frame concepts), subdivisions of slots, known as facets, link slots with these procedures and other data. An important relation in the knowledge base (KB) is known as "inherits-from," which links the entire KB into a single classification. Inheritance, whereby lower-level frames automatically assume descriptions of higher-level frames to which they are linked by this inherits-from relation, achieves a number of important KB functions. These include maintaining consistency of the KB, detecting redundancy in the KB, and simplifying algorithms for accessing frames based on these explicit hierarchical paths from higher-level frames to lower-level frames.

Indexers, with system guidance and help coming from the KB, create for each document indexed a set of indexing frames patterned after KB frames. These indexing frames are descriptions of instances of KB frames; these instances correspond to objects, events, procedures, and other specific descriptions as discussed in documents being indexed. Each indexing frame is linked to its corresponding KB frame by the same inherits-from relation used for linking frames in the KB classification. Indexing frames inherit slots from these KB frames, and since KB frames include executable procedures (indexing rules), this is how the indexing system can give help specific to the concept being indexed. Indexing assistance includes slot names as prompts for indexers to consider indexable aspects of a document, validating indexers' input, prescribing or suggesting slot values based on KB rules, and hierarchical KB displays for browsing permissible values for the current slot. The KB contains rules not only for creating and filling indexing frames, but also for generating in the background conventional MeSH indexing terms at the level of expert indexing. This output can be used to compare the system to conventional indexing, and would provide actual MEDLINE indexing for current retrieval systems.

Important features of the indexing program include detecting inconsistencies in previously stored indexing frames; retention of canceled frames for possible re-use; caching for quick access to large hierarchies; browsing of hierarchical displays, including automatic display of associated MeSH scope notes, regular-expression searching of these displays, and immediate display of the entire KB hierarchy; and word-level aliasing permitting truncation of individual words in a term, which then would be recognized by the system as lead-in vocabulary for frame terms. In addition, a KB Manager tool has been developed, designed to assist knowledge engineers in ensuring a consistent, compact, and syntactically correct KB. This software utilizes the inheritance feature of frame languages, and special scripts employing menu and cut & paste interfaces.

The prototype is written in Sun Common Lisp 4.0 and runs on SPARCstation 2 workstations under the SunOS operating system. Domain-independent project software includes a Lisp-based experimental frame language. The interface uses X Windows (X11 Release 5) and other public domain software (CLOS, CLX, CLUE). MedIndEx is designed to run similar indexing and KB Manager applications in other domains. As of late FY 1992, the KB contained more than 4800 frames (MeSH concepts).

In FY 1992, new features and capabilities for the indexer interface included permitting indexers to attach asterisks (central concept indicators) to conventional indexing output from MedIndEx; more efficient system handling of the consequences of indexers' deleting terms from already stored indexing frames; permitting the indexing of a document to begin by entering two or more concepts in parallel; immediately available display of the entire KB hierarchy including browsing, searching, and selecting functions; and display of MeSH scope notes on request and as part of the browsing of hierarchies.

As enhancements to the KB Manager made during 1992, several complex intellectual tasks were automated. These include generation of semantic labels for contextual hierarchies; detection of inconsistencies in restrictions on slot values that arise as a result of interactive editing of the KB over time; and creating and editing the alias table file used for word-level aliasing in the entry of frame-terms.

In 1992, the MedIndEx project received funding from NIH for the following evaluation project: Development of Experimental Design to Establish Methodology for Evaluating Knowledge-Based Expert Systems for Subject Indexing. Proposers Susanne M. Humphrey (MedIndEx project leader) and Dr. Nancy Wright (Head, Index Section) developed a Request for Proposals announced in July 1992 for accomplishing this work.

Other work during 1992 includes enhancement of the MeSH Query program. This program uses MedIndEx to develop a conventional MeSH-based search, automatically running the search and receiving postings in another workstation screen window where the user is connected to a retrieval service. In addition, there was progress toward using all of MeSH and incorporating elements of the UMLS Metathesaurus for MedIndEx. The 1992 MeSH was created as a frame knowledge base using MedIndEx software (the MedIndEx frame language on top of Lisp). This MeSH KB uses the MedIndEx system of semantic labels rather than MeSH tree numbers for generating the MeSH classification. It includes UMLS semantic types (for example, Pharmacologic Substance) as the basis for the inheritance hierarchy, and the UMLS semantic relations (for example, treated_by) as slots. The linkage of semantic types, as defined in the UMLS, can be expressed as slot restrictions in frames. For example, the Pharmacologic Substance inheritance hierarchy (which includes all MeSH terms that were assigned this semantic type), is expressed as a MedIndEx Lisp function in the restrictions subdivision of the treated by slot in the Pathologic Function frame. These restrictions, which also can be encoded to display a corresponding conventional MeSH classification hierarchy, would be inherited by all frames in the Pathologic Function inheritance hierarchy (i.e., all frames corresponding to MeSH terms assigned the semantic type Pathologic Function or an "isa" to it, such as Disease or Syndrome).

Work planned for 1993 includes enhancing the indexer interface to permit MeSH chemtool terms (from the MeSH Supplementary Chemical Records file) to be entered as indexing terms; designing an evaluation of the indexing system; and enhancing indexer interface software to collect data during system use as preparation for an evaluation of MedIndEx.

Information Technology Branch

The Information Technology Branch (ITB) pursues applied R&D in computer and information science with an emphasis on electronic information generation, storage, and retrieval. Major program areas at present are targeted toward the development of object-oriented full text and fielded data retrieval systems for both online and CD-ROM-based applications. Areas of activity within these current programs include development of generalized windowing interfaces across multiple platforms, object-oriented retrieval systems encompassing fielded data, full text, and graphics objects, editing workstations for manuscript preparation, computer-based publication, and CD-ROM technology. Within these activities, many areas of applied computer science must be addressed, including portability, object-oriented programming, multi-processing, client/server distributed processing models, and advanced memory management.

Full Text Retrieval

The Full Text Retrieval Program in ITB is targeted specifically to address the needs of searching, retrieving from, and updating online medical reference works. A medical reference work, in general, may contain voluminous amounts of text, structure (chapters, sections, sub-sections, etc.), and a variety of objects in addition to standard text such as table of contents, figures, tables, and footnotes. A previous ITB-developed full text retrieval system, IRx1 (Information Retrieval Experiment 1), allowed full text retrieval and maintenance of linear, non-structured, text. IRx1 has been the basis for IRx2, presently under development, a new object-oriented retrieval system designed to address the needs of the aforementioned, more general medical reference work. IRx2 supports Table of Contents (TOC) browsing, Natural Language Query (NLQ) searching, and a Client/Server architecture. The NLQ provides a ranked output of search results to the user (explicit Boolean operations are also available). When browsing the ranked list of matches to a query, the reader is apprised as to the context in which the query was resolved (chapter, section, etc.). This assists in helping the user more quickly identify relevant areas of text. The Client/Server architecture allows for the creation of multiple user-interfaces, across different personal computer systems.

An initial version of IRx2 has been employed to demonstrate prototype access to the Agency for Health Care Policy and Research (AHCPR) *Clinical Practice Guidelines*. Clinical practice guidelines contain all the attributes of other general medical reference works. In addition to the available AHCPR guidelines, two other related databases have been obtained and implemented as full text prototypes: the book, *Guidelines for Preventive Medical Services* (ca. 400 pages), and the database of 89 NIH Consensus Development Conference Reports.

Present IRx2 developments are targeted towards the addition of non-text objects, enhanced retrieval functions and user interfaces, portability to other-than-Unix computer systems, improved communication protocols in support of the Client/Server paradigm, and online text maintenance. Key to the latter is the use of an object-oriented database management system for document storage.

CD-ROM Program

NLM has a growing need to effectively disseminate large full-text databases and/or digitized images and/or

digitized audio in a number of program areas and across multiple platforms such as MS-DOS, Macintosh, and Sun/Unix. CD-ROM represents a unique storage medium for the dissemination of such information. Its salient features include a nominal computer storage of 600 Mbytes (over 1,000 Mbyte of compressed text) per side and a duplication cost of less than $2.00 per disk in large quantities. The effective utilization of CD-ROM, however, is encumbered by an access time much slower than magnetic hard disks. This impacts the response to the user of search software accessing data on a CD-ROM and, hence, requires special consideration in design and implementation of such applications.

In 1990, ITB established a laboratory for CD-ROM developments and initiated efforts to establish technical expertise in CD-ROM design and pre-mastering. The primary laboratory tool is a CD-ROM Pre-mastering and Simulation Workstation that will allow the formatting of tapes for mastering and the simulation of CD-ROM applications even prior to mastering. In addition, a CD-ROM write-once mastering unit allows the creation of a single CD-ROM on site for testing and evaluation. CD-ROMs have been mastered in the laboratory for several NLM programs areas.

In concert with the need to disseminate information on CD-ROM, ITB has extended the original object-oriented systems design for IRx2 to encompass fielded as well as full text data (see below).

Fielded-Data Information Retrieval

Many databases targeted for CD-ROM are of the fielded-data type rather than full-text. Examples of such databases are the NLM database of journal titles, Journal Information System (JIS), the NLM ChemID and TOXLINE databases, and the EPA Toxic Release Inventory (TRI) database. An object-oriented search engine capability, IRxFD (IRx Fielded Data), has been developed incorporating incremental searching of terms and user-defined data types. The latter allows the database fields to be of arbitrary hierarchy or complexity. The incremental search capability enables the user to select long and/or compound terms with the entry of a small number of characters. The search engines for the fielded-data databases have been designed to run on local personal computers.

The Journal Information System (JIS) was implemented during FY92 and made available to users in the NLM Reading Room. In addition to its application to other databases, further development of IRxFD is targeted towards generalizing and improving the initial object-oriented system implementation. In the future, this system will be merged with IRx2 to provide an integrated object-oriented retrieval system for full text and fielded data.

Communications Engineering Branch

Image engineering R&D projects in the Communications Engineering Branch focus on the capture, storage, processing, online retrieval, transmission, and display of both biomedical documents (mainly journals) and medical imagery. Areas of active investigation center on image compression, image enhancement, image understanding, pseudo-grayscale rendition, image transmission and networks, omnifont text recognition, and man-machine interface design. This applied R&D is directed toward such critical NLM tasks as document delivery, archiving, and preservation. In addition, there is research into imaging techniques that support medical educational packages employing digitized radiographic, dermatological, and other imagery.

System for Automated Interlibrary Loan (SAIL) Program

SAIL is an R&D program that has created a prototype automated document image delivery system to support the NLM's interlibrary loan (ILL) service. This collaborative program (with components of Library Operations) is motivated by the increasing burden faced by the Library in servicing the interlibrary loan requests in the traditional manual way. One difference between the approach adopted here and in earlier imaging projects is that R&D is not confined within the walls of the laboratory: *real users* are receiving actual documents, and hence it is possible to make an accurate assessment of the technical and operational problems in a scaled up system.

The pilot phase of this program was completed in FY 1992. More than 9000 ILL requests representing about 5 percent of the NLM's ILL load were received for document images stored in SAIL's optical disks; 82 percent of these were filled. The prototype system was gradually improved to overcome practical problems discovered during the evaluation phase. The next steps are to: expand coverage by including Internet as a delivery path; expand online content by incorporating a robotic jukebox; and introduce a higher degree of automation in the tagging, quality control, and browsing functions to reduce operator errors.

A number of engineering studies are being pursued in support of SAIL development and the library's ILL activity. Among these are the following:

Performance Studies: Beginning in a small way in mid-1991, SAIL delivery reached a high point in March 1992, receiving 62 requests a day and filling 55 of them (about 89 percent). This amounted to about 5 percent of the total requests received at NLM. In terms of overall fill rate, SAIL has filled 82 percent of all requests it received. Of the remainder, 9.5 percent were not filled because requesters declined to pay the standard NLM charges for the documents, and 8.5 percent were not filled for other reasons such as: the documents were not yet in SAIL, they had not been tagged, or the requester wanted more than the 50 pages that NLM's policy allows. SAIL fills 30 percent within 15 minutes, 64 percent within 60 minutes, and over 80 percent within 3 hours. Requests that take longer than 15 minutes to fill are those that require operator intervention either to

resolve ambiguities in the requests or to exchange optical disk platters in the limited number of drives.

Document conversion is the most time intensive of SAIL's operations. The original bound volume scanner converts 28 pages per hour. Since most journal issues could be guillotined, the fast scanner (DCW-2) developed later in the year was used extensively. The average rate using this scanner was 314 pages an hour.

Cost Studies: The cost of both conversion and delivery were calculated by a top-down and bottom-up approaches. Cost elements include labor, equipment, and supplies. The cost of converting paper documents to electronic form amounts to about 28 cents a page with the fast loose leaf scanner and about $1 a page with the bound volume scanner. Using the fast scanner, it costs about $3 to convert the average article. The cost of delivering the article varies significantly with the degree of operator intervention. Predictably, the least expensive operation is the completely automatic "online" case. More expensive are the "offline" cases where some degree of human intervention is required (for example, if key information identifying the article is missing or if a disk must be switched). Depending on these cases, the delivery costs range between $3.50 and $5 per article.

Artificial Neural Net (ANN) Studies: Certain "classification" problems encountered in SAIL operation lend themselves to neural network-based solutions. One such problem was that SAIL automatically filled requests that the requesters did not want filled. The requesters make this fact known by entering comments in an unstructured comments field in the standard ILL request. These comments might be of the form *NLM Do Not Fill*, but are inconsistent. Analyzing several thousand requests, researchers found nine keywords or word pairs that were strong indicators to fill or not fill the requests, and developed a subsystem consisting of a parser and an ANN of the back error propagation type. Capitalizing on the sparse pidgin-like language used in typical comments, the subsystem begins by extracting *NLM* and the other keywords from the comments field. These keywords are used as a nine-element input vector to the ANN whose output vector consists of one element: either *fill, do not fill*, or *uncertain*. Preliminary results show that the network correctly decides to fill or not fill 90 percent of the requests; operator intervention is required for the others. By significantly reducing the operator intervention required, this research promises to yield time and cost savings in future operational systems for automated document delivery.

Another ANN project is intended to aid scanning operators by automatically identifying a journal. This is accomplished by processing the image characteristics of a journal's cover page. In one approach the black pixel distribution of the digitized cover page serves as the input vector to an ANN. In a second approach, the black pixel distribution is initially processed by a Fast Fourier Transform whose first 35 coefficients serve as the input. The first approach was found to be successful in correctly classifying 25 different journals, but was slow to train. The second approach is less successful in classifying journals but is faster to train. Research is continuing to define the optimum ANN design to identify journals.

Automated Skew Detection: In document conversion, errors are detected at the quality control (QC) stage. When the bitmapped image is judged to be of poor quality (e.g., a skewed image), the operator rescans the page. Since rescanning after the QC stage is inconvenient and disruptive to the process, it is desirable to detect such errors during scanning. This is the motivation behind research in automated skew detection, which uses a multistage technique. This technique starts with the detection of page orientation (landscape vs. portrait, with an accuracy of 99 percent), followed by the detection of the bottom parts ("feet") of the characters (eliminating the rest of the pixels to increase computational speed), and then the application of the Hough transform that detects specific structural relationships between pixels in an image resulting in a numerical value for the skew angle as the output. Tests show the successful detection of angles as small as half a degree. Further research continues to optimize these algorithms and implement them in a subsystem that is part of the document capture process.

Document Identification Strategies: The hypothesis described in last year's report, that a combination of the issue identifier (a number called the MRI) and the starting page number of the article would adequately identify a requested article was confirmed by research in FY 1992.

Simulation Studies: Studies described in last year's report to predict a migration path for a scaled-up SAIL system were continued in FY 1992.

Digital X-ray Prototype Network (DXPNET) Program

The main goal of DXPNET, which also stands for *Digital X-ray Prototype workstations linked via InterNET*, is to investigate the technical feasibility and design issues in developing, maintaining and operating an archive of digitized radiographs. This is a collaborative program in which the Communications Engineering Branch on behalf of NLM serves as Technical Manager. The other participants are the National Center for Health Statistics (NCHS) and the National Institute of Arthritis, Musculoskeletal and Skin Diseases (NIAMS). The program supports the National Health and Nutrition Examination Surveys (NHANES) which NCHS periodically conducts to produce statistics on the health status of the U.S. population. One element of the collected data consists of radiographs, 17,000 from a survey already completed and an expected additional 10,000 from a current survey. The radiographs are of cervical and lumbar spine, and hands and knees.

The problems that DXPNET addresses are: *preservation* of the radiographic collection, *extracting information*

from this collection, and providing remotely located users with *access* to it. Specific program objectives include: 1) develop an Electronic X-ray Archive (EXA) for digitized NHANES x-rays implemented via an optical disk jukebox controlled by a UNIX workstation interfaced to the Internet; 2) develop image retrieval and display workstations, also called Standardized Readings Workstations (SRW), suitable for radiologists to remotely access the images from the archive, retrieve, display and manipulate them for reading; and 3) evaluate technical aspects of the design including compression technique, response time, and use patterns.

In FY 1992, the Communications Engineering Branch developed and tested an affordable prototype PC-based workstation that enables technicians from NCHS to perform quality control on the digital images produced by scanning the x-rays, an activity done under contract at UCLA's Division of Medical Imaging. Hardware components such as a 386 computer, a 1280 x 1024 pixel monitor, a WORM drive, and an imaging board were integrated and a complete image retrieval and display software system was developed. After an evaluation phase, the Branch deployed this workstation at NCHS where it is currently being used for quality checking the images received from UCLA on 5 1/4 inch WORM-type optical disks.

The next steps under way are: to develop the EXA based on a commercially available robotic optical disk jukebox to be located at NLM, to develop a pair of UNIX-based SRWs that will allow NCHS radiologists to create standardized readings by viewing the xray images on a high resolution (2048 x 2560 pixel) monitor, to develop the capability of linking the workstations over Internet, and to deploy one of the two SRWs at NCHS to allow image retrieval from the archive located at NLM.

Ongoing activities include the design of the software modules for the Standardized Readings Workstations and the Electronic X-ray Archive. Also, since image compaction is a key consideration both for maximizing the online storage capacity (a possible total of 224 gigabytes) as well as minimizing transmission time, compression studies are under way incorporating a community of collaborators at NIH, Stanford and Monash (Australia) Universities. While the JPEG technique employing the Discrete Cosine Transform is being investigated inhouse, our extramural collaborators are experimenting on the images we send them over the Internet with their implementations of the vector quantization, DPCM, and wavelet techniques. The objective of the compression studies is to identify the compression technique consistent with maximum compression ratio and necessary image quality.

The software engineering component of this project is both the most challenging and the most potentially useful. The key to success will be the identification and satisfaction of user needs. The software will make the image database accessible. It will allow the retrieval of classes of images based upon user-supplied search terms. The NHANES database contains all of the information on each participant; the x-ray images constitute just one element of a unit record. The ideal system will have the entire NHANES database searchable online so that all possible search criteria are available. Local retrieval of images and other health statistics could be achieved through remote searching of the database. If the NHANES database is not available online then an alternative search mechanism will be devised based on a subset of the database. Papers on the DXPNET activity were presented at the ASIS Mid Year Conference (May 1992) and the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference (September 1992).

Machine-Readable Archives in Biomedicine (MRAB) Program

The long-term goal is to build a prototype system for a machine-readable archive. An earlier imaging program addressed the problem of efficiently converting paper documents to electronic bitmapped image form for purposes of permanent storage. However, access to the stored material remained at the bitmapped image level. Although this level of access is certainly faster and more convenient than that available with paper and film archives, electronic imaging offers the potential for far greater access than merely to the image. It allows the creation of vast machine-readable text/image databases in very compact space (the equivalent of 350,000 to 400,000 document pages on a single disk, or about 50 million pages on a 128-disk jukebox), the capacity of rapidly and randomly navigating through this store, the ability to apply library processing (indexing and cataloging) to parts of the older literature, and the capability to access this store remotely.

The Center's program is motivated by several considerations:

1. The potential for making the biomedical literature available in compact and accessible electronic form will be demonstrated.
2. Certain objectives of the national preservation effort in biomedicine spearheaded by the NLM would be served. For example, medical libraries would not need to replace their brittle documents with thousands of reels of microfilm. Furthermore, their users would be better served by the access to both bibliographic pointers and content.
3. It would enable far greater access to the older biomedical literature than available today.
4. The availability of the older material in machine-readable form will allow it to be indexed and cataloged to more sophisticated and useful levels.
5. It would permit research into historic terms corresponding to modern medical vocabularies.

In principle, a document is made machine-readable by applying optical character recognition (OCR) technology to its bitmapped image. OCR in some form has been commercially available for several years. At the low end of the cost range, there are handheld and desktop OCR devices

that adequately handle typewritten, uniformly spaced material such as office correspondence. At the high end, there are devices that do omnifont recognition for typeset, proportionately spaced material such as books and journals. However, none can adequately meet all of the following conditions: bound documents, fragile paper, material with poor print to paper contrast, and compound documents with intermixed text and graphics where it is required to save these different data types in a form suitable for display in the original format. For the material under consideration here, all of these factors hold.

The long-term goal of the MRAB program is to address the problem by: 1) developing a prototype system to demonstrate the feasibility of creating, maintaining, and accessing a machine-readable archive of documents, 2) using this prototype as a testbed to answer questions concerning the process of creating this archive, 3) evaluating the role of such an archive at the NLM, and 4) defining the technical specifications for an operational system.

Last year's report described image enhancement studies. Current activities in MRAB include evaluations of commercial OCR devices and investigation of techniques for image enhancement and image segmentation.

OCR Evaluation Studies: Omnifont character recognition finds applicability at the Library. To investigate the areas where OCR may be applied, an inhouse system built around a Calera OCR engine as well as an off-site Kurzweil 5100 system were used to conduct performance studies. The material tested included the Surgeon General's *Index-Catalogue* and journals representative of those indexed for MEDLINE. The *Index-Catalogue* is a particularly difficult document from the character recognition standpoint because of the frequent occurrence of contiguous characters due to bleedthrough. Also, amino acid sequences printed in biotechnology journals were scanned and converted with the inhouse Calera-based system. The performance criteria included conversion error rate and conversion time. Results have been published in the literature. Last year's report described studies aimed at reducing the keyboarding activity currently done to enter fields such as author, article title and abstract, into NLM's indexing database.

Image Segmentation Studies: The decomposition of a bitmapped image into its constituent parts is of importance in a variety of EI functions such as image compression and border (page edge effects) removal. In association with the Undergraduate Research Study Program, a project was initiated to develop software to automatically remove unwanted borders. The techniques employed include: first order statistics, second order statistics and image morphology. The first method calculates the black pixel statistics and attempts to differentiate different regions on this basis. The second method uses an autocorrelation computation to get a measure of the inter-pixel relationships within each row and column of the image. The third method uses the morphologic dilation operator to create "blobs" of different sizes and shapes and differentiates regions on this basis. The first order statistics method was implemented and a sample of typical bitmapped document images was tested. The results indicate successful segmentation for images which are entirely text, but not for compound pages (text with interspersed graphics). Research will continue with the other two techniques this coming year.

Another approach to automatic image segmention was pursued using a particular implementation of an artificial neural network (ANN) model, the back error propagation paradigm, as a segmenting tool. This paradigm is known for its ability to "learn" to classify data. Two approaches were tried, both using unprocessed document image data generated in the lab. In one approach, the image was divided into tiles. The network then determined if the image data within the tile was text or nontext. In the other approach, histograms of the total numbers of black bits in vertical columns across the page image and in horizontal rows down the page were generated. These histograms were presented to the network, which then determined the location of the edges of the text. Error rates for both approaches ranged from 5 to 25 percent, encouraging enough to pursue additional study. Future efforts will focus on methods to preprocess image data to extract features that are more representative of each image data class.

Biomedical Digital Image Processing (BDIP) Program

This R&D area is directed toward the development and evaluation of improved techniques to capture, store and display medical images for computer-based educational systems. Biomedical images of interest include both multiple gray-level images (e.g., radiographs) and high-resolution color (e.g., dermatologic slides). Other images of interest include monochrome and color material from the historical collection. The results of research in this area are reported in the Proceedings of the Society of Photo-Optical Instrumentation Engineers and similar publications.

Digital Color Imaging for Dermatology: Research to determine the level of image compression acceptable for images used to teach dermatology showed unequivocally that dermatologists can identify skin lesion morphology from images compressed to about 10:1 with a Hadamard transform-based algorithm, as well as they can from the original 35mm slide. It also showed that, although compression ratios of about 20:1 and 30:1 did not degrade performance except in one case, subjective opinions indicate that these compression ratios result in images of slightly degraded quality. The images were digitized at 1000x1000 pixels over a 35mm slide with 5 bits for each of the red, green, and blue color components. The results of the study are reported in the Proceedings of the SPIE, vol. 1232.

Image Compression Techniques Evaluation: The results of a study investigating the use of the Discrete Cosine Transform (DCT) to compress digitized color dermatologic images were published in February 1991 following a

presentation at SPIE's Medical Imaging V Conference. 35mm slides of four morphologic types of skin lesions were digitized at 1000x1000 pixels with 8 bits for each of the red, green, and blue color components. Using the DCT algorithm, the resulting image files were compressed at compression ratios of about 7:1, 28:1, and 70:1. The original scans and the decompressed files were written to a 35mm film recorder. Together with the original photo slides, the slides resulting from digital images were evaluated in a study of morphology recognition and image quality assessment. Fifteen dermatologists were asked to identify the morphology depicted and to rate image quality in each slide. We conclude that the use of DCT-based compression yields acceptable performance for skin lesion images since differences in morphology recognition performance do not correlate significantly with the use of original photos versus compressed versions. Additionally, image quality does not appear to correlate significantly with level of compression. Future work in the area of image compression will include the evaluation of a high-speed nondestructive algorithm applied to radiographs.

Engineering Laboratories

To conduct the research and development outlined in the foregoing, the Communications Engineering Branch designs, equips and maintains two inhouse laboratories:

Signal Processing Laboratory. This laboratory houses advanced systems to electro-optically capture the digital images of documents, both in bound volume form and as loose sheets. Subsystems are available to perform image enhancement, segmentation, compression and storage on digital optical disk media. The laboratory also includes workstations for performing quality control and tagging for the captured documents. These bitmapped images may be retrieved in conjunction with a search of NLM's bibliographic databases or the DOCLINE document request system that serves the interlibrary loan function at NLM.

Specific equipment developed inhouse includes high performance loose-leaf and bound volume scanners using charge coupled devices; optical disk drives; laser printers; high-resolution (200 dpi) softcopy display devices. These are configured into systems that serve as laboratory testbeds to support research into automated document delivery, document preservation and document archiving, and techniques for image enhancement, manipulation, segmentation, compression for high density storage and high- speed transmission, omnifont text recognition, and related areas.

Image Processing Laboratory. The Image Processing Laboratory supports the investigation of image processing techniques for gray scale and color biomedical imagery at high resolution. It consists of computer resources and image processing equipment to capture, process, and display such high-resolution digital images. The equipment includes both Sun/Unix workstations, Sun-compatible workstations, and IBM PC-compatible computers, all linked via an Ethernet local area network.

The Sun/Unix machines include a SUN S390 SPARC server with a large disk storage capacity (3 gigabytes). This machine operates as a file server for Unix machines both inside and outside of the Image Processing Laboratory. Additional storage for the S390 is provided by a 9 1/2 inch tape drive and an 8mm tape drive.

One of the other Unix workstations is a Sun 4/260. This system supports development of the Standardized Readings Workstation for the DXPNET project. A very high resolution Megascan monitor is attached, capable of displaying 2048x2560 8-bit grayscale pixels. This monitor is intended primarily for displaying x-ray images for DXPNET. A conventional Sun color monitor is also attached, and this monitor displays the screens which provide the man-machine interface for the DXPNET users.

Two Sun-compatible Solbourne workstations in the lab provide development platforms for DXPNET and SAIL-related work, as well as support for the Undergraduate Student Research Program. In addition, the lab has an older Sun 3/50 system.

There are three PCs in the lab attached to the Ethernet. They are used for display and gray scale image processing operations, an Internet file server in order to provide access to medical images for researchers in remote locations, and a platform for the capture and display of gray scale or color images.

Educational Technology Branch

The Educational Technology Branch conducts research and development in computer and multimedia learning technologies, disseminates information about these technologies to NLM's various constituencies, and supports their application in health professions education. These endeavors represent a long-standing commitment to support and develop innovative approaches to training health care professionals. Activities have a significant audiovisual component, typically combining microcomputer and optical disc technologies and using analog videodisc-based images and varied digital image formats. The program includes intramural development of new educational technologies and liaison to health professions schools and professional societies for field testing and other collaboration. Programs developed both internally and elsewhere are demonstrated and information about the technology is shared through publications, seminars and other means.

The Learning Center for Interactive Technology (TLC) serves as the focus of these and other major branch activities which include the Computer-based Curriculum Delivery Systems (CCDS), Dermatology Visual Database, and Visible Human projects.

The Learning Center for Interactive Technology

The programs of The Learning Center for Interactive Technology include ongoing consultations and demonstrations, workshops and seminars, multimedia research and development, monographs, reports and other publications, database development and maintenance, and computer conferencing.

TLC is a "hands on" laboratory supporting these activities. The Center has three components: 1) a demonstration area; 2) a multimedia development area; and, 3) an interactive training facility. Since its establishment in 1985, its functions and size have steadily grown. The demonstration center has 22 carrels and more than 100 interactive programs. Large group demonstrations are presented from a carrel configured for video/data projection, and are also conducted in the training facility. Program formats include patient management problems, tutorials, evaluation, visual databases, expert, systems and information retrieval.

In FY 1992, TLC staff provided more than 400 demonstrations and "hands on" experiences for 1,215 visitors. This brings the total number of visitors since the Center opened in 1985 to 5,994. While the number of visitors is important to the accomplishment of TLC goals, outreach activities described below are equally important.

In FY 1992, four workshops were conducted in the training facility, at conferences, and at health science institutions. Thirty faculty were trained in videodisc repurposing workshops conducted at the NLM, the University of Alabama Medical Center, and at the Health Sciences Communication Association meeting. An authoring system workshop was offered to 30 attendees of an Association for Educational Communications and Technology/Health Sciences Communications Association meeting. This was an initial prototype conducted as part of the workshop development process. It will be enhanced and expanded to provide "hands on" learning experiences at The Learning Center. Several authoring tools are being evaluated for possible use in the workshops. In addition, 24 seminars were conducted in The Learning Center, at the Radiological Society of North America and at other sites. Also, 57 training sessions for 675 NLM staff were held.

Two major research reviews were completed in the areas of ergonomic factors influencing the design of learning environments and criteria and features used in the evaluation of computer-based education programs in nursing and other health science disciplines. The ergonomic study reviewed research on visual display workstations, specifically: 1) potential health hazards from electromagnetic radiation; 2) musculoskeletal disorders; 3) vision complaints; and, 4) psychosocial stresses. Ergonomic guidelines to balance human performance with learner satisfaction and well-being were presented. The evaluation study reviewed and analyzed assessment instruments used by health science professions education programs. Study results will appear in a monograph.

An extensive review of authoring system software begun in 1991 continued. AuthorBase, a database of authoring system software created in 1991, underwent a major update. The database was made available online via SprintNet (Telenet) and Internet. Information on existing programs was revised and records were added for new products. Authorbase currently includes information on over 100 authoring tools. Users can browse information about current and no longer supported systems and identify those that work with different platforms, operating environments, and multimedia. They can also retrieve citations to articles about different products. A database of courseware in The Learning Center also was developed.

Policies and procedures were devised for disseminating the Interactive Technology Sampler Videodisc completed last year. The disc contains portrayals of 20 interactive technology programs and serves as a general overview to interactive technology and The Learning Center itself. The disc is now routinely used in the Center and has been exhibited at national conferences and used in presentations at national meetings. One hundred copies have been distributed to health science libraries and institutions.

The Learning Center's online computer conferencing system, Educational Technology Network (E.T. Net), expanded and enhanced its service to the health professions community. E.T. Net came online over SprintNet (Telenet) in 1989, and in 1990 became available via Internet. E.T. Net is open to professionals engaged in either the development or use of interactive technology in health science education or nursing care research. It is available at no cost, 24 hours a day, 365 days a year.

Users join conferences as they participate in E.T. Net. As of September 1992, the active conferences and the number of registrants in each were as follows: AVLINE, 433; CAI, 462; Hardware, 205; Shareware, 519; Digital Images, 229; NUCARE (nursing), 373; UMLS User, 72 and General, 840. E.T. Net participation is growing steadily. Because SprintNet and Internet have international connections, E.T. Net registrants include colleagues in Canada, Europe, South Africa, Australia, and the United Kingdom.

Monographs and research reports written by the staff continue to be revised, reprinted and disseminated. Current titles include Optical Disc Technology, Interactive Technology, Videodisc Repurposing, Authoring Systems and Computer-Based Education in Nursing. A research report, Guidelines for Designing Effective and Healthy Learning Environments for Interactive Technologies, was completed.

Computer-based Curriculum Delivery Systems (CCDS)

The primary goal of CCDS is to design, produce, and test experimental technology-deliverable curricula for the health professions.

In June 1983, CCDS completed the first such prototype curriculum—a two-sided videodisc for Basic Medical

Pathology that was controlled by computer programs written for the Apple II-plus machine It was offered for field testing to all departments of pathology in U S medical schools and testing began at 12 schools in the autumn semester Since that time, CCDS recruited 10 additional pathologists to work on the project and produced 8 additional programs The field-testing network has grown from 12 to over 100 schools that have tested not only the technology-deliverable pathology curriculum but also prototype curricula in mental health (Teenage Depression and Suicide Risk Assessment) and in orthopaedics ("The Chronic Unstable Knee," "Knee and Forearm Anatomy," and "Magnetic Resonance Imaging of the Human Knee") There are now 78 health professions schools in the U S , Canada (6 schools), Europe (3 schools), Southeast Asia (3 schools), and the Caribbean (2 schools) involved in the pathology project The original 12 student work stations has grown to over 1,000 with new stations added almost daily

During FY 1992, CCDS brought out the second edition of the first two programs to make them conform to the format that was adopted after two years of testing The new videodiscs contain all of the images from all of the other videodiscs The database that describes this image collection has 8,810 records

During FY 1991-92, CCDS developed a new generation of the computer software to support learning experiences in basic pathology It is written in a relational database language and runs on local area networks (LANs) It gives local faculty complete control over presentation of the material by allowing modification of the databases (DBASE and Clipper) Also in FY 1992, CCDS acquired the necessary hardware and software and began experiments designed to determine whether this large image database can be converted to digital format to support future generations of technology-based curricula

The impact of the pathology programs can be judged in part on the basis of their rapid adoption and the following data Over 8,000 student evaluations of the programs have been received On a scale of 1 to 5, students rated the programs 4+ as a learning experience Comparison of pre- and posttest scores of those students not passing the pretest (a 70 percent score) show that students master the material in about one-third the amount of time devoted to it by traditional methods At least three schools have canceled all lectures on topics covered by CCDS lessons and others have reduced lecture time and made the CCDS lectures a required part of the course

CCDS staff, in collaboration with a guest investigator in the laboratory, completed two new orthopaedic prototype programs for use at the 1992 annual meeting of the American Academy of Orthopaedic Surgeons (AAOS)

Dermatology Visual Database Project

Commercial electronic technology for color image capture, recording, and display has proliferated rapidly Consonant with the recommendations of the Lister Hill Center Board of Scientific Counselors, the utility of this technology for dermatology education is being explored

Preliminary comparison of skin lesion photographic slides scanned and displayed in superVGA (640x480 pixel, 8- bit RGB palettized color) and Intel Digital Video Interactive (512 x 480 pixel, 16-bit YUV color) did not indicate advantage to the more expensive DVI technology The Sulzberger Institute for Dermatology Education of the American Academy of Dermatology produced a CD-ROM with a test collection of images in both formats to allow for more extensive evaluation Analog capture with component video of skin lesion photos also received a favorable preliminary evaluation by the AAD Subcommittee on Electronic Information, again a CD-ROM recording will facilitate a more extensive evaluation Prospective capture of skin lesion images by simultaneous photography and video will answer the question of whether photo to video transfer is the step in which critical information loss occurs

The melanoma interactive video tutorial is now being evaluated for learner performance in undergraduate and graduate medical education at the University of Arkansas, University of South Alabama, Uniformed Services University, George Washington University, and Johns Hopkins University Both digital IBM M-Motion and analog InfoWindow touchscreen versions are available to other institutions that wish to evaluate them

Audiovisual Program Development Branch

The Audiovisual Program Development Branch (AVPDB) conducts media development activities with three specific objectives As its most significant effort, the branch supports the LHC's research, development, and demonstration projects with high quality video, audio, and graphic materials From initial project conceptual images, through actual project implementation with image preservation, transfer, and display, to project evaluation and reporting, all forms and formats of imaging are produced

Creative consultation and materials development are also provided by the branch for the NLM's educational and information programs With the added mission requirement of the Library to increase its outreach activities, the support that the branch provides to these type of programs has increased significantly From optical media technologies to teleconference support, the graphic, video, and audio materials requirement has increased in quantity and diversified in format

The third area that the branch concentrates on is technical development issues such as image resolution, color fidelity, media transportability, media storage, and image communication In addition to technique development by the staff, the facilities and hardware systems must reflect state-of-the-art standards in a very rapidly changing field High Definition Television (HDTV) is a development area

being explored that represents the future for improved electronic image quality Multimedia techniques are being pursued for the educational and cost advantages that they offer. Three dimensional computer graphics, animation techniques, and photorealistic rendering methods have changed the tools and products of the graphic artists in the branch. Digital video and image compression techniques are central to projects being pursued to improve image storage and communication With all of the technologies being brought together in the development and support projects of the branch, the central core expertise remains the creative, artistic, communication skills of the staff

Cardiac Embryology

APDB is continuing work on the project begun last year in which 3-dimensional computerized representations of the embryonic development of the human heart are being created APDB is playing a leading role in a multi-institutional demonstration program to develop this scientific visualization project on cardiac embryology The graphics staff is transferring images from the Carnegie Embryology Collection, as computerized planar sections and assembling them in 3-dimensional wire frame forms on Macintosh computers. Techniques of solid rendering, transformation, animation, and levels of transparency of the developing organ are being constructed in a common file format for display on a microcomputer and, as motion sequences, on laser videodisc End products are intended for use as reference materials and as teaching aids in medical school curricula and in the continuing education of medical professionals beyond their pre-clinical training The project uses the latest computer graphics systems and high speed electronic networks for the purpose of markedly increasing the teaching effectiveness of classic collections of human embryonic material The final product will be a widely accessible, desktop-computer-based, interactive hypermedia presentation, available on laser videodisc or on CD-ROM Collaborating institutions include the Center for Advanced Instructional Design, Yale University School of Medicine, the Departments of Pathology and of Art As Applied to Medicine, the Johns Hopkins University School of Medicine, the National Museum of Health and Medicine; the Armed Forces Institute of Pathology, and the Walter Reed Army Medical Center

Cervical Cancer Success in Sight

Continuing a joint project with the Early Detection Branch of the National Cancer Institute, APDB has developed an educational program for health professionals on the early detection of cervical cancer Drawing on an approved base of information, an interactive laser videodisc was initially designed for the IBM M-Motion system The resulting program was awarded "Outstanding Scientific Exhibit" at the annual meeting of the American Academy of Family Physicians Making use of the same laser videodisc, a second program was written for operation on the IBM Infowindow System

The material is currently being adapted for the Compact Disc-Interactive (CD-I) format, to exploit the audio, video, and programmatic capabilities of that all-digital platform The Branch plans to redesign, develop, and produce additional instructional versions of this subject material to take advantage of the unique strengths of other emerging multimedia systems

Lister Hill Center R&D Program Modules

The branch produced 16 short (average length, less than two minutes) videotape modules, highlighting the principal research and development projects being conducted by the Lister Hill Center. These brief modules, designed for periodic modification as projects evolve, are also transferred to DRAW (Direct Read After Write) videodisc format, for pre-programmed selection and speed of access

Educational and Information Program Support

- Eight examples of remote location hospital information programs were documented on videotape, and an overview videotape unit was produced for use in a two-part satellite teleconference, titled "Information STAT Rx for Hospital Quality "
- 22-minute and 14-minute versions of a documentary videotape program, "Plagues and Politics," were produced to highlight the history of the U S Public Health Service The Service plans to distribute some 500 copies of the longer version
- A videotape program provided by the Society of Hospital Pharmacists, "Images of Hospital Pharmacy in America," was augmented with branch-produced introductory materials, converted to videodisc format, and mounted in a specially designed kiosk for use in a History of Medicine Division exhibit in the main lobby of the Library
- Members of the branch's graphics unit and production staff cooperated in the production of an interactive animated desktop computer module and a continuous loop videotape display, listing and illustrating various Lister Hill Center development projects These materials were used by LHC representatives at a locally held national meeting of SCAMC

Audiovisual Support Activities

The branch continues to upgrade the equipment used to support meetings held in the Lister Hill Center Auditorium and the NLM Board of Regents Room APDB also provides preventive maintenance for audiovisual recording, playback, and projection equipment used in other meeting rooms throughout the Library

NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION

David Lipman, M.D
Director

The National Center for Biotechnology Information (NCBI) was established by Public Law 100-607 in November, 1988, as a division of the National Library of Medicine The establishment of the NCBI reflects the importance of information science and computer technology in the understanding of the molecular processes that control health and disease. The Center has been given the responsibility to

- Create automated systems for storing and analyzing knowledge about molecular biology, biochemistry, and genetics, Perform research into advanced methods of computer-based information processing for analyzing the structure and function of biologically important molecules and compounds;
- Facilitate the use of databases and software by biotechnology researchers and medical care personnel,
- Coordinate efforts to gather biotechnology information worldwide.

  There are presently 40 senior scientists, postdoctoral fellows, and support staff working at the NCBI These scientists have backgrounds in medicine, molecular biology, biochemistry, genetics, biophysics, structural biology, computer science, and mathematics.

NCBI programs are divided into three areas: 1) building of new databases and enhancing existing ones which involve genomic information; this includes NLM-developed databases and extramural support for other research information resources; 2) basic research in computational molecular biology, and 3) dissemination and support of molecular biology databases and services. Within each of these areas, NCBI closely coordinates its activities with other NLM divisions and integrates information from key NLM databases such as MEDLINE into specialized data resources

Database Building and Enhancement

During 1992, responsibility for the NIH's GenBank Genetic Sequence Database was transferred from the National Institute of General Medical Sciences to NCBI. GenBank is an international database that collects all known DNA sequences and is a critical research tool in the analysis and discovery of gene function. NCBI will not only manage the database but will create new records from the journal literature and distribute the data to researchers A group of indexers in Library Operations with specialty training in molecular biology creates new sequence records from the journal literature These records are supplemented with the direct submissions from authors that are processed and reviewed by Los Alamos National Laboratory. The combined set of records was officially released as the NCBI-GenBank database on October 1, 1992 and is released every two months with daily updates.

Comprehensive coverage of all sequence data, protein as well as DNA, is provided along with the corresponding MEDLINE bibliographic information, including abstracts. At the NLM, more than 3600 journals are scanned for sequence data, and the NLM has expanded its journal coverage to include all journals that regularly contain sequence data even if they are in nonmedical domains, e.g , plant science. An integral component of the database is the inclusion of abstracts and indexing terms from the MEDLINE records of sequence-containing articles.

Specialized sequence indexers at NLM with advanced training in biological disciplines, are responsible for identifying and annotating sequence data from the MEDLINE literature The workflow involves.

- MeSH indexers identify articles containing sequence data as part of normal MEDLINE processing;
- Corresponding MEDLINE records are loaded into NCBI's relational database;
- Journal issues with sequences are scanned by sequence indexers and selected sequences are entered into the database with the addition of biological annotation, keyboarding of sequences is performed by contractor;
- Accuracy of biological information is reviewed by NCBI biologists,

This sequence database known as NCBI-GenBank is a key component in an integrated sequence database system that NCBI is developing as a single, comprehensive source of all known DNA and protein sequence information The purpose of the integrated database is to offer researchers the capability to perform seamless searching across all available data, including the journal abstracts linked to the sequence data.

Cooperative arrangements are being used to augment the in-house data capture operation. A collaboration with the National Agricultural Library as part of the Plant Genome Project furnishes coverage of the plant sequence literature. Similarly, an agreement with the U S. Patent and Trademark Office has been established to capture sequence information from issued patents. Finally, a series of R&D contracts were issued in FY 1992 for the express purpose of adding additional data to the sequence database

NCBI staff also are active in creating smaller-scale, special-purpose databases. These databases include transcription factors, and an integrated view of E. coli genetic map and sequence data. A new Expressed Sequence Tags (ESTs) database named 'dbEST' has been developed for publicly available collections of cDNA sequences. Participants submit data electronically to NCBI and are provided with access to the BLAST network service, plus software tools to assist in their EST analyses. Another new database, 'Contacts,' automatically notifies registered users of new MEDLINE citations that are related to a particular molecular biology citation of interest. It uses the clustering techniques developed for the *Entrez* system to identify groups of related documents based upon statistical text analysis. NCBI is collaborating with an international working group to design a structure and an ASN.1 definition for an integrated database of genetic and molecular data on the drosophila genome. This project received funding on July 1, 1992 from the NIH Human Genome program, and NCBI will continue to collaborate in an advisory capacity as this project moves into the implementation phase.

The NLM also provides funding and technical support to producers of molecular biology databases. The Brookhaven Protein Data Bank of 3-D protein structures and the CarbBank carbohydrate structure database are supported through interagency agreements with NSF and DOE, respectively. These databanks are major resources for research on the structure and function of large biomolecules. Other support for outside projects was provided through a Broad Agency Announcement for enhancements to NCBI's database projects. The announcement was issued in March and three awards were made in late September 1992. The awards support investigator-initiated projects to add specialized data and annotation such as vector sequences, repetitive DNA, and protein sequences from the back literature.

Software Toolkit

Equally important as building databases for molecular sequence information is the ability to access and retrieve the information using automated systems. The software toolkit concept addresses this need by focusing on the creation of software modules that provide a set of high-level functions to assist developers in building application software. Among these tools are a Portable Core Library of functions in the C language that facilitate writing software for different hardware platforms and operating systems, and AsnLib, a collection of routines for handling ASN.1 data and developing ASN.1 applications. The ASN.1 (Abstract Syntax Notation) tools use an International Standards Organization data description language to provide a mechanism for defining and structuring data as well as a set of program definitions which can interact with databases structured in ASN.1.

NCBI's adoption of ASN.1 for database output has several advantages for users as well as developers. The data definitions in ASN.1 for biological objects enables the representation and structuring of complex biological data in data files without the need for a specific database management system. Manipulation of the complex objects is performed through the ASN.1 software tools which are all in the public domain and freely distributed to the biology community. Thus, complicated analysis programs can be readily constructed from pre-existing sets of modular tools, saving considerable time and programming effort.

The major reason for adopting a standardized representation of biological objects is to facilitate the exchange of data; currently ASN.1 specifications have been developed for the major sequence databases (GenBank, PIR, and Swiss-Prot), the protein structure database at Brookhaven, the Protein Data Bank, and the OMIM database of Dr. Victor McKusick. ASN.1 is also being used as a distribution format for NCBI sequence and bibliographic data. The NCBI-Sequences CD-ROM contains sequence and MEDLINE data in the ASN.1 format to provide a structured version of the databases that are hardware- and software-independent.

User Retrieval Tools

A major application based upon the toolkit approach is a retrieval tool called *Entrez* that searches nucleotide and protein sequence databases and MEDLINE citations in which the sequences were published. With *Entrez* and a database on a CD-ROM or a local network, a user can rapidly search several hundred megabytes of sequence and literature data using techniques that are fast and intuitive. A key feature of the system is the concept of "neighboring," which permits a user to locate related references or sequences by asking for all papers or sequences that resemble a given paper or sequence. Neighbors are pre-computed using statistical algorithms developed at the NCBI. The ability to traverse the literature and molecular sequences via neighbors and links provides a very powerful yet intuitive way of accessing the data. The software also runs in a network environment and versions are available which allow a user (client) anywhere on the Internet to formulate queries on a local workstation and send off requests to a server at NCBI.

The system for building, maintaining, and performing quality control over the NCBI-GenBank database has been tested and refined with the production of five pre-releases of the *Entrez*: Sequences CD-ROM, NCBI-GenBank, which contains DNA sequence data in the traditional GenBank flat file format, and NCBI-Sequences, which contains an integrated sequence database and MEDLINE citations in ASN.1 format. The CD-ROMs have been sent to more than 1,000 users. The first official release is scheduled for October 15, 1992 and regular releases will follow every two

months. To date, more than 800 orders have been placed with the Government Printing Office.

NCBI is also using a modular approach in building systems for online retrieval. The BLAST sequence searching server has been implemented as a network-based retrieval system. Over 70 major sequencing centers and research institutions around the country have been provided with versions of this software that enables them to transmit a query sequence from their local computer over the Internet network to a BLAST server running on a computer at the NCBI. In a few seconds, the BLAST server executes the user's query and returns the results to the client program for viewing by the user. More than 1,200 search requests per day are processed, with typical response times of under 15 seconds. The BLAST network server has been recognized by the research groups as an essential laboratory tool not only to analyze data but to aid in setting directions for wet bench research. It has been a critical tool in the analyses of the several thousand human cDNAs sequenced by the Venter laboratory at NIH.

Two electronic mail servers were made available to the scientific community in August. The first, RETRIEVE, is used to retrieve records from several sequence databases, including GenBank, EMBL, SwissProt, and PIR, by sending a mail message containing the query to the server. The second, BLAST, is used to compare nucleic acid or protein sequences with the BLAST algorithm. Any user in the world with e-mail access can submit a query to the servers and have an answer returned within minutes. Over 1400 queries are handled daily.

NCBI-GenBank is also distributed over the networks through the standard File Transfer Protocol (FTP) program. Several hundred requests are processed daily for downloading files from NCBI's public data repository. The entire database of over 300 Mbytes as well as daily updates are available to network users.

Basic Research

Basic research is at the core of NCBI's mission. A group of multidisciplinary scientists work on fundamental biomedical problems at the molecular level using mathematical and computational methods. The basic research and database/software development have been found to be mutually reinforcing: the drive to make and substantiate new discoveries in biomedical science gives a rich stimulus, urgency, and direction to the development of new methods.

Although the expertise of the group is concentrated in sequence analysis, protein structure, and gene identification, research under way covers a wide range of topics in computational biology and information science, including mathematical models of evolution, statistical evaluations of large-scale drug trials, dynamical behavior of chemical reaction systems, and statistical text-retrieval algorithms. The basic research has strengthened the applications and database work by providing innovative algorithms and approaches that have been transformed into end-user software applications. Moreover, basic researchers in the group have made significant contributions in gene analysis over the past year by applying NCBI databases and software to sequence interpretation and by proposing experimental strategies for gene identification.

Algorithms, software tools, databases, and integrated systems developed at NCBI have had a wide-ranging impact on the advancement of biomedical research over the past two years. Major advances have resulted from, or have been assisted by, NCBI programs. For example, Dr. Francis Collins, discoverer of the cystic fibrosis and neurofibromatosis genes and Howard Hughes Investigator at the University of Michigan, used NCBI's BLAST program and sequence alignment tools to identify the biochemical function of neurofibromin, the first step toward developing drugs to treat "Elephant Man's disease" that afflicts one in 3000 Americans. Dr. Jeffery Gordon, Washington University, is using NCBI's alignment software to develop inhibitors of the AIDS virus. Dr. Bert Vogelstein, Johns Hopkins, used BLAST to analyze the gene responsible for colon cancer, the second most common cancer in the United States, with 180,000 new cases and 58,300 deaths projected for 1992. Dr. Frank McCormick, Chiron Corporation, is using NCBI's database and technology to study proteins implicated as causes of cancer and blindness.

In the genome research area, NCBI's experimental BLAST network service has become a critical component that laboratories rely on to rapidly interpret the vast quantities of data that are resulting from these projects. Genome research centers in Paris, Boston, Rockville, St. Louis, Houston, Utah, Oklahoma, and San Diego use BLAST to identify coding regions of genes and assign probable biological functions. NCBI's new database of "expressed sequence tags" has received enthusiastic support from plant genome researchers and is expected to have a major impact on agricultural biotechnology over the next several years.

The intramural group is engaged in more than 40 projects, many of which involve collaborations with NIH and other research laboratories. The work is reviewed by a Board of Scientific Counselors of distinguished extramural scientists (see Appendix 7 for list of members). The high caliber of the work has been evidenced by the number of peer-reviewed publications and the requests for outside collaborations. Examples of ongoing projects include: new statistical methods for AIDS vaccine testing that will minimize the use of rare and expensive animals necessary for the testing protocols; an automated method for modeling 3-D structures of RNA that appeared as a cover article for a special issue of Science on computational biology; identification of fossil sequences by comparing protein sequences from evolutionary distant organisms which predicts that today's databases already contain nearly 90% of all strongly con-

served sequences; sequence analysis combined with site-directed mutagenesis work to identify functional domains of a class of proteins (GTPase-activating proteins) that regulate cancer genes or have cancer-causing properties themselves; and ongoing collaborations with the French genome center, CEPH/Genethon, in identification of human disease genes including the Kallmann syndrome gene.

Communication

As part of its mandate to support the development of new information technologies of relevance to biology and genetics, the NCBI has exercised a leadership role in sponsoring forums for the exchange of information among leading scientists from the fields the computer science and biology. NCBI has also extended its outreach to the library science community by invited presentations and workshops on biotechnology information topics.

Workshops have been organized for software developers to have an opportunity to work with the NCBI software tools and the ASN 1 specifications for database objects. The first of the workshops was held in the fall of 1991 to assist academic and commercial software developers to make direct use of NCBI databases, especially CD-ROM versions. CD-ROMs of the NCBI-GenBank database began distribution in August through subscriptions to the Government Printing Office.

The Visitors Program, established last year, has been successful in bringing members of the scientific community to the NCBI to engage in collaborative research in the bioinformatics area as well as joint activities in database design and implementation. Selection of candidates is coordinated through the NLM's Extramural Program and involves a peer review of candidates' experience and proposed program. Over 30 visits from senior researchers were arranged in the second year of the program.

Information dissemination is also being provided by the Gen-Info Data Repository, a service for distributing software and databases produced by the NCBI as well as by outside groups. This repository has grown significantly and now contains 30 molecular biology databases that are freely available for network users through Internet FTP. A CD-ROM version with 20 databases was produced this year and distributed to 300 users. NCBI is continuing to work with Dr. Elvin Kabat to update and maintain the database of protein sequences of immunological interest, an important resource for immunologic researchers.

NCBI's expanded products and services has necessitated the establishment of a framework for producing and distributing information to the existing and potential user community as well as handling user service calls and producing user-oriented documentation. To this end, a program of participation in exhibits at major meetings of interest to molecular biologists was begun with an exhibit at the American Association of Microbiologists in May 1992 Staff of NCBI participated in the NLM's booth at the annual Medical Library Association meeting and presented a day-long tutorial on molecular biology and NCBI's services to a group of medical librarians and information specialists. Three issues of a newsletter were distributed to a mailing list of 13,000 biologists and institutions. On the NIH campus, 500 scientists are supported through online access to 20 databases under the IRX system.

The NCBI also participates in an advisory role with other government agencies such as the Patent and Trademark Office and the Department of Agriculture on programs involving biotechnology information. Special workshops on sequence analysis have been organized for the patent examiners at the U.S. Patent and Trademark Office who evaluate sequence data. Within the NIH, the NCBI coordinates with the National Institute for General Medical Science and the National Center for Human Genome Research on databases and informatics programs which impact information exchange on a national level

Extramural Programs

The NLM's Extramural Programs division has a program of grants for computer analysis of molecular biology data. Its scope is quite broad and includes research into methods and algorithms for improving the efficiency of information retrieval and improving the efficiency of analytical operations that are computationally intensive. Research applications to develop expert systems for annotating and linking databases are encouraged, as are proposals for work on algorithms for structure and function prediction Software development for newer machine architectures is within the scope of the program as well, including molecular analysis by neural net techniques, and multiprocessor programming Postdoctoral training in the cross-disciplinary areas of biology, medicine, and computer science is also supported through the NLM's informatics fellowship program.

Biotechnology Information in the Future

The design of molecular biology databanks and their resulting utility will, of necessity, need to follow scientific trends in research The stimulus to maintain state-of-the-art systems will come from an intramural and extramural program supporting scientific discovery

The NCBI will continue to develop software tools to assist researchers and software developers in gaining access to the growing volume of sequence and gene data and will encourage the widespread distribution of software and databases as essential components in the research process.

With the rapid advances of molecular biology research and information technologies, NCBI will be engaged in developing and employing new methods for disseminating knowledge to the biomedical community.

EXTRAMURAL PROGRAMS

Milton Corn, M D
Acting Associate Director

The Extramural Programs Division provides support to the health science community in the biomedical areas for which the National Library of Medicine takes particular responsibility NLM support for extramural programs stems from two sources from the Medical Library Assistance Act (MLAA) of 1965 and its extensions, and from Section 301 of the Public Health Service Act as amended Because MLAA is awaiting reauthorization by the Congress, EP's MLAA budget was kept at FY 1991 levels during the current year with resulting strain on the Division's ability to fund eligible grants appropriately The dual basis of the funding sources as well as the historic mission of the Library explain the eclectic variety of the funded projects for which the Division takes responsibility

Overview

The Extramural Programs Division is deeply involved in such major NLM themes as outreach, the High Performance Computer and Communications (HPCC) initiative, and biotechnology However, because much of EP's existing program structure is already related to outreach, HPCC, and biotechnology-related activities, new funds provided to support such areas are included among the "traditional" programs described in specific sections below rather than in an artificial separate listing

As developed by the Board of Regents' Planning Panel, the outreach program involves a number of EP initiatives, including professional training, IAIMS, and resource grants to improve access to biomedical information by health libraries and physicians

FY 1992 was the first year in which significant amounts of new funds were assigned to EP for use in HPCC programs EP allocated the funds to two distinct areas an Internet Connections program, considered as a resource grant and administered for NLM by the connections program already in operation at the National Science Foundation, and an HPCC demonstration project competition restricted to IAIMS institutions

Biotechnology is supported with both resource grants and research grants which are reviewed by a specific subcommittee of the BLRC

An important goal in FY 1992 concerned an ongoing effort to broaden support for informatics by encouraging other NIH institutes and other agencies to fund informatics-related projects, and by cooperation between NLM's Extramural Programs Division and other organizations Examples include support from the National Cancer Institute for additional training slots in EP's institutional grants; agreement with the National Center for Human Genome Research about respective areas of interest in biotechnological informatics, cooperative efforts with the National Science Foundation in the Internet Connections program; and in a database support program which may be useful in supporting some UMLS-related research EP was also successful in persuading such other institutes as NHLBI and NCI to fund projects originally assigned to NLM, which had high priority scores and informatics themes of interest to the other Institute EP also cooperated with the Radiation Oncology Division of the NCI in sponsoring a conference on computer imaging

Although budget restrictions, including a mandated cutback in travel funds, were experienced by all divisions, the relative lack of travel funds caused particular hardship for EP's vital site visit program Critical visits to evaluate applications were made, albeit with minimal staff attendance However, judiciously selected trips are also of great importance for evaluation of ongoing grants, of prospective grantees, and for promulgating the activities of NLM's grant programs, such trips were significantly curtailed

The Research Grants Section of this report summarizes some recent activities in the area of basic and applied information science The application of computers to biomedical information storage and retrieval has revolutionized the operations of biomedical libraries and has engendered the useful term, "medical informatics," to describe the theory and practice of providing information and decision support accurately and usefully to health workers Such research is vital now when the volume of biomedical information is growing at a rate that threatens our ability to keep track of what we know, and to use what we do know most efficiently.

Training efforts also merit specific description Training of competent professionals in medical informatics must remain an important goal of the Division. This new field needs scientists who can exploit the enormous potential for improvement in health delivery which medical informatics is capable of providing Merging information science with the peculiar complexities of modern health care and research poses complex problems whose solution will depend on well-trained specialists. NLM supports both institutional training programs and a fellowship program

Of particular importance during FY 1992 was the completion of a new competition for institutional training grants which resulted in 5-year awards to ten universities to support training in informatics research. Also new in FY 1992 was a fellowship in applied informatics, open to most health care professionals, and designed to promote training of experts in applying informatics to the clinical, educational, research, and administrative problems of health centers

The NLM's Integrated Advanced Information Management Systems (IAIMS) program addresses the insufficiently

appreciated but vital issue of integrating usefully the myriad information systems which have sprung up at most of our medical centers. Such databases are useful, to be sure, but all too often are unrelated, isolated, and very far from taking advantage of the synergism that could be realized by linking the various academic information systems present in our health centers, including the library, research material, academic administration information, medical education, and hospital information systems.

To respond appropriately to NLM experience and changing biomedical culture/technology during the past ten years, IAIMS was extensively revised during FY 1992 with important changes in scope, objectives, and funding levels. Because of scheduling problems and the proven expertise of IAIMS institutions, the FY 1992 HPCC demonstration grant competition was restricted to IAIMS grantees, past and present. (Should additional funds become available in future years for demonstration projects, eligibility for application will probably be extended to all health institutions.)

Medical Library Resource Grants have been an essential element of the Division's activities for years. It is clearly an NLM mission to make biomedical information easily available to health professionals. This emphasis was heightened when the NLM adopted outreach as a major new initiative. Improvement of access by physicians to medical information was specifically addressed by a recent revision of the Resource Grant Program designed to expand the ability of hospital libraries, particularly in rural or underserved areas, to establish facile contact with the national biomedical library system.

A revised program announcement describing Resource Grants was issued in 1992 to emphasize the current NLM policy of promoting access to national networks as the cornerstone of biomedical information management in the future.

Grants in support of publications have little to do with medical informatics but are a time-honored, important commitment by the Division to the scholarly activities which lie at the heart of libraries everywhere.

The support provided for the bioethics center is self-explanatory, as is the section on the Division's committee activities, and on the conferences supported.

Budget information is summarized in Table 10.

Regional Medical Library support, as authorized by the Medical Library Assistance Act, is described in the chapter on Library Operations. The Special Foreign Currency Program, administered by the Extramural Programs' International Programs Branch, is described elsewhere under International Programs.

Training

Research issues in the health information and health computer sciences call for highly trained creative scientists, able to link medicine with computers, and health care with information science. There is a particular need in academic medicine for the discipline of medical informatics. Through its research training program, NLM provides grants for research career training in this field.

Late in 1991, a Request for Applications (RFA) was issued and 24 responses were received. The primary intent of this program, as stated in the RFA, is the preparation of individuals for an academic career in Medical Informatics. Ten grants were awarded in FY 1992; because some programs are multi-institutional, 17 institutions in all are actively participating in the training program. These 10 grants provide the support for 60 postdoctoral trainees and 29 predoctoral students. In addition to these full-time trainees, the grants offer short-term training opportunities for students in medical or other professional schools of the health sciences. Such training is intended to give students summer or "off-quarter" opportunities for involvement in research so that they might consider medical informatics as a career option. Opportunities for 13 such individuals are included in the grants.

The National Cancer Institute has expressed an interest in increasing the pool of individuals who can deal with the application of medical informatics to the problems of cancer diagnosis and therapy. As evidence of this interest, NCI transferred to the NLM program sufficient funds to support four postdoctoral trainees.

Each site offers an excellent setting for didactic instruction, involvement in major ongoing computer science studies, and opportunities for work in advanced information science research. Subsequent to their training period, these newly trained investigators will contribute to the growth of science by their studies of the role of knowledge in professional life, by analyses of the social structures for managing knowledge, and by advancing the frontiers of the computer sciences for organizing, retrieving, and utilizing health knowledge.

The 10 training grant awards were issued to:

Paul D. Clayton, Ph.D.
Columbia-Presbyterian Medical Center

Charles P. Friedman, Ph.D.
University of North Carolina, Chapel Hill/
Duke University

Lael C. Gatewood, Ph.D.
University of Minnesota/Mayo Medical School

G. Anthony Gorry, Ph.D.
Rice University/Baylor College of Medicine

Robert A. Greenes, M.D., Ph.D.
Harvard Medical School/Brigham and Women's
Hospital/Massachusetts General Hospital/
New England Medical Center/Massachusetts
Institute of Technology

Perry L. Miller, M.D., Ph.D.
Yale University School of Medicine

Randolph A. Miller, M.D., Ph.D.
University of Pittsburgh

Joyce A. Mitchell, Ph.D.
University of Missouri-Columbia

Edward Shortliffe, M.D., Ph.D.
Stanford University Medical Center

Kent A. Spackman, M.D., Ph.D.
The Oregon Health Sciences University

In addition to these 10 institutional grants, the NLM supports an individual fellowship program. In 1992, the goals of the fellowship are the same as for the institutional training grants, i.e., to produce people for an academic career in medical informatics. In FY 1992, 11 individual fellowships awards were made. These fellows are receiving their training at Columbia University, Ohio State University, University of Utah, New York University, University of Washington, University of California-San Francisco, Sante Fe Institute, University of Pittsburgh, and the University of Illinois-Chicago.

Research Grants

The Library continues its support of research in the areas of medical informatics, biotechnology informatics, and health library information science.

Medical Informatics

NLM was able to sponsor only a small number of additional new projects this year. EP made four new traditional R01s, two FIRST awards to new investigators, and one large resource award (P41). The field, however, is beginning to attract the interest of other institutes at NIH. Several investigators, who have received earlier NLM support, are funded by the National Cancer Institute, the National Heart, Lung and Blood Institute, and the Agency for Health Care Policy and Research.

An example of a supported project is one at the Massachusetts General Hospital (MGH) entitled "Problem-Based Knowledge Access." Its aim is to develop computer-based clinical support systems for organizing and accessing problem-based knowledge in clinical settings. The intent is to improve software interfaces for rapid and intuitive query formulation. This project requires the development of powerful, but convenient, indexing and retrieval techniques. A related aim is to refine methodologies that facilitate the role of subject matter experts in the synthesis of new knowledge bases. To evaluate the effectiveness of these innovative methodologies, prototypic systems in three clinical domains will be developed and evaluated. The one most fully developed, pulmonary artery catheter management, is installed at the MGH and undergoing beta testing. Early reports are that the physicians and staff of the ICU would like to have it installed permanently.

With the costs of electronic storage decreasing, large databases containing a variety of images are becoming prevalent. One of the vexing issues facing the image database managers is how to index the individual images for rapid and accurate retrieval. Departments of radiology, pathology, and anatomy are particularly concerned with this problem. Dr. Carl Jaffe, Yale University, has found that geometry forms an inevitable part of almost all queries addressed to an image database. He and his colleagues are developing a robust geometric reasoning environment for medical image databases which will enhance the retrieval of particular items.

The National Library of Medicine believes it is important to support younger researchers getting started in their academic careers. For example, Dr. Michael Lincoln of the University of Utah received such support in FY 1992. His research involves finding new means or tools for the development of medical knowledge by "knowledge engineers." The intent is to improve reliability and validity of medical knowledge by making expert judgement more consistent, objective, and accessible.

Biotechnology Informatics

Four new R01s were funded in this category. Recent improvements in computer algorithms for comparing DNA and protein sequences have dramatically decreased the amount of time required to compare an unidentified sequence to a DNA or protein sequence library. Time and computational expense are no longer significant factors in sequence comparison. Dr. William Pearson, University of Virginia, is therefore focusing his attention on improving the quality of comparisons. This will be accomplished by examining approaches to improve the sensitivity, selectivity, and amount of information that can be inferred from similarity scores.

An example of a younger investigator in the biotechnology informatics area is Dr. Michael Mavrovouniotis, University of Maryland. He plans to develop a qualitative approach, using appropriate algorithms, that will help point the way to specific experimental studies to test theoretical predictions. Aims include modeling metabolic pathways that carry out a specific metabolic function and consideration of how defects in given enzymes alter flow and pattern of metabolites. The relevance of these aims is in regard to genetic defects in human metabolism.

Medical Library and Information Science

In this program area, NLM supports research projects concerning medical bibliography. More effective ways of organizing bibliographic databases and retrieving from them are a continuing program interest. The analysis of the

literature to yield new knowledge is emerging as a useful methodology for assessing the implications of new diagnostic and treatment modalities. One of these methodologies, meta-analysis, was the subject of several earlier NLM sponsored research grants.

Another question of interest concerns the users of medical information. How do health professionals perceive that they need information, and how do they go about satisfying this need? Once they complete a bibliographic search, what do they do with articles they find? How does bibliography really fit into their professional lives? Although there have been earlier studies of information-related behaviors (some funded by this program), it seems appropriate to seek new investigative approaches in multidisciplinary endeavors.

The single grant awarded in this area, to Professor Amy Warner of the University of Michigan, involves mapping synonymous and quasi-synonymous expressions in the MEDLINE bibliographic database. The goal is to discover algorithms that will perform this mapping of variant expressions automatically. If this experiment succeeds, users will be able to cope better with variant expressions or "linguistic synonymy," a frequent barrier to successful searches.

Conferences

To foster multidisciplinary interest and dialogue in biomedical knowledge issues, Extramural Programs supported two conferences this year. In both cases, other programs joined the Library to share sponsorship responsibilities.

On November 13-14, 1991, NLM and the National Science Foundation sponsored a workshop on "Creating An Infrastructure For Intelligent Systems In Molecular Biology." Forty-three participants met at NLM to discuss cooperative effort between computational science and molecular biology. Significant areas of scientific interest to both disciplines include datasets derived from intelligent systems, and data exchange and analysis by high performance computing networks. The workshop papers and the discussions they stimulated have been reported at two national conferences and will figure prominently in a forthcoming book on artificial intelligence and molecular biology. A follow-up conference is planned for 1993. Plans are also under way for a summer school to orient the different specialists to each other's methodologies. The school will also introduce students to an area where both molecular biologists and computer scientists find common interests.

On March 26-27, NLM jointly sponsored a symposium on "Medical Imaging Databases" with the National Cancer Institute and the Agency for Health Care Policy and Research. There were approximately 50 participants. As at the earlier workshop, NLM provided meeting spaces and supportive services; Extramural Program staff coordinated NLM's participation. The symposium discussions and conclusions will be reported in a forthcoming journal article and will be reported at several scientific meetings.

These small conferences, while demanding of staff time and resources, are invaluable sources of information for program planning and development for Extramural Programs. They identify biomedical problem areas where advances in computer and information science could lead to resolution or improvement. They provide a peer forum where the current state of the art is examined, and likely directions for new research are identified. The meeting spaces and the supportive resources of NLM's Lister Hill Center facilitate these discussions; they offer a high level of support at relatively modest cost.

Bioethics

Support for the National Reference Center for Bioethics at Georgetown University was continued this year, as in past years, by a specialized center grant. The collection of material which this grant makes possible was indexed and made available through NLM's BIOETHICSLINE bibliographic database. The indexing activity is supported by contract with NLM's Library Operations.

The material in this collection is devoted to a single subject, bioethics, and ranges widely because papers significant for this field often appear in literature other than biomedical. With NLM's support, the Center has also initiated an archive activity. The National Center for Human Genome Research awarded the Center a grant this year to expand this activity to include documents related to ethical issues concerning research on the human genome.

Resource Grants

NLM continued to support electronic access to information resources and services by means of the Information Access and Information Systems Grants. In FY 1992, 10 grants were awarded, and some 42 additional institutions were beneficiaries through consortium projects. Information Access Grants were awarded to two single institutions: Pennsylvania College of Podiatric Medicine in Philadelphia to create an online public access catalog to be connected to a local area network for greater access; and to the Whitman Hospital and Medical Center in Colfax, Washington, to provide Grateful Med access to this rural health care facility of 48 beds. Information Access Grants were also awarded to 4 consortia located in Billings, Montana; Missoula, Montana; Morehead, Kentucky; and Cumberland, Maryland. Some 46 institutions benefited from these consortium awards to provide Grateful Med, DOCLINE, and Loansome Doc capabilities.

Four Information Systems Grants were funded: 1) Gary Freiberger of the University of Maryland at Baltimore to develop a UMLS interface to the Health Sciences Library's

bibliographic databases; 2) Carol H. Fenichel of Hahnemann University in Philadelphia to develop a research-oriented information service containing such specialized information as internal research resources, GenBank and similar gene-sequence databases, online "NIH Guide to Grants and Contracts," bulletin board capabilities for researchers, and a calendar of research-related meetings, programs, etc.; 3) Janet M. Bostrom of Stanford University Hospital to develop an information system for transferring nursing research outcomes to practice settings; and 4) Michele Klein of Children's Hospital of Michigan in Detroit to interface the hospital's information system with literature, drug information, and some full texts.

As part of NLM's HPCC activities, EP supervised an interagency agreement with the National Science Foundation to award $395,000 for 10 grants to health science institutions to acquire basic connections to the National Research and Educational Network (NREN). Five institutions received approximately $30,000 each to link up to the Internet directly or through the main campus: Geisinger Hospital System (Danville, Pa.), Morehouse School of Medicine (Atlanta, Ga.), Kaiser Foundation Research Institute (Oakland, Ca.), Lovelace Medical Foundation (Albuquerque, NM), and the University of Connecticut (Farmington). Another five were granted approximately $50,000 each to extend an existing connection to outlying sites such as an academic health sciences center providing linkages to its affilited hospitals; the recipients were the University of Arizona (Tucson), Hahnemann University (Philadelphia), Washington University (St. Louis), University of Utah (Salt Lake City), and the Medical College of Wisconsin (Milwaukee).

Publication Grants

The Publication Grant Program provides selective short-term financial support for not-for-profit, biomedical scientific publications. Studies prepared or published under this NLM program include critical reviews or monographs on special areas of medical research and practice; secondary literature tools (such as atlases and catalogs); research monographs in the history of medicine; publications on medical informatics, health information science and biotechnology; pilot or temporary support for secondary periodicals; and the proceedings of scientifically significant symposia related to U.S. health needs. Because funds for publication support have dwindled, available resources in recent years have been used principally for history of medicine projects. The Publication Grant Program is supplemented by NLM's Special Foreign Currency Program, authorized under Public Law 480. (The Special Foreign Currency Program is described in the annual report's chapter on International Programs.)

During FY 1992 NLM awarded 10 Publication Grants totaling $283,000. Of these, seven were new awards. This small grant program has a current self-imposed annual ceiling on direct costs per grant of $25,000. The average grant awarded in FY 1991, including both direct and indirect costs, was under $23,000.

Among the new awards made in FY 1992 was a one-year award to complete the research and manuscript for a history of neurology in the United States between 1863 and 1945 by Dr. Bonnie E. Blustein. Her study emphasizes what neurologists did, their organizations, the development of neurology as a clinical specialty and its shift into the university milieu, and the divergence of neurology from psychiatry, neurosurgery, and internal medicine—rather than being a study of the growth of knowledge and concepts. The research will be one of the first case studies of a clinical specialty that address important issues in the development and institutionalization of scientific knowledge in a particular social context.

Another new award was to Dr. Michael J. Donnelly for a two-year study of the contributions of medicine to early social statistics. Recognizing that physicians played a preeminent role in the development of population-based social analyses, he will examine the rise of "moral statistics"—data on births and deaths, insanity, suicide, and poverty. The study promises to elucidate a series of critical questions, not only for the history of social medicine, but for the history of the relationship of medicine to social science and social epidemiology.

Among the books published in FY 1992 funded through the Publication Grant Program in previous years was Drs. Dan L. Lindsley and Georgianna G. Zimm's *The Genome of Drosophila melanogaster*. The work is a revision of an earlier catalog of mutations and chromosome rearrangements of Drosophila, and contains some 4,000 genes and 9,000 chromosomes. It is 1,133 pages and includes cyto genetic maps. (A list of supported publications received in FY 1992 is in Appendix 3.)

IAIMS

Integrated Advanced Information Management Systems (IAIMS) are institution-wide computer networks that link and relate library systems with a variety of individual and institutional databases and information files for patient care, research, education, and administration. Resource grants have been made to assist medical centers and health science institutions in planning and development projects that will lead to the implementation of IAIMS. The goal is to create organizational mechanisms within health institutions to manage more effectively the knowledge of medicine, and to provide for a system of comprehensive information access.

In FY 1992 NLM announced a revision of the existing program, first announced in 1982, with 1) a name change of replacing "academic" with "advanced" to recognize the wide applicability of the IAIMS concept and the need to incorporate new technologies, 2) retaining an initial planning phase, but fusing the old second and third phases (modeling and implementation) into one new operational phase, 3) changes in the level of support, to provide for up to $150,000 per year for one to two years for planning, and up to $500,000 per year for one to five years for operations, 4) some modification of the scope and conditions of the grant that include, among other things, the flexibility to have some operational elements introduced during planning, a clear relationship to clinical aspects of the health sciences, and incorporation of current NLM objectives, such as direct access to MEDLINE or extensive use of Grateful Med, and 5) incorporation of the High Performance Computing and Communications initiative into IAIMS planning and operations.

The revised program also includes the option in the operational phase to have an IAIMS apprenticeship through a position entitled "IAIMS Assistant." A structural learning experience must be provided. An additional $50,000 for each year can be requested to support salary, travel and supply costs.

Response to the announcement has been very high with more than 200 inquiries having been received in the first month following publication. The majority of inquirers have represented individual hospitals or associations of hospitals. The first applications to the revised program are expected in FY 1993.

During FY 1992 continuation awards were made to Baylor College of Medicine, Columbia University, Duke University, and Georgetown University for IAIMS implementation, and to the American College of Obstetricians and Gynecologists for IAIMS model development (a contract to the Oregon Health Sciences University for IAIMS implementation also remained active). Three institutions had been approved for the old Phase II modeling award, but the revision of the IAIMS program eliminated such support. Consequently, NLM awarded the University of Michigan, Yale University, and the University of Pittsburgh a one-year "transition" award to enable development of a new competitive grant application for the revised program and to complete current activity.

HPCC and IAIMS

NLM is the lead biomedical organization in the Federal Government's High Performance Computing and Communications (HPCC) initiative. There is a section on HPCC in the chapter on Special Initiatives. For the IAIMS Program, certain aspects of HPCC are readily adaptable, and others are clearly a natural step in evolution. Accordingly, in mid-year FY 1992, when limited funds became available to support the first HPCC grants, and in recognition of the requirement to make awards in FY 1992, NLM issued a one-time program announcement to 21 current and former IAIMS grantees for HPCC demonstration grant applications to investigate ways of solving significant problems in the health sciences. A partial list of possible project areas was identified that include collaboration enabled by high speed networks, distance learning, solution of computationally intensive problems in molecular biology in a distributed environment, visualization techniques, and network-based digital imaging.

Twenty-one institutions received the announcement. Fourteen applications were received and reviewed, and eight received priority scores between 100 and 250. From these, three were selected for funding: Dr. Edward L. Chaney, University of North Carolina, for a demonstration of using HPCC resources in radiation therapy entitled, "HPCC technology for real time decision support"; Dr. Walter B. Panko, University of Michigan, for a demonstration entitled, "A network-base image repository for biomedical researchers" in the rapidly developing area of confocal microscopy; and Dr. Gordon K. Springer, University of Missouri-Columbia, for a project to demonstrate the value of high performance computing, innovative software, and high speed communications entitled, "Integrating the tools for molecular biology."

Minority Support Activities

EP provides additional funds to resource grant P.I.'s who wish to train minority students. EP also strongly encourages resource grant applications from health institutions that serve minority populations. When resources permit, EP also provides additional funds to R01 P.I.'s who wish to add a minority scientist to the project; one such was funded in FY 1992.

Highlights of Committee Activities

NLM's scientific merit peer review group, the Biomedical Library Review Committee (BLRC) met three times in FY 1992 and reviewed 122 applications; 85 were recommended for further consideration. The Committee operates as a "flexible" review group; i.e., it is composed of three standing subcommittees, consisting of seven members each: Medical Library Resource Subcommittee, Medical Informatics Subcommittee, and Biotechnology Information Subcommittee.

A final peer review of applications is performed by the Board of Regents, which meets three times a year, approximately three months after the Biomedical Library Review

Committee. One of the Board's subcommittees, the Extramural Programs Subcommittee, meets the day before the full Board for the review of "special" grant applications. Examples of "specials" include applications for which the recommended amount of financial support is larger than some predetermined amount, when at least two members of the scientific merit review group dissented from the majority, when a policy issue is identified, and when an application is from a foreign institution. The Extramural Programs Subcommittee makes recommendations to the full Board which votes on the applications.

Plans for FY 1993

Internal reorganization of EP is needed.

Detailed grant-funding plans depend on the actual amounts made available by Congress. In general, all of the existing grant programs will be continued.

High Performance Computing and Communication funds, if allocated to EP, will be distributed among research, training, and resource programs to help develop the various elements needed for useful application of HPCC to biomedical needs.

Table 10
Extramural Grant and Contract Program
(dollars in thousands)

Category	*FY 1990*		*FY 1991*		*FY 1992*	
	No.	*$*	*No.*	*$*	*No.*	*$*
Research	47	11,343	49	11,231	51	11,863
Resource projects	15	4,492	17	4,721	14	4,802
Resource access	15	519	17	684	7	493
Training	7	2,886	7	2,714	10	3,520
Fellowships	5	189	9	306	9	298
Regional Medical Libraries	7	3,772	8	5,500	8	5,482
Publications	17	390	13	365	10	283
(IAIMS projects)	(8)	(3,821)	(9)	(3,693)	(3)	(3,938)
(Med. info. research)	(24)	(6,030)	(26)	(6,066)	(29)	(6,894)
(Biotech. research)	(17)	(4,179)	(16)	(4,192)	(13)	(3,759)
Totals:	113	$23,591	120	$25,521	109	$26,741

OFFICE OF COMPUTER AND COMMUNICATION SYSTEMS

Harry D. Bennett
Acting Director

The Office of Computer and Communication Systems (OCCS) provides information processing capabilities to meet NLM needs and, in so doing, determines and meets the data processing and data communication requirements for: 1) disseminating biomedical information to thousands of institutional and individual health professionals around the nation and the world; 2) operating the world's largest library in a single technical area—biomedicine; and 3) providing Management Information System (MIS) services, including office automation, to NLM.

OCCS: 1) implements computer and communication systems using state-of-the-art technology and techniques; 2) analyzes, plans, and provides real-time, on-line, around-the-clock information services for increasingly sophisticated users; 3) schedules and controls maintenance and publication of dozens of databases, each measured in billions of bytes (characters); 4) operates a modern computer center; 5) conducts performance measurement and capacity planning for computer hardware, operating systems, database management systems, and transaction processors; and 6) produces and distributes data and software products to thousands of institutions and health professionals. The organization of OCCS is a direct reflection of these responsibilities. Computer and communication systems are:

- developed and implemented by the Development Branch
- enhanced and maintained by the Application Services Branch
- executed on computers under operating system control of the Systems Support Branch
- provided as an around-the-clock service by the Computer Services Branch.

Development Branch

The Development Branch is responsible for analyzing, designing, and implementing computer-based systems to support NLM's requirements. Development activities during the past year included the implementation of a major new release for the Technical Services System (TESS), implementation of a new release of Grateful Med and expansion of Grateful Med services, extensions to the local area network services, developing initial Information Systems Laboratory (ISL) projects, and developing a new Online Public Access Catalog (OPAC).

The Technical Services System (TESS), under development since 1989 to integrate various functions of the Technical Services Division, is a distributed processing system that integrates mainframe computer, personal computer, local area network (LAN), and database technologies.

The initial implementation under TESS provided for the creation and maintenance of original cataloging. The next major release of TESS provided the framework for the integration of acquisition and cataloging activities (1990). In early FY 1992, subject and name authority control for the cataloging function was integrated into TESS. This release also included the capability for creation and maintenance of the Name Authority File.

During 1992, the principal effort on TESS has been directed toward the conversion of the entire CATLINE/AVLINE files into the TESS environment. TESS will be enhanced to support all file maintenance activities of CATLINE/AVLINE, and to support all cataloging distribution activities. This release is scheduled for implementation in spring 1993.

The scope and coverage of Grateful Med continue to grow. A new version for the PC (version 6.0) and an Apple Macintosh update were released in 1992. PC Version 6.0 incorporated new user communications capabilities including TCP/IP access via the Internet, and Local Area Network (LAN) access via Novell software. Grateful Med can now provide broader access to the NLM with the addition of new communication scripts. For example, U.K. access using Janet/Internet was made available in the fall of 1992. More than 2.5 million searches were performed via Grateful Med this year by over 36,000 registered owners of the software.

Loansome Doc provides a link to DOCLINE enabling the user to electronically order journal articles found during a Grateful Med search. These articles are ordered from an affiliated library. Loansome Doc became fully operational in 1992. It has been integrated into Grateful Med Version 6.0 and is now distributed to all Grateful Med users. Over 52,000 document requests were submitted by over 2,100 users since the project was begun. The reports from those health professionals who have ordered documents using Loansome Doc continue to be positive about this new capability.

A collection of Local Area Networks (LANs) form the internal communications systems of NLM. Together, these various LANs support approximately 700 NLM staff, con-

tractors, and visiting patrons They provide access to all of NLM's data processing resources as well as access to external computer networks and data systems Access is provided not only to the MEDLARS system, but also to file servers, minicomputers, and other systems used for library operations, office automation, research and development There is little doubt that external and internal demands on NLM's computer and communications resources will continue to grow in the future

In order to accommodate the requirements for growth, higher speeds, better reliability, and more efficient sharing of resources, NLM has begun to upgrade its LAN systems to incorporate newer technology NLM is currently installing a high-speed backbone network that will interconnect medium speed Ethernet subnets and thus provide a single, consolidated architectural approach to its LAN needs The high-speed backbone will also provide connections to other high-speed networks such as the NIH RESNet and, eventually, NREN During 1992, efforts began to transition computer communications systems from the NLM broadband system to Ethernet systems A second direct NLM T1 connection was also established to the Internet via SURANet to better support Internet access to MEDLARS and other LAN-connected resources Internet access was also provided to TOXNET

The Information Systems Laboratory (ISL) was created within the Development Branch in 1991 It is intended as a core facility to help OCCS modernize and enter the emerging technology domains of distributed processing, open systems, high-speed networks, and worldwide connectivity and service provision Although it was not fully staffed until June 1992, the ISL did distribute beta releases of its first software product (TC_COMM) in November 1991 TC_COMM is designed to free application developers from developing their own communications modules TC_COMM provides an Application Programmer's Interface (API) to a Berkeley sockets interface, the industry defacto standard for TCP/IP drivers The first formal release of TC_COMM was in September 1992

The ISL is introducing open systems computers and workstations to support operational requirements Various efforts are supporting the redesign of existing systems and the development of new systems which use multiplatform open system servers, TCP/IP communications and Internet connectivity These new systems should facilitate the provision of remote NLM services both domestically and internationally An additional goal of these new systems, for internal NLM processing requirements, is to provide Internet-based remote use and ultimately to provide better support for remote cataloging and indexing activities

In 1991 development started on a prototype Online Public Access Catalog (OPAC) system that utilizes existing NLM bibliographic databases and retrieval engines, such as ELHILL In FY 1992, the project became a full-scale development effort and substantial progress was made in creating the OPAC server software and the end user workstation software The OPAC server is based on an open systems design running on a Unix computer, the Reading Room workstations will be DOS PC's. The initial design goal was to develop a simple (to the end user), easy to use, intuitive system to support the first-time and infrequent visitor to the NLM Reading Room During FY 1992 a demonstration OPAC system was developed and testing was started by the Public Services and History of Medicine Divisions Later phases of the OPAC development will also support terminal emulation, as well as DOS software-based PC access via TCP/IP communications over the NLM LAN, NIH Campus networks, and the Internet

Application Services Branch

The Application Services Branch (ASB) supports the various NLM programs and serves as the nucleus of all automated programming support services

- In FY 1992, the DOCLINE Interlibrary Loan (ILL) system, NLM's online facility for requesting library materials was converted to a Virtual Telecommunications Access Facility (VTAM) environment Users realize a much better response time and also use the more powerful IBM/3090 mainframe computer installed this past year The hours of system availability were increased and the number of possible users was increased from 2,500 to 5,000 as a result of these changes Software was written to allow Clinical Alert data to be ordered free from NLM New versions of the software, including report software about journal articles requested by the network and general statistical reports about Loansome Doc activity were created and put in use

New generations of software for subsystems of the Automated Indexing Management System (AIMS) became operational in FY 1992 AIMS is an IBM mainframe computer application that runs under the Customer Information Control System (CICS) It provides access to the Inquire Data Base Management System (DBMS) for storing and retrieving new records and maintaining old records Subsystems affected were 1) Journal Control, 2) Indexing, 3) Checkin; 4) Bibliographic Processing, 5) Binding, and 6) Capping Also, new generations of software for the In Process (INPROC) and Literature Selection (LSTRC) CICS systems were created

Many software enhancements were made to the Model 204 DBMS- based MeSH system This system provides for data entry, verification, and validation of the NLM controlled thesaurus, MeSH Data are extracted daily from the Model 204 MeSH database and updated to the

MEDLARS information retrieval MeSH database. These thesaurus data are used to formulate searches of the MEDLARS online databases.

As planned, the Bioethic Citation Maintenance System (BCMS) became fully operational in September 1991. This software is PC based and provides for individual citation maintenance and new record creation of Bioethics related data. Providing this capability allowed NLM to format data as input to the annual Bibliography of Bioethics publication in 1992. Data was sent to the Government Printing Office and the publication was created.

Again, in FY 1992 more than 10% of the NLM bibliographic data of almost 7 million records were class maintained. Class maintenance is the adding of new terms, deleting old terms and replacing terms with preferred ones in MEDLARS records. New data fields are introduced to the records as required. Major software enhancements to support the class maintenance effort were the processing of additional "Publication Type" data, mapping of MeSH to Chemical Names, and identifying inconsistent data between MeSH Chemical records and data carried in the MEDLINE family of files.

There are currently more than 63,500 users of NLM online services. Over 700 new codes are assigned by MEDLAR Management each month. Software was created to accommodate new fields resulting from use of the NLM Bulletin Board System (BBS), a facility for registering as a user of NLM online services.

Systems Support Branch

The Systems Support Branch is responsible for hardware analysis, system software, and data communications. The current NLM mainframe configuration is an IBM 3090-300J with MVS/ESA (multiple virtual systems/enterprise systems architecture).

During FY 1992 the NLM mainframe computer system was upgraded to provide increased capacity and improved reliability and availability. The Systems Support Branch provided planning, coordination, and software conversion for this major upgrade. New teleprocessing interfaces were developed and implemented for the information retrieval system (ELHILL) and interlibrary loan system (DOCLINE) which improved reliability and increased capacity for online users.

Accomplishments for FY 1992 include:

- Replacement of the IBM 3084-Q and IBM 3081-K mainframe with an IBM 3090-300J.
- Development and implementation of an improved teleprocessing interface for interlibrary loan (DOCLINE).
- TSO, PDQ, and DOCLINE access were converted to the resource access control facility (RACF). This improved security for online user access.
- Expanded mainframe interface to the Internet. A second gateway (IBM 3172) was installed to eliminate the single point of failure as well as provide increased capacity.
- The teleprocessing interface for ELHILL was rewritten as direct VTAM application. TCAM 3 was discontinued.
- Automated operations software was implemented to reduce manual tasks and expedite the resolution of problems by automatically calling appropriate personnel.
- Remote job entry (RJE) was converted from emulation program to the Network Control Program. This permits the phase-out of obsolete IBM 3705 communication controllers.
- Support was provided for more than 100 software products used by programmers, users and System Support staff.
- Support was provided for connecting new terminals and work-stations throughout the NLM.
- Development and distribution of procedures and status information for the mainframe system.
- Program changes to system software as required by NLM developers and users.
- Office automation support for personal computers and the PROFS calendaring and message system.

Computer Services Branch

The Computer Services Branch provides data processing services and support for subscribers and users of MEDLARS, DOCLINE and other databases through the use of a large mainframe computer system installed at the NLM.

The system now installed is an IBM 3090-300J with performance characteristics of processing 62 million instructions per second (MIPS) operating under MVS/ESA. In addition, the Computer Services Branch maintains an IBM 9370 in support of the Library's PROFS calendaring and message system. Operational staff support is provided on a 24-hour day, 6-day a week basis. Operator staffing is also provided for most Sundays and government holidays, as most subscribers, both domestic and foreign, continue to use the online system on those days.

The peripheral equipment attached to the IBM 3090 mainframe consists of 220 direct access storage devices (DASD) with a total online storage capacity of approximately 260 billion bytes or characters of data. In addition, subscriber support of requested database files is performed through the use of 14 magnetic tape and cartridge drives. Also installed are many telecommunication units to provide easy and quick access into the main MEDLARS and DOCLINE database files for worldwide use.

Printer output support exceeded 13 million pages or 679 million lines of output printed locally as well as over

remote printers. Both high-speed fan fold, and cut sheet laser printers as well as impact high speed printers are attached to the mainframe system. Thirty-five percent of this printout is subscriber requested for which they pay a charge under terms of their agreements with NLM.

During the past fiscal year the Computer Services Branch created and mailed out 6,500 magnetic tapes of MEDLARS and TOXNET database information and files to both domestic and international subscribers.

INTERNATIONAL PROGRAMS

Richard K. C. Hsieh, Dr. P.H.
Director, International Programs

During the past year, NLM continued its international collaboration with individual countries, international government organizations such as the World Health Organization (WHO) and the Pan American Health Organization (PAHO), and international nongovernmental organizations such as the International Council for Scientific and Technical Information (ICSTI). The Special Foreign Currency Program was active in the support of critical reviews and history of medicine projects. Other NLM international activities included training for colleagues from abroad, the NLM publication exchange program (with 169 institutions in 51 countries, including the U.S.), as well as receiving numerous professional visitors from abroad.

Collaboration with Individual Countries

A feasibility study with USAID support was initiated in Egypt with the Academy of Scientific Research and Technology to establish a National Library of Medicine. The goal of this project is to improve Egyptian medical information services. Egypt has an established International MEDLARS Center, but the library will be one more important step toward improved health information services in the country. This project will improve the collections in three major health libraries.

International MEDLARS Agreements

The Library has MEDLARS agreements with partners in 15 foreign countries and with two international organizations (Table 11).

The National Informatics Centre (NIC) in New Delhi, India has begun to provide MEDLARS search services for the public. In addition to their use of BRS search software and NLM tapes for the most recent five years (1988 to 1992), there is also online service to the NLM computers.

The New Zealand Ministry of Health and the Seoul National University Medical College in Korea are the two newest International MEDLARS Centers. Both Centers are still experimenting with Grateful Med as a means of access. Canada, China, the United Kingdom, Taiwan, and Sweden have also begun to distribute Grateful Med to their MEDLINE users.

NLM also has a MEDLARS agreement with the Pan American Health Organization (PAHO), intergovernmental health organization. In 1989, PAHO amended its leasing agreement with NLM to provide online access to MEDLARS databases from Argentina, Chile, Jamaica, and Costa Rica. In 1990, NLM continued a collaborative project with PAHO and the University of Chile in the improvement of a gateway system named BITNIS. This new system demonstrated the capability for health professionals to conduct MEDLINE searches from Argentina, Chile, Costa Rica, Mexico and Venezuela. In 1992, the BITNIS gateway software was ported into a Sun Spartan Workstation which includes many new features. A Beta test was conducted from June to October 1992 by 150 users in 12 participating institutions.

To use BITNIS, a MEDLINE search is initiated by using Grateful Med to send search commands to NLM through the Internet network. The search results obtained from the NLM computer are transmitted back to the originator through Internet; and Grateful Med is used again to edit and present the results. The objective of the BITNIS project is to provide NLM's MEDLINE to health professionals in all Latin American countries where the high cost of international communication services inhibits online access to the database.

We were pleased to receive a telefacsimile message from the Kuwait International MEDLARS Center in which the Center requested to become active again following the liberation of that country.

NLM also accepted the first International Associate into its 1992 domestic Associate Program. A Chinese applicant from the School of Library and Information Science, China Medical University was selected from a group of many qualified applicants.

Collaboration with the World Health Organization

The NLM and the World Health Organization continued to cooperate in the publication of the *Quarterly Bibliography of Major Tropical Diseases* and the *Bibliography of Acute Diarrhoeal Diseases*. NLM prepares camera-ready copy from the MEDLINE system, and WHO prints and distributes these to thousands of institutions in the developing countries.

NLM and WHO also continued a collaborative interlibrary loan arrangement in which photocopies of journal articles are provided to WHO-referred requestors at a reduced rate. Library resources in developing countries are usually insufficient and the need for biomedical and health information can be met only by drawing on the collections of the developed world. Even though NLM and WHO continue to provide some photocopies of journal articles to developing countries, this arrangement can only partially meet the demand. Unless other resources in developed countries can be found, the need for interlibrary loans to developing countries will continue to grow.

Table 11
International MEDLARS Centers

Tapes	*Tapes/Software*	*Online NLM*
France	Australia	PAHO*
Germany	China	Canada
Japan	Sweden	Egypt
PAHO (BIREME)*		France*
Switzerland*		India*
India*		Italy
		Korea
		Kuwait
		Mexico
		New Zealand
		South Africa
		Switzerland
		Taiwan
		United Kingdom

*Combined online/tapes

Special Foreign Currency Program

Authorized under Public Law 83-480, as amended, the Library's Special Foreign Currency Program utilizes U S -owned local foreign currencies to prepare and publish biomedical scientific publications for the health-science community This program, active since 1962, is the oldest of NLM's extramural support activities Although over the years NLM has sponsored collaborative PL-480 projects in seven countries, support is presently available only in India

During FY 1992, 19 projects totaling $330,000 (equivalent in foreign currency) were active in India About 15 percent supported the translation and publication of biomedical monographs and bibliographies by noted foreign scientists The remainder funded the translation and publication of major historical monographs These classics in the history of medicine are selected in collaboration with the American Association for the History of Medicine

Among the publications received in FY 1992 was a translation from the Russian of the autobiography of Nikolai I Pirogov, *Questions of Life Diary of an Old Physician*, edited and with a new introduction by Dr Galina V Zarechnak Pirogov is known as the "Father of Russian Surgery," and has been called one of the greatest of military surgeons Unlike Pirogov's numerous books on various fields of medicine and surgery, the *Diary* records his observations, reactions to events of those times and reflections on a multitude of questions such as preventive medicine, the psychological impact of war on soldiers, women's role in society and the wars, women's rights, the role of religion in the state, the origin of the universe, ethical problems of medicine and education, and the doctor-patient relationship. His "questions of life" focus on values and ideas which are as interesting today as when written more than 100 years ago

Also published in FY 1992 was the three-volume translation of Professor Eugen Korschelt's *Regeneration and Transplantation*, edited by Dr Bruce M Carlson Korschelt, a well-known German zoologist at the University of Marburg, which summarized what was known in the fields of regeneration and transplantation in the late 1920s The volume on regeneration describes a wide variety of regenerative phenomena through a large spectrum of animal species, plants, and even inorganic crystals Some of the topics discussed, e g , multiple regeneration, have been passed from the cataloging stage to the point where now many of the phenomena are dealt with in contemporary theoretical models Other topics, such as the origin of regenerating cells, are still the subject of debates similar to those carried out in the 1920's

Transplantation, the subject of the remaining volumes, is treated in the broad biological context of the pre-immunology era, with both a large section on transplantation in invertebrates and a volume on transplantation phenomena and techniques in vertebrates The latter range from the grafting of embryonic organ primordia to the autotransplantation of human skin and bones These volumes provide important historical coverage that will be of interest to scientists working on development, regeneration, and transplantation

International Meetings and Visitors

The Library is a member of the International Council for Scientific and Technical Information (ICSTI) This organization serves as a meeting ground for information and abstracting agencies, both commercial and governmental Common interests include economics of primary and secondary publications, transborder flow of information, electronic publication, standardization and the information needs of developing countries At the 1992 general meeting of ICSTI held in Berlin, NLM was represented by the NLM Deputy Director Kent Smith, who also assumed the position of ICSTI President

NLM organized a Workshop in International Collaboration in the Application of Medical Informatics with joint sponsorship of the Fogarty International Center, National Cancer Institute, and NLM The 1992 Workshop was held prior to the MEDINFO 92 meeting in Geneva, Switzerland, with emphasis on collaboration between newly democratized Central and Eastern European countries and the United States Major topics included were training health professionals in medical informatics, assisting health professional education through multimedia educational technologies, improving health professionals' access to published information, and strategies in evaluating medical informatics systems

The Library continues to attract many foreign visitors each year, including medical librarians, health professionals, and government officials. Many of these visitors have responsibility for medical, scientific or technical information in their own countries. Their interest in NLM is more than cursory, and they are officially received and briefed on relevant aspects of NLM operations and research. In 1990 visitors came from the following countries:

Argentina, Australia, Bahrain, Bangladesh, Belgium, Benin, Bolivia, Botswana, Brazil, Bulgaria, Burkina Faso, Central African Republic, Chile, China, Colombia, Congo, Croatia, Cuba, Czech and Slovak Federal Republic, Djibouti, Ecuador, Egypt, Estonia, Finland, Gabon, Germany, Ghana, Greece, Honduras, Hungary, India, Indonesia, Iran, Israel, Japan, Kenya, Kuwait, Lebanon, Madagascar, Malaysia, Mali, Mexico, Morocco, Nepal, New Zealand, Nigeria, Oman, Pakistan, Panama, Peru, Philippines, Poland, Qatar, Romania, Russia, Rwanda, Saudi Arabia, Singapore, South Africa, Spain, Sri Lanka, Sudan, Swaziland, Sweden, Switzerland, Taiwan, Thailand, The Netherlands, Uganda, United Arab Emirates, United Kingdom, Uruguay, Yugoslavia, Zambia, and Zimbabwe.

ADMINISTRATION

Kenneth G. Carney
Executive Officer

Financial Resources

In FY 1992, the Library had a total appropriation of $99,088,000. Table 12 displays the FY 1992 budget authority plus reimbursements from other agencies, and the allocation of these resources by program activity.

Table 12
Financial Resources and Allocations, FY 1992
(In Thousands of Dollars)

Budget Authority:	
Appropriation, NLM	$99,088
Plus: Reimbursements	15,864
Total	114,952
Budget Allocation:	
Extramural Programs	26,058
Intramural Programs	80,921
Library Operations	(49,342)
Lister Hill National Center Biomedical Communications	(15,242)
National Center for Biotechnology Information	(9,214)
Toxicology Information	(7,123)
Research Management and Support	7,973
Total	$114,952

Personnel

The NLM operated under an employment freeze imposed by the Public Health Service during the final weeks of the fiscal year. The NLM expects to close the fiscal year with 603 full time equivalents (FTEs), 10 less than its allocated ceiling.

The White House announced the appointment of NLM's Director, Donald A.B. Lindberg, M.D., as the first Director of the National Coordination Office for High Performance Computing and Communications. The announcement was made by Dr. Allen Bromley, Assistant to the President for Science and Technology Policy. Dr. Lindberg will hold both directorships concurrently.

This year, the NLM Director announced the creation of a new office within the Office of the Director. The Office of Health Information Programs Development (OHIPD) is under the direction of Elliot R. Siegel, Ph.D., who was appointed to the new SES position of Associate Director for Health Information Programs Development. Dr. Siegel will direct the efforts of the three units within OHIPD: the Office of Outreach Development, the Office of Planning and Analysis, and the Office of International Programs.

Henry M. Kissman, Ph.D., the NLM Associate Director for Specialized Information Services (SIS) since 1970, retired from the Federal service this year. The Director, NLM, selected Michael Ackerman, Ph.D., to be detailed from the Lister Hill Center to SIS to serve as Acting Associate Director while recruitment for the Senior Executive Service position is conducted.

Two Special Experts joined the NLM this year. Dr. Stephen Strickland was appointed to the NLM Visiting Historical Scholar Program. The Program brings recognized historical scholars to the Library's History of Medicine Division to engage in historical research, scholarly interchange, and staff consultation. Also appointed as a Special Expert, Ms. Karen Ginter joined Library Operations where she will coordinate programs that expand access to the NLM's online services. Ms. Ginter will serve as the focal point for new ventures that demonstrate the value and efficiency of online searching and the use of Greatful Med.

Eugene V. Koonin, Ph.D., was appointed as a Visiting Scientist with the National Center for Biotechnology Information. Dr. Koonin is a graduate of Moscow State University where he conducted research in the replication of genomic RNA in encephalomyocarditis virus. He has worked as a senior researcher at the Institute of Poliomyelitis and Viral Encephalitis of the USSR Academy of Medical Sciences, and at the Institute of Microbiology of the USSR Academy of Sciences.

The NLM staff were saddened by the unexpected death on May 16 of Charles A. Walker, Ph.D., Director of NLM's Office of Outreach Development. Dr. Walker had been Chancellor of the University of Arkansas at Pine Bluff before he joined the NLM staff in 1991.

Awards

The NLM Board of Regents Award for Scholarship or Technical Achievement was awarded to Karen J. Patrias, Library Operations, in recognition of scholarly achievement in developing NLM's recommended formats for bibliographic citations.

The Frank B. Rogers Award recognizes employees who have made significant contributions to the Library's fundamental operational programs and services. In 1992, William Willmering, Library Operations, received the award for his work in the development of automated systems that

improve the Library's ability to acquire and provide access to the biomedical serials literature.

The NIH Director's Award was presented to Bruno Vasta, Specialized Information Services, in recognition of his work in building and reconfiguring TOXNET into a microprocessor system providing valuable toxicological information to the biomedical community.

The NLM Director's Honor Award, presented in recognition of exceptional contributions to the NLM mission, was awarded to two NLM employees this year. Maria G. Farkas, Ph.D., Library and Operations, and Henry Kissman, Ph.D., Specialized Information Services (SIS). Dr. Farkas was recognized for her substantial contributions to the Library's mission through her indexing and linguistic skill. Dr. Kissman, the Associate Director, SIS, received the honor award in recognition of his outstanding leadership in developing the NLM's information program in toxicology.

This year, two employees were honored for their contributions in Equal Employment Opportunity (EEO). The NLM's EEO Special Achievement Award was presented to Cynthia B. Gaines, Specialized Information Services, for her leadership role in NLM's EEO program. The Philip Coleman Award was presented to Alvin Barnes, Library Operations, for his contributions in establishing a positive work environment at the NLM.

* * * * *

Table 13
Staff, FY 1992 Full-Time Equivalents

Program	*Full-Time Permanent*	*Other*
Office of the Director	20	2
Office of Public Information	5	1
Office of Administration	51	4
Office of Computer and Communications Systems	63	7
Extramural Programs	16	3
Lister Hill National Center for Biomedical Communications	73	8
National Center for Biotechnology Information	18	15
Specialized Information Services	34	4
Library Operations	249	30
TOTAL	529	74
TOTAL FTEs	603	

Equal Employment Opportunity

Fiscal Year 1992 was a period of much activity and accomplishment in the area of Equal Employment Opportunity. Among the events:

Nov 24 Two EEO awards were presented at the NLM annual awards ceremony: Cynthia B. Gaines, SIS, received the EEO Special Achievement Award; the Philip Coleman Award was presented to Alvin Barnes, LO

January Schomburg Exhibit—"African-Americans in Space Science" and "Countdown to Eternity"

Jan 29: Rep. Louis Stokes address honors Martin Luther King, Jr.

Feb 5: Lectures by Dr. Vanessa Gamble and Dr. Todd Savitt in observance of African-American History Month

March NLM exhibit at National Association For Equal Opportunity

Mar 20: Address by Senator Nancy Kassebaum in observance of Women's History Month

Apr 10: Visit of students from the Amidon Elementary School, Washington, D.C.

Apr 16: Student National Medical Association, New Orleans

Apr 30: NIH Human Resource Management Professional Development Conference, Gaithersburg, MD

May 19 MLA Annual Meeting, Capital Ideas Breakfast 2, "Health Issues of Race and Gender" Speakers Dr Elliot Siegel, NLM Associate Director, Health Information Programs Development, Dr John Ruffin, NIH Associate Director for Minority Programs, Dr Vivian Pinn, Director of NIH Office of Research on Women's Health

Jun 11 Career Day, sponsored by the NLM EEO Advisory Committee

Jun 12 Women in Biomedical Careers Dynamics of Change Strategies for 21st Century Workshop, sponsored by NIH Office of Research on Women's Health

Jun 29 NLM Cultural Diversity Seminar, Dr Edwin Nichols

Aug 24 Federal Dispute Resolution Conference, Atlanta

Sep 4 National Conference of African American Librarians, Columbus, OH

Oct 28 U S Patent and Trademark Office, National Disability Awareness Employment Month Program, met with Director, W H Williams, Office of Civil Rights

David Nash
EEO Officer

APPENDIX 1: ACRONYMS, ABBREVIATIONS, AND INITIALISMS

AAOS	American Academy of Orthopaedic Surgeons
AAT	Art and Architecture Thesaurus
ADA	Americans with Disabilities Act
AHA	American Hospital Association
AHCPR	Agency for Health Care Policy and Research
AI/COAG	Artificial intelligence hemostasis consultant system
AI/RHEUM	Artificial intelligence rheumatology consultant system
AIDSDRUGS	AIDS drugs
AIDSLINE	AIDS information onLINE
AIDSTRIALS	AIDS Clinical TRIALS
AIMS	Automated Indexing Management System
AKAT	Audio Knowledge Acquisition Tool
ANN	Artificial neural network
ANSWER	ATSDR/NLM's Workstation for Emergency Response
APDB	Audiovisual Program Development Branch
ARC	Annual Review of Carcinogens
ARL	Association of Research Libraries
ASB	Applications Services Branch
ASN	Abstract Syntax Notation
ATSDR	Agency for Toxic Substances and Disease Registry
AVLINE	AudioVisuals onLINE
BBS	Bulletin Board System
BCMS	Bioethic Citation Maintenance System
BDIP	Biomedical Digital Image Processing
BI	Biotechnology Informatics
BICC	Biomedical Information Communications Center
BIOETHICSLINE	BIOETHICS onLINE
BIREME	Biblioteca Regional de Medicina - NLM's International MEDLARS Center in Brazil
BITNET	Because It's Time Network
BITNIS	BITNET NLM Intercommunication System
BLAST	Basic Local Alignment Search Tool
BLRC	Biomedical Library Review Committee
BOSC	Board of Scientific Counselors
CAFE	Cataloging Front-End
CANCERLIT	CANCER LITerature
CAS	Chemical Abstracts Service
CASE	Computer Assisted Software Engineering
CATLINE	CATalog onLINE
CBM	Current Bibliographies in Medicine
CC	Chemline's Classification Code
CCDS	Computer-based Curriculum Delivery Systems
CCEHRP	Committee to Coordinate Environmental Health and Related Programs
CCRIS	Chemical Carcinogenesis Research Information System
CD-I	Compact Disc-Interactive
CD-ROM	Compact Disk-Read Only Memory
CENDI	Commerce, Energy, NASA, NLM and Defense Information
CHEMID	Chemical Identification File
CHEMLEARN	Microcomputer-based training for CHEMLINE
CHEMLINE	CHEMical Dictionary OnLINE
CICS	Customer Information Control System
CLINPROT	CLINical cancer PROTocols
COACH	Expert searcher system prototype. To improve MEDLINE retrieval with Grateful Med
CODATA	Committee on Data for Science and Technology
COSTART	FDA's thesaurus of adverse reaction terms
CPT	The AMA's Current Procedural Terminology
CRISP	Computer Retrieval of Information on Scientific Projects
CROSSFILE	Permits TOXNET users to search for and display data from multiple files simultaneously
CSB	Computer Science Branch
CTX	Criteria Table Expert Systems
DART	Developmental and Reproductive Toxicology
DASD	Direct access storage devices
DBIR	Directory of Biotechnology Information Resources

DBMS	Database Management System
DCT	Discrete Cosine Transform
DCW	Document Capture Workstation
DENTALPROJ	Dental Projects database
DHHS	Department of Health and Human Services
DIRLINE	Directory of Information Resources Online
DOCLINE	DOCuments onLINE
DOCUSER	DOCument delivery USER
DOE	Department of Energy
DRAW	Direct Read After Write
DRW	Document Request Workstation
DSM-IIIR	American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders
DSRT	Document Storage, Retrieval, and Transmission
DXP	Digital X-ray Prototype
DxPLAIN	Massachusetts General Hospital's expert diagnostic system
DXPNET	Digital X-ray Prototype Network
E.T. Net	Educational Technology Network
ECR	Emergency Care Research Institute
EDDS	Electronic Document Delivery System
EDSR	Electronic Document Storage and Retrieval
EEO	Equal Employment Opportunity
EINECS	European Inventory of Commercial Chemical Substances
ELHILL	MEDLARS software named after Senator Lister Hill
EMIC	Environmental Mutagen Information Center
EMICBACK	Environmental Mutagen Information Center Backfile
EPA	Environmental Protection Agency
ER	Entity Relationship
ETICBACK	Environmental Teratology Information Center Backfile
EXA	Electronic X-ray Archive
FASEB	Federation of American Socities for Experimental Biology
FCCSET	Federal Coordinating Committee for Science, Engineering and Technology
FEDRIP	Federal Research-In-Progress
FIRST	First Independent Research Support and Transition
FLICC	Federal Library and Information Center Committee
FTE	Full-time equivalents
FTP	File Transfer Protocol
GenBank	National, NIH-supported DNA sequence database
GenInfo	Databank providing a core of biological information about sequences, including the sequence itself, that accurately reflects the journal literature
GM	Grateful Med
GRAS list	Generally Recognized as Safe List
HAP	Hazardous Air Pollutants List
HBCU's	Historically Black Colleges and Universities
HCTA	Health Care Technology Assessment
HDTV	High Definition Television
HEALTH	HEALTH planning & administration database
HISTLINE	HISTory of medicine onLINE
HOPE	Health Omnibus Programs Extension Act
HPCC	High Performance Computing and Communications
HSDB	Hazardous Substances Data Bank
HSTAR	Health Services and Technology Assessment Research (database)
IAIMS	Integrated Advanced Information Management System
IARC list	International Agency for Research on Cancer List
ICD-9-CM	International Classification of Diseases, 9th Edition, Clinical Modification
ICSTI	International Council for Scientific and Technical Information
IEEE	Institute for Electrical and Electronics Engineers
ILAR	Institute of Laboratory Animal Research
ILL	Interlibrary Loan
IMIA	International Medical Informatics Association
IMPAG	International MEDLARS Policy Advisory Group
INTELSAT	International Telecommunications Satellite Organization
INTROMED	A training/practice database
INTROTOX	A practice subset of HSDB for new users of TOXNET
INVESTIGATOR	A research program for knowledge acquisition planning
IOM	Institute of Medicine
IRIS	Integrated Risk Information System
IRW	Image Retrieval Workstation

IRx	Information Retrieval Experiment
IRxFD	IRx Fielded Data
ISL	Information Systems Laboratory
ISW	Image Server Workstation
ITB	Information Technology Branch
JHU	Johns Hopkins University
JIS	Journal Information System
KB	Knowledge Base
LAN	Local Area Network
LC	Library of Congress
LCSH	Library of Congress Subject Headings
LEXTOOL	An interactive lexicon building tool for adding entries to the SPECIALIST lexicon
LHNCBC	Lister Hill National Center for Biomedical Communications
LIS	Library Information Sciences
LO	Library Operations
LSTRC	Literature Selection Technical Review Committee
MACAW	Multiple Alignment Construction and Analysis Workbench
MARC	Machine-Readable Catalog
MedIndEx	Medical Indexing Expert
MEDLARS	MEDical Literature Analysis and Retrieval System
MEDLINE	MEDlars onLINE
MEDSTATS	Medical Statistics Expert System
MEDTUTOR	Microcomputer-based tutorial for MEDLINE
MeSH	Medical Subject Headings
MGH	Massachusetts General Hospital
MH	MeSH Heading
MI	Medical Informatics
Micro-CSIN	Chemical Substances Information Network
MIIS	Modified Interpretative Information System
MIM	Mendelian Inheritance in Man
MIPS	Million instructions per second
MIS	Management Information System
MisHIN	Mississippi Health Sciences Information Network
MLAA	Medical Library Assistance Act
MRAB	Machine-Readable Archives in Biomedicine
MRI	Magnetic resonance imaging
MUMPS	Massachusetts Utility Multi-Programming System
MX	CHEMLINE's Name of Mixture field
NAC	National Audiovisual Center
NARIC	National Rehabilitation Information Center
NCBI	National Center for Biotechnology Information
NCHS	National Center for Health Statistics
NEMA	National Electrical Manufacturers Association
NHANES	National Health and Nutrition Examination Surveys
NIAMS	National Institute of Arthritis, Musculoskeletal and Skin Diseases
NIC	National Informatics Center
NICHD	National Institute of Child Health and Human Development
NIEHS	National Institute of Environmental Health Sciences
NIH	National Institutes of Health
NIK	NLM Information Kiosk
NIOSH	National Institute for Occupational Safety and Health
NISO	National Information Standards Organization
NLQ	Natural Language Query
NLS	Natural Language Systems
NM	CHEMLINE's Name of Substance field
NN/LM	National Network of Libraries of Medicine
NREN	National Research and Education Network
NTIS	National Technical Information Service
NUCARE	NUrsing CAre REsearch
OCCS	Office of Computer and Communications Systems
OCR	Optical character recognition
OHIPD	Office of Health Information Programs Development
OHSU	Oregon Health Sciences University
OMIM	Online version, Mendelian Inheritance in Man
OPAC	Online Public Access Catalog
ORAU	Oak Ridge Associated Universities
ORNL	Oak Ridge National Laboratory
ORW	Online Reference Works
PA	CHEMLINE's MeSH Pharmacological Action Field
PADS	Packet Assembler-Disassemblers
PAFA list	Priority Based Assessment of Food Additives List
PAHO	Pan American Health Organization
PAM	Principals of Ambulatory Medicine
PDQ	Physician Data Query

PIR	Protein Identification Resource
POPLINE	POPulation information onLINE
QC	Quality Control
RACF	Resource Access Control Facility
RDBMS	Relational Database Management System
REFLINE	Subset of MEDLINE for NLM patrons
RelTox	Relational Toxicology Project
RFA	Request for Applications
RJE	Remote job entry
RML	Regional Medical Library
RMPs	Regional Medical Programs
RN	Registry Numbers
RQ List	Hazardous Substances Reportable Quantities List
RTECS	Registry of Toxic Effects of Chemical Substances
SAAS	Selection and Acquisition Subsystem
SAIL	System for Automated Interlibrary Loan
SDILINE	Selective Dissemination of Information onLINE
SERHOLD	Serial Holdings
SIC	Subcommittee on Information Coordination
SIDE	Sulzberger Institute of Dermatologic Education
SIS	Specialized Information Services
SNOMED	College of American Pathologists' Systematized Nomenclature of Medicine
SPECIALIST	Experimental system for parsing, analyzing, and accessing biomedical text
SPIE	Society of Photo-optical Instrumentation Engineers
SRW	Standardized Readings Workstations
STIC	Science and Technology Information Center
SUPERLIST	Important chemicals found on one or more of 16 Federal and state government lists
TESS	Technical Services System
TIP	Toxicology Information Program
TLC	The Learning Center for Interactive Technology
TOC	Table of Contents
TOXLEARN	Microcomputer-based training for TOXLINE
TOXLINE	TOXicology Information OnLINE
TOXLIT	TOXicology LITerature from special sources
TRI	Toxic Chemical Release Inventory
TRIFACTS	Toxic Chemical Release Inventory Facts
TSCA	Toxic Substances Control Act
UMDNS	ECRI's Universal Medical Device Nomenclature System
UMLS	Unified Medical Language System
URSP	Undergraduate Research Study Program
USAN	United States Adopted Names
VAMIS	Virginia Medical Information System
VTAM	Virtual Telecommunications Access Facility
WHO	World Health Organization
WORM	Write Once Read Many—Disc

APPENDIX 2: STAFF BIBLIOGRAPHY

Akrigg D, Attwood TK, Bleasby AJ, Findlay JB, North AC, Maughan NA, Parry SD, Perkins DN, Wootton JC. SERPENT—an information storage and analysis resource for protein sequences. Comput Appl Biosci 1992;8(3):295-6.

Boguski MS, Bairoch A, Attwood TK, Michaels GS. Proto-vav and gene expression [letter]. Nature 1992;358(6382):103.

Boguski MS, Hardison RC, Schwartz S, Miller W. Analysis of conserved domains and sequence motifs in cellular regulatory proteins and locus control regions using new software tools for multiple alignment and visualization. New Biol 1992;4:247-60.

Boguski MS, Murray AW, Powers S. Novel repetitive sequence motifs in the alpha and beta subunits of prenyl-protein transferases and homology of the alpha subunit to the MAD2 gene product of yeast [letter]. New Biol 1992;4(4):408-11.

Boguski MS, States DJ. Molecular sequence databases and their uses. In: Rees AR, Wetzel R, Sternberg MJE, ed. Protein engineering: a practical approach. Oxford: IRL Press, 1992;57-88.

Boguski MS. A molecular biologist visits Jurassic Park. Biotechniques 1992;12(5):668-9.

Boguski MS. Computational sequence analysis revisited: New databases, software tools and the research opportunities they engender. J Lipid Res 1992;33:957-74.

Boyko VP, Karasev AV, Agranovsky AA, Koonin EV, Dolja VV. Coat protein gene duplication in a filamentous RNA virus of plants. Proc Natl Acad Sci U S A 1992;89:9156-60.

Bryant SH, Lawrence CE. The frequency of ion-pair substructures in proteins is quantitatively related to electrostatic potential: a statistical model for nonbonded interactions. Proteins 1991;9(2):108-19.

Cassedy JH. A country doctor in Connecticut medicine and politics: William L. Higgins, 1867-1951. Conn Med 1992; 56:363-9.

Cassedy JH. Medicine in America: a short history. Baltimore: Johns Hopkins University, 1991;187.

Cassedy JH. Numbering the North's medical events: humanitarianism and science in Civil War statistics. Bull Hist Med 1992; 66:210-33.

Cheh ML, Kingsland LC III. Criteria table representation in the medical domain. In: Kahn M, Smith J Jr, Buchanan B, Musen M, Szolovits P, eds. Working notes of the AAAI Spring symposium series 1992, artificial intelligence in medicine. Menlo Park, CA: American Association for Artificial Intelligence, 1992;11-24.

Choi TJ, Kuwata S, Koonin EV, Heaton LA, Jackson AO. Structure of the L (polymerase) protein gene of sonchus yellow net virus. Virology 1992;189(1):31-9.

Claverie JM. Identifying coding exons by similarity search: Alu-derived and other potentially misleading protein sequences. Genomics 1992;12:838-41.

Colaianni LA. Retraction, comment, and errata policies of the US National Library of Medicine. Lancet 1992;340:536-7.

Colaianni LA. That vision thing. Bull Med Libr Assoc 1992 Jan;80(1):1-8.

Cookson J, Sneiderman C, Rivera C. Discrete cosine based image compression applied to dermatology. In: Proceedings of the Society of Photo-Optical Instrumentation Engineers, medical imaging V (1991). Bellingham WA: International Society for Optical Engineering 1991;374-8.

Dimri GP, Rudd KE, Morgan MK, Bayat H, Ames GF. Physical mapping of repetitive extragenic palindromic sequences in Escherichia coli and phylogenetic distribution among Escherichia coli strains and other enteric bacteria. J Bacteriol 1992;174(14):4583-93.

Doniger J, Landsman D, Gonda MA, Wistow G. The product of unr, the highly conserved gene upstream of N-ras, contains multiple repeats similar to the cold-shock domain (CSD), a putative DNA-binding motif. New Biol 1992;4(4):389-95.

Doszkocs TE. Neural networks in libraries: the potential of a new information technology. In: Chen CC, ed. NIT [New Information Technology]: fourth international conference on new information technology, Budapest, Hungary. West Newton MA: MicroUse Information, 1991;27-33.

Duronio RJ, Gordon JI, Boguski MS. Comparative analysis of the beta transducin family with identification of several new members including PWP1, a nonessential gene of Saccharomyces cerevisiae that is divergently transcribed from NMT1. Proteins 1992;13(1):41-56.

Ghosh D. Glucocorticoid receptor binding site in the human immunodeficiency virus long terminal repeat. J Virol 1992;66:586-90.

Ghosh D. New developments of a transcription factors database. Trends Biochem Sci 1991;16:445-7.

Ghosh D. TFD: the transcription factors database. Nucleic Acids Res 1992;20(Suppl):2091-3.

Harris N, Hunter L, States D. Megaclassification: discovering motifs in massive datastreams. In: AAAI-92: Proceedings of the tenth national conference on artificial intelligence. Menlo Park, CA: AAAI Press, 1992;836-42.

Hauser SE, Rivera C, Thoma G Factors affecting the performance of a DOS-based WORM file server In Digest of papers of the eleventh IEEE symposium on mass storage systems Los Alamitos CA IEEE Computer Society, 1991,33-7

Hendrickson W, Rudd KE Physical map location of the argFGH operon of Escherichia coli J Bacteriol 1992,174 (11) 3836-7

Humphrey SM Evolution toward knowledge-based indexing for information retrieval In Deogun J, ed Workshop on future directions in text analysis, retrieval and understanding Washington, DC National Science Foundation, 1991,132-9

Humphrey SM Interactive knowledge-based systems for improved subject analysis and retrieval In Lancaster FW, Smith LC, eds Artificial intelligence and expert systems will they change the library? Papers presented at the 1990 clinic on library applications of data processing, March 25-27, 1990 Urbana-Champaign, IL Graduate School of Library and Information Science of the University of Illinois at Urbana-Champaign, 1992,81-117

Humphreys BL, Lindberg DAB The Unified Medical Language System Project a distributed experiment in improving access to biomedical information In Lun KC, Degoulet P, Piemme TE, Rienhoff O, eds MEDINFO 92 Proceedings of the seventh World Congress on Medical Informatics New York North-Holland, 1992,1496-500

Hunter L, Harris N, States D Efficient classification of massive, unsegmented datastreams In Sleeman D, Edwards P, eds Machine learning Proceedings of the ninth international workshop (ML92) San Mateo, CA Morgan Kaufman, 1992,224-32

Hunter L, States D Bayesian classification of protein structure IEEE Expert 1992,7(4) 67-75

Hunter L Artificial intelligence and molecular biology In AAAI-92 Proceedings of the tenth national conference on artificial intelligence Menlo Park, CA AAAI Press, 1992,866-8

Hunter L Classifying for prediction a multistrategy approach to predicting protein structure In Michalski RS, Tecuci G, eds Proceedings of the first international workshop on multistrategy learning Fairfax, VA George Mason University Center for Artificial Intelligence, 1991,394-402

Hunter L Knowledge acquisition planning using multiple sources of knowledge to answer questions in biomedicine Mathematical & Computer Modelling 1992,16(6/7) 79-91

Ilyina TV, Gorbalenya AE, Koonin EV Organization and evolution of bacterial and bacteriophage primase-helicase systems J Mol Evol 1992,34 351-7

Ilyina TV, Koonin EV Conserved sequence motifs in the initiator proteins for rolling circle DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archaebacteria Nucleic Acids Res 1992,20(13) 3279-85

Jayakumar A, Rudd KE, Fabiny JM, Barnes EJ Localization of the Escherichia coli amtA gene to 95 8 minutes J Bacteriol 1991,173(5) 1572-3

Jin T, Rudd KE, Inouye M The nlpA lipoprotein gene is located near the selC tRNA gene on the Escherichia coli chromosome J Bacteriol 1992,174(11) 3822-3

Karlin S, Bucher P, Brendel V, Altschul SF Statistical methods and insights for protein and DNA sequences Annu Rev Biophys Biophys Chem 1991,20(175) 175-203

Kingsland LC III, Dos/kocs TD, Syed EJ, Schuyler PL, Lindberg, DAB COACH an expert searcher adjunct program for users of Grateful Med Program of the Medical Library Association (MLA) 1991 annual meeting Chicago, IL Medical Library Association, 1991,2a

Kingsland LC III, ed Comput Methods Programs Biomed Special Issue SCAMC Student Paper Competition 1990 Comput Methods Programs Biomed 1991,35(4) 233-334

Kingsland LC III, Syed EJ, Lindberg DAB Coach an expert searcher program to assist Grateful Med users searching MEDLINE In Lun KC, Degoulet P, Piemme TE, Rienhoff O, eds MEDINFO 92 Proceedings of the seventh world congress on medical informatics Amsterdam North-Holland, 1992,382-6

Kingsland LC III Finalists' papers, 1990, from the student paper competition in medical informatics at the fourteenth annual symposium on computer applications in medical care (SCAMC) Special issue, Comput Methods Programs Biomed 1991,35(4) 233-4

Koonin EV, Gorbalenya AE, Purdy MA, Rozanov MN, Reyes GR, Bradley DW Computer-assisted assignment of functional domains in the nonstructural polyprotein of hepatitis E virus delineation of an additional group of positive-strand RNA plant and animal viruses Proc Natl Acad Sci U S A 1992,89(17) 8259-63

Koonin EV, Gorbalenya AE An insect picornavirus may have genome organization similar to that of caliciviruses FEBS Lett 1992,297 81-6

Koonin EV, Gorbalenya AE Autogenous translation regulation by Escherichia coli ATPase SecA may be mediated by an intrinsic helicase activity of this protein FEBS Lett 1992,298 6-8

Koonin EV, Senkevich TG, Chernos VI A family of DNA virus genes that consists of fused portions of unrelated cellular genes Trends Biochem Sci 1992,17(6) 213-4

Koonin EV, Senkevich TG Evolution of thymidine and thymidylate kinases the possibility of independent capture of TK genes by different groups of viruses Virus Genes 1992,6(2) 187-96

Koonin EV, Senkevich TG Vaccinia virus encodes four putative DNA and/or RNA helicases distantly related to each other J Gen Virol 1992,73(Pt 4) 989-93

Koonin EV Archaebacterial virus SSV1 encodes a putative Dna-A protein Nucleic Acids Res 1992,20(5) 1143

Koonin EV. DnaC protein contains a modified ATP-binding motif and belongs to a novel family of ATPases including also DnaA. Nucleic Acids Res 1992;20(8):1997.

Landsman D. RNP-1, an RNA-binding motif is conserved in the DNA-binding cold shock domain. Nucleic Acids Res 1992;20(11):2861-4.

Lawrence CE, Bryant SH. Hydrophobic potentials from statistical analysis of protein structures. Methods Enzymol 1991;202(20):20-31.

Layne SP, Merges MJ, Dembo M, Spouge JL, Conley SR, Moore JP, Raina JL, Renz H, Gelderblom HR, Nara PL. Factors underlying spontaneous inactivation and susceptibility to neutralization of human immunodeficiency virus. Virology 1992;189(2):695-714.

Lightfoot DA, Baron AJ, Cock JM, Wootton JC. A nitrate reductase gene of the cyanobacterium Synechococcus PCC6301 inferred by heterologous hybridization, cloning and targeted mutagenesis. Genetica 1992;85(2):107-17.

Lindberg DA, Humphreys BL. The Unified Medical Language System (UMLS) and computer-based patient records. In: Ball MJ, Collen MF, eds. Aspects of the computer-based patient record. New York: Springer-Verlag, 1992;165-75.

Lindberg DA, Siegel ER. On assessing the impact of medical information: does MEDLINE make a difference? [editorial]. Methods Inf Med 1991 Oct;30(4):239-40.

Lindberg DA, West RT, Corn M. IAIMS: an overview from the National Library of Medicine. Bull Med Libr Assoc 1992 Jul;80(3):244-6.

Lindberg DA. Enrollment growth continues at NLM. US Med 1992 Jan:15-6.

Lo R, Locatis C, Ullmer E, Carr V, Banvard R, Williamson M, Ackerman M. Developing a database of authoring system software. Educ Tech Res Dev 1992;40(2):69-76.

Lo R, Locatis C, Ullmer E, Carr V, Banvard R, Le Q, Williamson M, Ackerman M. AuthorBase: a database of authoring system software. In: Clayton P, ed. Proceedings of the fifteenth annual symposium of computer applications in medical care. New York: McGraw-Hill, 1991;875-6.

Locatis C, Ullmer E, Carr V, Banvard R, Le Q, Lo R, Williamson M. Authoring systems. LHNCBC Monograph 92-1. Bethesda, MD: National Library of Medicine, 1992;27.

Locatis C, Ullmer E, Carr V, Banvard R, Le Q, Lo R, Williamson M. Authoring systems: an introduction and assessment. J Comput Higher Educ 1991;3(1):23-35.

Locatis C, Ullmer E, CarrV, Banvard R, Le Q, Lo R, Williamson M. Authoring systems reassessed. Educ Tech Res Dev 1992;40(2):77-82.

Major F, Lapalme G. Derivation d'une structure de controle de retour-arriere. In: Queinnec C, ed. Actes des journees JFLA92: Avancees applicatives. Perros-Guirec, France: BIGRE, 1992:202-19.

McCray AT, Aronson A. Mapping the language of users to the language of documents. In: Griffiths J-M, ed. Proceedings of the 54th ASIS annual meeting. Medford, NJ: Learned Information, 1991;328.

McCray AT, ed. Proceedings of the workshop on language and information processing [at the 54th annual ASIS meeting]. Silver Spring, MD: American Society for Information Science, 1991;126.

McCray AT. Natural language processing for intelligent information retrieval. In: Nagel JH, Smith WM, eds. Proceedings of the annual international conference of the IEEE Engineering in Medicine and Biology Society, Volume 13. New York: IEEE, 1991;1160-1.

McCray AT. Extending a natural language parser with UMLS knowledge. In: Clayton PD, ed. Proceedings of the fifteenth annual symposium on computer applications in medical care (SCAMC). New York: McGraw-Hill, 1991;194-8.

McCray AT. The potential for NLP in information retrieval. In: Deogun J, ed. Workshop on future directions in text analysis, retrieval and understanding. Washington, DC: National Science Foundation, 1991;43-7.

Mehnert RB. MEDLARS/MEDLINE. In: Ralston A, Reilly ED, eds. Encyclopedia of Computer Science. New York: Van Nostrand Reinhold, 1993;847-8.

Mehnert RB. The National Library of Medicine. Bowker Annual 1992;37:152-5.

Morgan PF, Fine RL, Montgomery P, Marangos PJ. Multidrug resistance in MCF-7 human breast cancer cells is associated with increased expression of nucleoside transporters and altered uptake of adenosine. Cancer Chemother Pharmacol 1991;29(2):127-32.

Mortimer RK, Schild D, Contopoulou CR, Kans JA. Genetic and physical maps of Saccharomyces cerevisiae. Methods Enzymol 1991;194:827-63.

Parascandola J, Harris M. Images of hospital pharmacy in America. Am J Hosp Pharm 1992;49:1413-21.

Parascandola J. Historical perspectives on in vitro toxicology. In: Goldberg AM, ed. In vitro toxicology: mechanisms and new technology. Alternative methods in toxicology; vol 8. New York: MA Liebert, 1991;87-96.

Parascandola J. L.J. Henderson and the mutual dependence of variables: from physical chemistry to Pareto. In: Elliott CA, Rossiter MW, eds. Science at Harvard: historical perspectives. Bethlehem: Lehigh University, 1992;167-90.

Parascandola J. The development of American pharmacology: John J. Abel and the shaping of a discipline. Baltimore: Johns Hopkins University, 1992;212.

Prabhu VV, Young L, Prohofsky EW. Observation of preferred direction melting in a lattice model of poly(dA-dC).poly(dG-dT). Biopolymers 1991;31(3):355-8.

Ram A, Hunter L. The Use of explicit goals for knowledge to guide inference and learning. Journal of Applied Intelligence 1992;2(1):47-73.

Ramsingh A, Araki H, Bryant S, Hixson A. Identification of candidate sequences that determine virulence in Coxsackievirus B4. Virus Res 1992;23:281-92.

Recipon H, Perasso R, Adoutte A, Quetier F. ATP synthase subunit c/III/9 gene sequences as a tool for interkingdom and metaphytes molecular phylogenies. J Mol Evol 1992;34(4):292-303.

Rindflesch TC, Reeves JE, Kac MB. Theoretical implications of disordered syntactic comprehension. Language Sciences 1992;14(1):1-25.

Rozanov MN, Koonin EV, Gorbalenya AE. Conservation of the putative methyltransferase domain: a hallmark of the 'Sindbis-like' supergroup of positive-strand RNA viruses. J Gen Virol 1992;73(8):2129-34.

Rudd KE, Bouffard G, Miller W. Computer analysis of E. coli restriction maps. In: Davies KE, Tilghman SM, ed. Genome analysis. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 1992;1-38. (Genome Analysis; vol 4; Strategies for Physical Mapping).

Rudd KE, Schneider TD. Compilation of E. coli ribosome binding sites. In: Miller J, ed. A short course in bacterial genetics: a laboratory manual and handbook for Escherichia coli and related bacteria. Cold Spring Harbor, NY: Cold Spring Harbor Press, 1992;17.19-17.45.

Rudd KE. Alignment of E. coli DNA sequences to a revised, integrated genomic restriction map. In: Miller J, ed. A short course in bacterial genetics: a laboratory manual and handbook for Escherichia coli and related bacteria. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 1992;2.3-2.43.

Schuyler PL, Sinn SK, Weiss PJ. Proposal #4: Arguments in favor. In: Conway MO, ed. The future of subdivisions in the Library of Congress subject headings system: report from the Subject Subdivision Conference. Washington, D.C.: Library of Congress, 1992;75-9.

Schwartz R. When we last met ... Networking, telecommunications and networked information resources. PVC [Potomac Valley Chapter. American Society for Information Science] Currents 1992 Feb:8.

Schwartz R. When we last met ... NREN: Information access for the future. PVC [Potomac Valley Chapter. American Society for Information Science] Currents 1992 Jan:4.

Seibert WF, Kuenz MA. Growth and change in 67 medical school libraries, 1975-1989. Bethesda, MD: NIH, National Library of Medicine, 1992;37.

Seibert WF, Kuenz MA. Statistics and library history. [Letter to the editor]. Libraries and Culture 1992;27:110-3.

Shin DG, Lee C, Zhang J, Rudd KE, Berg CM. Redesigning, implementing and integrating Escherichia coli genome software tools with an object-oriented database system. Comput Appl Biosci 1992;8(3):227-38.

Siegel ER, Rapp BA, Lindberg DA. Evaluating the impact of MEDLINE using the Critical Incident Technique. In: Clayton PD, ed. Fifteenth Annual Symposium on Computer Applications in Medical Care. Washington, DC: McGraw-Hill, 1991:83-7.

Sikorski RS, Michaud WA, Wootton JC, Boguski MS, Connelly C, Hieter P. TPR proteins as essential components of the yeast cell cycle. Cold Spring Harb Symp Quant Biol 1991;56:663-73.

Sinn SK. The Development of classification and subject heading systems for medicine. In: Weinberg BH, ed. Cataloging heresy: challenging the standard bibliographic product. Medford, NJ: Learned Information, 1992;39-55.

Sneiderman C, Cookson J, Hood A. Use of computer graphic images in teaching dermatology. Comput Med Imaging Graph 1992;16:151-2.

Sneiderman CA, Bicknell EJ. Computer-assisted dynamic integration of multiple medical thesauruses. Comput Biol Med 1992;22(1-2):135-45.

Sneiderman CA, Hood AF. A comparison of video and photographic display of skin lesions for morphology recognition. J Biocommun 1992;18(4):22-5.

Sparks S. Connect with latest health science information . Int Nurs Rev 1991;38(4):99.

Sparks S. New electronic nursing care research conference. Nursing Educators Microworld 1992;6(2):13.

Spouge JL. Fast optimal alignment. Comput Appl Biosci 1991;7(1):1-7.

Spouge JL. Statistical analysis of sparse infection data and its implications for retroviral treatment trials in primates. Proc Natl Acad Sci U S A 1992;89(16):7581-5.

States D, Hunter L. Models and the Bayesian classification of protein structural elements. In: Buntine W, Del Alto M, eds. Collected notes on the workshop for pattern discovery in large databases. Moffett Field, CA: NASA Ames Research Center, 1991. Technical Report FIA-91-07;135-9.

States DJ, Botstein D. Molecular sequence accuracy and the analysis of protein coding regions. Proc Natl Acad Sci U S A 1991;88(13):5518-22.

States DJ, Gish W, Altschul SF. Improved sensitivity of nucleic acid database searches using application-specific scoring matrices. Methods Companion to Methods Enzymol 1991;3:66-70.

States DJ. Molecular sequence accuracy: analysing imperfect data. Trends Genet 1992;8(2):52-5.

Stepien PP, Margossian SP, Landsman D, Butow RA. The yeast nuclear gene suv3 affecting mitochondrial post-transcriptional processes encodes a putative ATP-dependent RNA helicase. Proc Natl Acad Sci U S A 1992;89 (15):6813-7.

Thoma GR. Automated document delivery. Quarterly Bulletin of the International Association of Agricultural Information Specialists 1992;37(1-2):84-8.

Thoma GR. Automated electronic document delivery: system design considerations. In: Nelson LJ, ed. Proceedings: image compression conference. Norwell, MA: BIS Strategic Decisions, 1991;v.1-5.

Toone WM, Rudd KE, Friesen JD. Mutations causing aminotriazole resistance and temperature sensitivity reside in gyrB, which encodes the B subunit of DNA gyrase. J Bacteriol 1992;174(16):5479-81.

Ullmer E. Learning environments: the technology-cognition connection. In: Simonson MR, Jurasek K, eds. Proceedings of selected research paper presentations at the annual meeting of the Association for Educational Communications and Technology. Washington, DC: AECT, 1992;834-42.

Vakharia DD, Bryant SH, Dias JA. Topographic analysis of human follicle stimulating hormone-beta using anti-peptide antisera. Mol Cell Endocrinol 1992;85:89-97.

vanBemmel J, McCray AT, eds. Yearbook of medical informatics 92: advances in an interdisciplinary science. Stuttgart: Schattauer Verlag, 1992;523.

Vasta BM, Casey SM. Downsizing: a case study of the National Library of Medicine's TOXNET system. MUG (Mumps Users' Group) Q 1991;21(5):11-4.

Walker FL, Thoma GR. A prototype workstation for accessing and using an optical disc-based database of biomedical documents. In: Dillon M, ed. Interfaces for information retrieval and online systems. Westport CT: Greenwood, 1991;207-18.

Weiss PJ. MARC Notes. LITA [Library & Information Technology Association]Newsletter 1992;13(1):11-3.

Weiss PJ. MARC Notes. LITA [Library & Information Technology Association]Newsletter 1992;13(3):20-1.

Weiss PJ. MARC Notes. LITA [Library & Information Technology Association] Newsletter 1992;13(4):19-21.

White CT, Singh RK, Reintjes PB, Lampe J, Erickson BW, Dettloff WD, Chi VL, Altschul SF. BioSCAN: a VLSI-based system for biosequence analysis. Proceedings of the IEEE international conference on computer design: VLSI in computers and processors. Boston: IEEE Computer Society Press, 1991:504-9.

Wilbur WJ, Sirotkin K. The automatic identification of stop words. J Inf Sci 1992;18:45-55.

Wilbur WJ. An information measure of retrieval performance. J Info Systems 1992;17(4):283-98.

Wilbur WJ. Retrieval testing by the comparison of statistically independent retrieval methods. J Am Soc Info Sci 1992;43(5):358-70.

Wootton JC, Federhen S. Segmentation by sequence complexity: the significance of residue clusters in protein sequences (extended abstract). In: Konopka AK, Salamon P, ed. Open problems of computational molecular biology. Telluride, Colorado: Telluride Summer Research Center, Telluride Academy, 1992:77-81.

APPENDIX 3: FY 1992 EXTRAMURAL PROGRAMS-SUPPORTED PUBLICATIONS

Amero SA, Kretsinger RH, Moncrief ND, Yamamoto KR, Pearson WR. The origin of nuclear receptor proteins: a single precursor distinct from other transcription factors. Mol Endocrinol; 1992; 6(1): 3-7. (R01 LM 04969).

Barker WC, George DG, Hunt LT, Garavelli JS. The PIR protein sequence database. Nucleic Acids Research; 1991; 19 Suppl: 2231-6.(P41 LM 05206).

Bergeron B. Challenges associated with providing speech recognition user interfaces for computer-based educational systems. Collegiate Microcomputer; 1991 Aug; 9(3): 129-43. (R29 LM 04715).

Bergeron B. The effects of stereophonic priming on the development of auscultatory pattern recognition skills. Collegiate Microcomputing; 1991 Feb; 9(1): 25-35. (R29 LM04715).

Bergeron BP, Shiffman RN, Rouse RL, Greenes RA. Composing user models through logic analysis. In: Clayton PD, ed. Assessing the value of medical informatics. 15th Annual Symposium on Computer Applications in Medical Care; a conference of the American Medical Informatics Association; 1991 Nov 17-20; Washington, DC. New York: McGraw-Hill; [1992]: 681-5. (R01 LM 04572 and 4715 and T15 LM 07037).

Berner ES, Shugerman AA. Needs assessment for diagnostic decision support systems (DDSS). In: Clayton, PD, ed. Assessing the value of medical informatics. 15th Annual Symposium on Computer Applications in Medical Care; a conference of the American Medical Informatics Association; 1991 Nov 17-20; Washington, DC. New York: McGraw-Hill; [1992]: 605-8. (R01 LM 05125).

Brinkley JF. Semi-automatic medical image segmentation using knowledge of anatomic shape. In: Proceedings, SPIE Medical Imaging V: Image Processing; 1991 Feb 27-Mar 1; San Jose, CA: 78-87. (R01 LM 04925).

Brinkley JF. Spatial anatomic knowledge for 2-D interactive medical image segmentation and matching. In: Clayton, PD, ed. Assessing the value of medical informatics. 15th Annual Symposium on Computer Applications in Medical Care; a conference of the American Medical Informatics Association; 1991 Nov 17-20; Washington, DC. New York: McGraw-Hill; [1991]: 460-4. (R29 LM 04925).

Broering NC, Bagdoyan HE. The impact of IAIMS at Georgetown: strategies and outcomes. Bull Med Libr Assoc; 1992 Jul; 80(3): 263-75. (G08 LM 04392).

Broering NC. An Integrated Academic Information Management System (IAIMS) at Georgetown University. In: Lun KC, Degoulet P, Piemme TE, Rienhoff O, eds. MEDINFO 92; proceedings of the Seventh World Congress on Medical Informatics; 1992 Sep 6-10; Geneva, Switzerland. Amsterdam: North-Holland; 1992: 112-6. (G08 LM 04392).

Brunet LW, Morrissey CT, Gorry GA. Oral history and information technology: human voices of assessment. J Organizational Computing; 1991; 1(3): 251-74. (G08 LM 04905).

Christensen C, Elstein AS, Bernstein LM. Formal decision supports in medical practice and education. Teaching and Learning in Medicine; 1991; 3(2): 62-70. (R01 LM 04583).

Christensen C, King A, Fetzer B. Medical students' reactions to AIDS: influence of patient characteristics on hypothetical treatment decisions. Teaching and Learning in Medicine; 1991; 3(3): 138-42. (R01 LM 04583).

Cimino JJ, Elkin PL, Barnett GO. As we may think: the concept space and Medical Hypertext. Comp Biomed Res; 1992; 25(3): 238-63. (G08 LM 04419).

Cimino JJ, John SB, Hripcsak G, Sideli RV, Fink DJ, Friedman C, Clayton PD. One year's experience with the Unified Medical Language System (UMLS) in academia and patient care. In: Lun KC, Degoulet P, Piemme TE, Rienhoff O, eds. MEDINFO 92; proceedings of the Seventh World Congress on Medical Informatics; 1992 Sep 6-10; Geneva, Switzerland. Amsterdam: North-Holland; 1992: 1501-5. (G08 LM 04419 and R29 LM 05397).

Clayton PD, Nobel JJ. Cost and cost justification for integrated information systems in medicine. In: Bakker AR, Ehlers CT, Bryant JR, Hammond WE, eds. Proceedings of the IMIA Working Conference on Hospital Information Systems: Scope-Design-Architecture; 1991 Sep 7-11; Goettingen, Germany. Amsterdam: North-Holland; 1992: 133-40. (G08 LM 04419).

Clayton PD, Grossman DD. Current developments and future trends in information technology in healthcare. In: Bakker AR, Ehlers CT, Bryant JR, Hammond WE, eds. Proceedings of the IMIA Working Conference on Hospital Information Systems: Scope-Design-Architecture; 1991 Sep 7-11; Goettingen, Germany. Amsterdam: North-Holland; 1992: 19-25. (G08 LM 04419).

Clayton PD, Anderson RK, Hill C, McCormack M. An initial assessment of the cost and utilization of the Integrated Academic Information System (IAIMS) at Columbia- Presbyterian Medical Center. In: Clayton PD, ed. Assessing the value of medical informatics. 15th Annual Symposium on Computer Applications in Medical Care; a conference of the American Medical Informatics Association; 1991 Nov 17-20; Washington, DC. New York: McGraw-Hill; [1992]: 109-13. (G08 LM 04419).

Clayton PD, Sideli RV, Sengupta S. Open architecture and integrated information at Columbia-Presbyterian Medical Center: the role of networks, workstations and databases. MD Comput; 1992 Sep-Oct; 9(5): 297-303. (G08 LM 04419).

Deibel SRA, Greenes RA. The DeSyGNER data access element: a readily reusable component for the construction of data-compatible multimedia programs. In: Clayton PD, ed. Assessing the value of medical informatics. 15th Annual Symposium on Computer Applications in Medical Care; a conference of the American Medical Informatics Association; 1991 Nov 17-20; Washington, DC. New York: McGraw-Hill; [1992]: 776-80. (R01 LM 04572).

Deshpande AS, Richards DS, Pearson WR. A platform for biological sequence comparison on parallel computers. Cabios; 1991; 7(2): 237-47. (R01 LM 04969).

The digital anatomist: human brain animations [[computer+videodisc]]. Sundsten JW, Kastella KG. Seattle, WA: Health Sciences Center for Learning Resources, Univ. of Wash.; 1991; (R01 LM 04925).

Doyle J. The foundations of psychology: a logico-computational inquiry into the concept of mind. In: Cummins R, Pollock J, eds. Philosophy and Artificial Intelligence: Essays at the Interface. Cambridge, MA: MIT Press; 1991: 40-77. (R01 LM 04493).

Doyle J, Shoham Y, Wellman MP. A logic of relative desire (preliminary report). In: Ras ZW, Zemankova M, eds. Methodologies for Intelligent Systems. 6 ed. Berlin: Springer-Verlag; 1991: 16-31. (R01 LM 04493).

Doyle J, Wellman MP. Rational self-government and universal default logics. In: Bourgine P, Walliser B, ed. Economics and cognitive science. Oxford: Pergamon; 1991: 5-13. (R01 LM 04493).

Doyle J. Reason maintenance and belief revision; foundations vs. coherence theories. In: Gardenfors P, ed. Belief Revision. Cambridge: Cambridge Univ. Press; 1992: 29-51. (R01 LM 04493).

Eckman MH, Wong JB, Pauker SG. The role of clinical decision analysis in medical quality management. In: Couch JB, ed. Health Care Quality Management for the 21st Century. Tampa, FL: American College of Physician Executives; 1991: 253-80. (R01 LM 04493).

Edison AS, Westler WM, Markley JL. Elucidation of amino acid spin systems in proteins and determination of heternuclear coupling constants by carbon-proton-proton three-dimensional NMR. J Magnetic Resonance; 1991; 92: 434-8. (R01 LM 04958).

Elstein AS. On the psychology of clinical intuition; discussion about intuition in surgery as a strategy of medical decision making: its potency and limitations. Theor Surg; 1991; 6: 95-97. (R01 LM 04583).

Eno K, Sundsten JW, Brinkley JF. A multimedia anatomy browser incorporating a knowledge base and 3D images. In: Clayton PD, ed. Assessing the value of medical informatics. 15th Annual Symposium on Computer Applications in Medical Care; a conference of the American Medical Informatics Association; 1991 Nov 17-20; Washington, DC. New York: McGraw-Hill; [1991]: 727-31. (R01 LM 04925).

Fejzo J, Westler WM, Markley JL, Macura S. Complete elimination of spin diffusion from selected resonances in two-dimensional cross-relaxation spectra of macromolecules by a novel pulse sequence (SNOESY). J Am Chem Soc; 1992; 114: 1523-4. (R01 LM0 4958).

Fejzo, J, Krezel AM, Westler WM, Macura S, Markley JL. Direct cross-relaxation NOESY (D.NOESY). A method for removing spin-diffusion cross peaks from two-dimensional NOE spectra of macromolecules. J Magnetic Resonance; 1991; 92: 651-7. (R01 LM 04958).

Fejzo J, Westler WM, Macura S, Markley JL. Elimination of chemical-exchange-mediated spin diffusion from exchange spectra of macromolecules. Exchange-decoupled NOESY (XD-NOESY). J Magnetic Resonance; 1991; 92: 195-202. (R01 LM 04958).

Fejzo J, Zolnai Z, Macura S, Markley JL. Quantitative evaluation of two-dimensional cross-relaxation NMR spectra of proteins. Interproton distances in turkey ovomucoid third domain. J Magnetic Resonance; 1990; 88: 93-110. (R01 LM 04958).

Fejzo, J, Krezel AM, Westler WM, Macura S, Markley JL. Refinement of the NMR solution structure of a protein to remove distortions arising from neglect of internal motion. Biochemistry; 1991 Apr 23; 30(16): 3807-11. (R01 LM 04958).

Friedman CP, de Blick R, Gilmer JS, Twarog RG, File DD. Influence of a computer database and problem exercises on students' knowledge of bacteriology. Acad Med; 1992 May; 67(5): 332-8. (R01 LM 04843).

Garavelli JS. Molecular modeling on the Commodore Amiga. J Mol Graphics; 1991 Mar; 9: 24-6, 36. (P41 LM 05206).

Georgeson S, Linzer M, Griffith JL, Weld L, Selker HP. Acute cardiac ischemia in patients with syncope: importance of the initial electrocardiogram. J Gen Intern Med; 1992 Jul/Aug; 7: 379-86. (T15 LM 07044 and R01 LM 04493).

Georgeson S, Coombs AT, Eckman MH. Prophylactic use of the intra-aortic balloon pump in high-risk cardiac patients undergoing noncardiac surgery: a decision analytic view. Am J Med; 1992 Jun; 92: 665-78. (R01 LM 04493 and T15 LM 07044).

Gorry GA. Information technology and the academic medical center. Acad Med; 1992; 67(1): 18-21. (G08 LM 04905).

Greenes RA, Deibel SRA. Application development in the DeSyGNER environment: dynamic assembly of independent information resources via direct manipulation authoring. In: Clayton PD, ed. Assessing the value of medical informatics. 15th Annual Symposium on Computer Applications in Medical Care; a conference of the American Medical Informatics Association; 1991 Nov 17-20;

Washington, DC New York McGraw-Hill, [1992] 913-5 (R01 LM 04572 and T15 LM 07037)

Greenes, RA Approaches to sharing and collaboration through modular system design a focus on knowledge management In Timmers T, Blum BI, eds Software engineering in medical informatics Amsterdam North-Holland, 1991 409-28 (R01 LM 04572 and T15 LM 07037)

Greenes RA A "building block" approach to application development for education and decision support in radiology implications for integrated clinical information systems environments J Digit Imaging, 1991 Nov, 4(4) 213-25 (R01 LM 04572 and T15 LM 07037)

Grewal R, Arcus J, Bowen J, Fitzpatrick KT, Hammond WE, Hickey L, Stead WW Bedside computerization of the ICU, design issues benefits of computerization versus ease of paper & pen In Clayton PD, ed Assessing the value of medical informatics 15th Annual Symposium on Computer Applications in Medical Care, a conference of the American Medical Informatics Association, 1991 Nov 17-20, Washington, DC New York McGraw-Hill, [1992] 793-7 (G08 LM 04958)

Guide to the American Medical Association historical health fraud and alternative medicine collection Hafner AW, Carson JG, Zwicky JF, eds Chicago American Medical Association, 1991 215p (G08 LM 04637)

Guo D, Lincoln MJ, Haug PJ, Turner CW, Warner HR Exploring a new best information algorithm for Iliad In Clayton PD, ed Assessing the value of medical informatics 15th Annual Symposium on Computer Applications in Medical Care, a conference of the American Medical Informatics Association, 1991 Nov 17-20, Washington, DC New York McGraw-Hill, [1991] 624-8 (R29 LM 05260 and R01 LM 04604)

Haimowitz IJ, Kohane IS Influences on the performance of hospital clinical event monitoring In Clayton PD, ed Assessing the value of medical informatics 15th Annual Symposium on Computer Applications in Medical Care, a conference of the American Medical Informatics Association, 1991 Nov 17-20, Wash , DC New York McGraw-Hill, [1992] 614-18 (R01 LM 04493)

Hammond WE, Stead WW Adopting TMR for physician/nurse use In Clayton PD, ed Assessing the value of medical informatics 15th Annual Symposium on Computer Applications in Medical Care, a conference of the American Medical Informatics Association, 1991 Nov 17-20, Washington, DC New York McGraw-Hill, [1992] 833-7 (G08 LM 04613)

Haug P, Beesley D Automated selection of clinical data to support radiographic interpretation In Clayton PD, ed Assessing the value of medical informatics 15th Annual Symposium on Computer Applications in Medical Care, a conference of the American Medical Informatics Association, 1991 Nov 17-20, Washington, DC New York McGraw-Hill, [1992] 593-7 (R01 LM 04932)

Haug P, Hoak S Veristat a support tool for knowledge development In Miller RA, ed SCAMC proceedings 14th Annual Symposium on Computer Applications in Medical Care Standards in Medical Informatics, 1990 Nov 4-7, Washington, DC Los Alamitos (CA) IEEE Computer Society Press, [1990] 650-4 (R01 LM 04932)

Haug PJ, Clayton PD, Tocino I, Morrison JW, Elliot CG, Collins DV, Harada SK, Frederick PR Chest radiography a tool for the audit of report quality Radiology, 1991, 180(271-6 (R01 LM 04932))

Haug PJ Diagnostic coding in the radiology department Curr Imaging, 1991, 3 99-105 (R01 LM 04932)

Haug PJ, Frederick PR, Tocino I Quality control in a medical information system Med Decis Making, 1991, 11(suppl) S57-S60 (R01 LM 4932)

Haynes RB, Ramsden MF, McKibbon KA, Walker CJ Online access to MEDLINE in clinical settings impact of user fees Bull Med Libr Assoc, 1991 Oct, 79(4) 377-81 (R01 LM 04696)

Healy JC, Bozek SA Wright-Giemsa a computer-based system for the creation and presentation of peripheral blood smears Lab Med, 1991 Oct, 22(10) 728-31 (T15 LM 07044)

Heckerling PS, Elstein AS, Terzian CG, Kushner MS The effect of incomplete knowledge on the diagnoses of a computer consultant system Med Inform, 1991, 16(4) 363-70 (R01 LM 04583)

Hendrickson G, Anderson RK, Clayton, PD, Cimino JJ, Hripcsak G, Johnson SB, McCormack M, Sengupta S, Shea SJC, Sideli R, Roderer, N The Integrated Academic Information Management System at Columbia- Presbyterian Medical Center MD Comput, 1992 Jan-Feb, 9(1) 35-42 (G08 LM 04419)

Herrmann FR, Safran C, Levkoff SE, Minaker KL Serum albumin level on admission as a predictor of death, length of stay, and readmission Arch Intern Med, 1992 Jan, 152 125-30 (R23 LM 04260)

Hillson SD, Dowd B, Rich EC, Luxenberg MG Call nights and patient care effects on inpatients at one teaching hospital J Gen Intern Med, 1992 Jul/Aug, 7 405-10 (T15 LM 07041)

Hillson SD, Feldman R, Wingert TD Economies of scope and payment for physician services Med Care, 1992 Sep, 30(9) 822-31 (T15 LM 07041)

Hlatky MA, Lipscomb J, Nelson C, Califf RM, Pryor DB, Wallace AG, Mark DB Resource use and cost of initial coronary revascularization Circulation, 1990 Nov, 82(Suppl IV)(5) 208-13 (G08 LM 04613)

Hoogstraal H Bibliography of ticks and tickborne diseases New Delhi, India Amerind, 1992, Vol 8 327p

Hripcsak G Arden Syntax for Medical Logic Modules MD Comput, 1991 Mar-Apr, 8(2) 76, 78 (G08 LM 04419)

Hripcsak G, Cimino JJ, Johnson SB, Clayton PD The Columbia-Presbyterian Medical Center decision-support system as a model for implementing the Arden Syntax In

Clayton PD, ed. Assessing the value of medical informatics. 15th Annual Symposium on Computer Applications in Medical Care; a conference of the American Medical Informatics Association; 1991 Nov 17-20; Washington, DC. New York: McGraw-Hill; [1992]: 248-52. (G08 LM 04419).

Jayes RL, Beshansky JR, D'Agostino RB, Selker HP. Do patients'coronary risk factor reports predict acute cardiac ischemia in the emergency department? a multicenter study. J Clin Epidemiol; 1992; 45(6): 621-6. (R01 LM 04493 and T15 LM 07044).

Jayes RL, Larsen GC, Beshansky JR, D'Agostino RB, Selker HP. Physician electrocardiogram reading in the emergency department—accuracy and effect on triage decisions: findings from a multicenter study. J Gen Intern Med; 1992 Jul/Aug; 7: 387-92. (T15 LM 07044 and R01 LM 4493).

Johnson SB, Clayton PD, Fink D, Sengupta S, Shea S, Bourne P, Sideli RV, Aguirre A, Cimino JJ, Hripcsak G, McCormack M, Hill C. Achievements in Phase III of an Integrated Academic Information Management System. In: Lun KC, Degoulet P, Piemme TE, Rienhoff O, eds. MEDINFO 92; Proceedings of the Seventh World Congress on Medical Informatics; 1992 Sep 6-10; Geneva, Switzerland. Amsterdam: North-Holland; 1992: 117-23. (G08 LM 04419).

Johnson SB, Friedman C, Cimino JJ, Clark A, Hripcsak G, Clayton PD. Conceptual data model for a central patient database. In: Clayton PD, ed. Assessing the value of medical informatics. 15th Annual Symposium on Computer Applications in Medical Care; a conference of the American Medical Informatics Association; 1991 Nov 17-20; Washington, DC. New York: McGraw-Hill; [1992]: 381-5. (G08 LM 04419).

Kleinmuntz B. Can computers be clinicians? Theory and design of a diagnostic system. In: Hersen M, Kazdin AE, Bellack AS, eds. The clinical psychology handbook. 2nd ed. New York: Pergamon; 1991: 506-15. (R01 LM 04583).

Korschelt E. Regeneration and transplantation. Carlson BM, ed. Canton, MA: Science History Publications; 1992; 3 volumes.

Larsen GC, Manolis AS, Sonnenberg RA, Beshansky JR, Estes NAM, Pauker SG. Cost-effectiveness of the implantable cardioverter-defibrillator: effect of improved battery life and comparison with amiodarone therapy. J Am Coll Cardiol; 1992 May; 19(6): 1323-34. (R01 LM 04493).

Lau LM, Warner HR. Performance of a diagnostic system (Iliad) as a tool for quality assurance. In: Clayton PD, ed. Assessing the value of medical informatics. 15th Annual Symposium on Computer Applications in Medical Care; a conference of the American Medical Informatics Association; 1991 Nov 17-20; Washington, DC. New York: McGraw-Hill; [1991]: 1005-10. (R29 LM 05260).

Lindsley DL, Zimm GG. The genome of Drosophila melanogaster. San Diego, CA: Academic Press; 1992: 1133p. (R01 LM 03689).

Long WJ, Naimi S, Criscitiello MG. Development of a knowledge base for diagnostic reasoning in cardiology. Comput Biomed Res; 1992; 25: 292-311. (R01 LM 04493).

Long WJ. The probability of disease. In: Clayton PD, ed. Assessing the value of medical informatics. 15th Annual Symposium on Computer Applications in Medical Care; a conference of the American Medical Informatics Association; 1991 Nov. 17-20; Wash., DC. New York: McGraw-Hill; [1992]: 619-23. (R01 LM 04493).

Myers E, Blanco C, Hallick RB, Jahnke J. MacMolecule: interactive, 3D molecular modeling curriculum development program, featuring a video animation generation program, and a molecular image database. 1.7 ed. Tucson, AZ: Univ. of Arizona; 1992; (R01 LM 04960).

Mark DB, Shaw L, Harrell FE, Hlatky MA, Lee KL, Bengtson JR, McCants CB, Califf RM, Pryor DB. Prognostic value of a treadmill exercise score in outpatients with suspected coronary artery disease. N Engl J Med; 1991 Sep 19; 325: 849-53. (G08 LM 04613).

Markley JL, Darba P, Fejzo, Krezel AM, Macura S, McNemar CW, Mooberry ES, Seavey BR, Westler WM, Zolanae Z. Computational aspects of multinuclear NMR spectroscopy of proteins at NMRFAM. In: Hoch JC et al., eds. Computational aspects of the study of biological macromolecules by nuclear magnetic resonance spectroscopy. New York: Plenum; 1991: 39-50. (R01 LM 04958).

Markley JL, Seavey BR, Alexandrescu AT, Darba P, Hinck AP, Loh SN, McNemar CW, Mooberry ES, Oh B-H, Stockman BJ, Wang J, Westler WM, Zehfus MH, Zolnai Z. Multinuclear magnetic resonance spectroscopy of proteins: information content, data extraction and analysis, and database design. In: Ikehara M, Oshima T, Titani K, eds. Protein engineering, protein design in basic research, medicine, and industry. New York: Springer Verlag; 1990: 285-90. (R01 LM 04958).

McClure RC, Schwartz L, Greenes RA, Deibel SRA, Marquis A, McClay JC. Education and decision support in radiology: authoring using a building block architecture. In: Brody WR, Johnston GS, eds. Computer applications to assist radiology. Carlsbad, CA: Symposia Foundation; 1992: 377-83. (R01 LM 04572 and T15 LM 07037).

Merriam J, Ashburner M, Hartl DL, Kafatos FC. Toward cloning and mapping the genome of drosophila. Science; 1991 Oct 11; 254: 221-5. (R01 LM 04896).

Myers EW, Huang X. An 0(N2 log N) restriction map comparison and search algorithm. Bull Math Biol; 1992; 54(4): 599-618. (R01 LM 04960 and 5110).

Myers G. A four Russians algorithm for regular expression pattern matching. J Assoc Computing Machinery; 1992 Apr; 39(4): 430-48. (R01 LM 04960).

Pearson WR. Identifying distantly related protein sequences. Current Opinion in Structural Biology; 1991; 1: 321-6. (R01 LM 04969).

Pearson WR. Searching protein sequence libraries: comparison of the sensitivity and selectivity of the Smith-

Waterman and FASTA algorithms Genomics, 1991, 11(635-50 (R01 LM 04969))

Pirogov NI Questions of life, diary of an old physician Zarechnak GV, ed Canton, MA Science History Publications, 1992 480p

Poses RM, Cebul RD, Wigton RS, Centor RM, Collins M, Fleischli G Controlled trial using computerized feedback to improve physicians' diagnostic judgments Acad Med, 1992 May, 67(5) 345-7 (R01 LM 04321)

Pryor DB, Shaw L, Harrell FE, Lee KL, Hlatky MA, Mark DB, Muhlbaier LH, Califf RM Estimating the likelihood of severe coronary artery disease Am J Med, 1991 May, 90 553-62 (G08 LM 04613)

Pryor DB, Lee KL Methods for the analysis and assessment of clinical databases the clinician's perspective Stat Med, 1991, 10 617-29 (G08 LM 04613)

Puerta AR, Egar JW, Musen MA Automated general of adaptable knowledge-acquisition tools with Mecano Stanford, CA Knowledge Systems Laboratory Medical Computer Science, 1991 Dec, KSL-91-62, 8p (R01 LM 05157)

Puerta AR, Egar JW, Tu SW, Musen MA A multiple-method knowledge-acquisition shell for the automatic generation of knowledge-acquisition tools Knowledge Acquisition, 1992, 4 171-96 (R01 LM 05157 and R01 LM 05208)

Roderer NK, Clayton PD IAIMS at Columbia-Presbyterian Medical Center accomplishments and challenges Bull Med Libr Assoc, 1992 Jul, 80(3) 253-62 (G08 LM 04419)

Seavey BR, Farr EA, Westler WM, Markley JL A relational database for sequence-specific protein NMR data J Biomol NMR, 1991, 1 217-36 (R01 LM 04958)

Shahar Y, Tu SW, Musen MA Knowledge acquisition for temporal-abstraction mechanisms Knowledge Acquisition, 1992, 4 217-36 (R01 LM 05157 and R01 LM 05208)

Shahar Y, Tu SW, Musen MA Temporal-abstraction mechanisms in management of clinical protocols In Clayton PD, ed, Assessing the value of medical informatics 15th Annual Symposium on Computer Applications in Medical Care, a conference of the American Medical Informatics Association, 1991 Nov 17-20, Washington DC New York McGraw-Hill, [1992] 629-33 (R01 LM 05157)

Shiffman RN, Greenes RA Use of augmented decision tables to convert probabilistic data into clinical algorithms for the diagnosis of appendicitis In Clayton PD, ed Assessing the value of medical informatics 15th Annual Symposium on Computer Applications in Medical Care, a conference of the American Medical Informatics Association, 1991 Nov 17-20, Washington, DC New York McGraw-Hill, [1992] 686-90 (T15 LM 07037)

Sonnenberg FA An intelligent decision system for lung disease in AIDS In Knowledge-based construction of probabilistic and decision models Workshop notes from the Ninth National Conference on Artificial Intelligence, 1991 Jul 16, Anaheim, CA Am Assoc for Artificial Intelligence, [1991] 123-7 (R29 LM 05266)

Stead WW, Borden RB, Boyarsky MW, Crow DS, Mears TP, Stone AA, Woods PJ A system's architecture which dissociates management of shared data and end-user functions In Clayton PD, ed Assessing the value of medical informatics 15th Annual Symposium on Computer Applications in Medical Care, a conference of the American Medical Informatics Association, 1991 Nov 17-20, Washington, DC New York McGraw-Hill, [1992] 475-80 (G08 LM 04613)

Stein LD, Snydr-Michal J, Greenes RA Realistic viewing and manipulation of radiographic images on a personal computer - a user interface for educational applications J Digit Imaging, 1991 Aug, 4(3) 169-76 (R01 LM 04572 and T15 LM 07037)

Sundsten JW, Kastella KG, Conley DM Videodisc animation of 3D computer reconstructions of the human brain J Biocommun, 1991, 18(2) 45-9 (R01 LM 04925)

Szolovits P Compilation for fast calculation over pedigrees Cytogenet Cell Genet, 1992, 59 136-8 (R01 LM 04493)

Tcheng JE, Jackman JD, Nelson CL, Gardner LH, Smith LR, Rankin JS, Califf RM, Stack RS Outcome of patients sustaining acute ischemic mitral regurgitation during myocardial infarction Ann Intern Med, 1992 Jul 1, 117(1) 18-24 (G08 LM 04613)

Tu SW, Shahar Y, Dawes J, Winkles J, Puerta AR, Musen MA A probem-solving model for episodic skeletal-plan refinement Knowledge Acquisition, 1992, 4 197-216 (R01 LM 05157 and R01 LM 05208)

Turner CW, Lincoln MJ, Haug P, Williamson JW, Jessen S, Cundrick K, Warner H Iliad training effects a cognitive model and empirical findings In Clayton, PD, ed Assessing the value of medical informatics 15th Annual Symposium on Computer Applications in Medical Care, a conference of the American Medical Informatics Association, 1991 Nov 17-20, Washington, DC New York McGraw-Hill, [1992] 68-72 (R01 LM 04604)

Vries JK, Marshalek B, D'Abarno JC, Yount RJ, Dunner LL An automated indexing system utilizing semantic net expansion Comput Biomed Res, 1992, 25 153-67 (R01 LM 04635)

Wagner MM An automatic indexing method for medical documents In Clayton PD, ed Assessing the value of medical informatics 15th Annual Symposium on Computer Applications in Medical Care, a conference of the American Medical Informatics Association, 1991 Nov 17-20, Wash , DC New York McGraw-Hill, [1992] 1011-7 (F37 LM 00013 and T15 LM 07059)

Wagner MM, Cooper GF Evaluation of a meta-1-based automatic indexing method for medical documents Comput Biomed Res, 1992, 25 336-50 (T15 LM 07059)

Wellman MP, Doyle J Modular utility representation for decision-theoretic planning In Hendler J, ed Artificial

Intelligence Planning Systems: Proceedings of the First International Conference; 1992; College Park, MD; 1992: 236-42. (R01 LM 04493).

Wu TD. Probabilistic evaluation of candidate sets for multidisorder diagnosis. In: Bonissone PP, Henrion M, Kanal LN, Lemmer JF, eds. Uncertainty in Artificial Intelligence. 6 ed. Amsterdam, North-Holland: Elsevier; 1991: 107-15. (R01 LM 04493).

Yeh A. Automatically analyzing a steadily beating ventricle's iterative behavior over time. Artif Intel Med; 1991; 3: 313-23. (R01 LM 04493).

Zolnai Z, Westler WM, Ulrich EL, Markley JL. Drafting table and light-box software for multidimensional NMR spectral analysis (PIXI). The personal computer workstation. J Magnetic Resonance; 1990; 88: 511-22. (R01 LM 04958).

Zolnai Z, Macura S, Markley JL. Phasing two- and three-dimensional NMR spectra by use of the Hilbert transform can save computer time and space. J Magnetic Resonance; 1990; 89: 94-101. (R01 LM 04958).

APPENDIX 4: REGIONAL MEDICAL LIBRARIES IN THE NATIONAL NETWORK OF LIBRARIES OF MEDICINE

1 MIDDLE ATLANTIC REGION
The New York Academy of Medicine
2 East 103rd Street
New York, NY 10029
(212) 876-8763 FAX (212) 534-7042
States served: DE, NJ, NY, PA

2 SOUTHEASTERN/ATLANTIC REGION
University of Maryland at Baltimore
Health Sciences Library
111 South Greene Street
Baltimore, MD 21201-1583
(410) 706-2855 FAX (410) 706-0099
States served: AL, FL, GA, MD, MS, NC, SC, TN, VA, WV, DC, the Virgin Islands, and Puerto Rico

3 GREATER MIDWEST REGION
University of Illinois at Chicago
Library of the Health Sciences
P.O. Box 7509
Chicago, IL 60680
(312) 996-2464 FAX (312) 996-2226
States served: IA, IL, IN, KY, MI, MN, ND, OH, SD, WI

4 MIDCONTINENTAL REGION
University of Nebraska Medical Center
Leon S. McGoogan Library of Medicine
600 South 42nd Street
Omaha, NE 68198-6706
(402) 559-4326 FAX (402) 559-5482
States served: CO, KS, MO, NE, UT, WY

5. SOUTH CENTRAL REGION
Houston Academy of Medicine-Texas Medical Center Library
1133 M.D. Anderson Boulevard
Houston, TX 77030
(713) 790-7053 FAX (713) 790-7030
States served. AR, LA, NM, OK, TX

6 PACIFIC NORTHWEST REGION
University of Washington
Health Sciences Center Library, SB-55
Seattle, WA 98195
(206) 543-8262 FAX (206) 543-2469
States served: AK, ID, MT, OR, WA

7 PACIFIC SOUTHWEST REGION
University of California, Los Angeles
Louise Darling Biomedical Library
10833 Le Conte Avenue
Los Angeles, CA 90024-1798
(310) 825-1200 FAX (310) 825-5389
States served· AZ, CA, HI, NV and U S. Pacific Territories

8. NEW ENGLAND REGION
University of Connecticut Health Center
Lyman Maynard Stowe Library, ASB-3
263 Farmington Avenue
Farmington, CT 06030-5370
(203) 679-4500 FAX (203) 679-1305
States served: CT, MA, ME, NH, RI, VT

APPENDIX 5: BOARD OF REGENTS

The NLM Board of Regents meets three times a year to consider Library issues and make recommendations to the Secretary of Health and Human Services on matters affecting the Library.

Appointed Members:

COHN, Lawrence H., M.D. (Chair)
Chief of Cardiac Surgery
Brigham and Women's Hospital
Boston, MA

ALLEN, Beverly E.
Director, Multi-Media Center
Morehouse School of Medicine
Atlanta, GA

ANDERSON, Rachael K.
Director, Health Sciences Center Library
University of Arizona
Tucson, AZ

BOOKER, Naomi C.
Chair and President
Marketing and Management Innovations
Baltimore, MD

DeNARDIS, Lawrence J., Ph.D.
President, University of New Haven
West Haven, CT

JOYNT, Robert J., M.D., Ph.D.
Vice President and Vice Provost for Health Affairs
University of Rochester
Rochester, NY

KAHN, Robert E., Ph.D.
President, Corporation for National Research Initiatives
Reston, VA

NEWTON, Carol M., M.D., Ph.D.
Professor of Biomathematics
University of California
Los Angeles, CA

WALKER, H. Kenneth, M.D.
Professor of Medicine
Emory University School of Medicine
Atlanta, GA

Ex Officio Members:

Librarian of Congress

Surgeon General
Public Health Service

Surgeon General
Department of the Air Force

Surgeon General
Department of the Navy

Surgeon General
Department of the Army

Under Secretary for Health
Department of Veterans Affairs

Assistant Director for Biological Sciences
National Science Foundation

Director
National Agricultural Library

Dean
Uniformed Services University of the Health Sciences

APPENDIX 6: BOARD OF SCIENTIFIC COUNSELORS/ LISTER HILL CENTER

The Board of Scientific Counselors meets periodically to review and make recommendations on the Library's intramural research and development programs.

Members:

FRYBACK, Dennis G., Ph.D. (Chair)
Professor, Preventive Medicine and
Industrial Engineering
University of Wisconsin-Madison
Madison, WI

BRUTLAG, Douglas L., Ph.D.
Associate Professor of Biochemistry
Stanford University School of Medicine
Stanford, CA

CIMINO, James J., MD
Assistant Professor of Medicine
College of Physicians and Surgeons
Columbia University
New York, NY

FRISSE, Mark E., M.D.
Associate Professor of Medicine
Washington University School of Medicine
St. Louis, MO

HUNTLEY, Joan S., Ph.D.
Research and Development Project Leader
Weeg Computing Center
Iowa City, IA

LEHNERT, Wendy G., Ph.D.
Professor of Computer and Information Science
Department of Computer and Information Science
University of Massachusetts
Amherst, MA

MUN, Seong Ki, Ph.D.
Director, Division of Imaging Physics
Georgetown University
Washington, D.C.

PETERSON, George D., Ph.D.
Asst. Vice President for Academic Affairs
Morgan Stgate University
Baltimore, MD

APPENDIX 7. BOARD OF SCIENTIFIC COUNSELORS/ NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION

The National Center for Biotechnology Information Board of Scientific Counselors meets periodically to review and make recommendations on the Library's biotechnology-related programs.

Members:

SAUER, Robert T., Ph.D. (Chairman)
Professor, Department of Biology
Massachusetts Institute of Technology
Cambridge, MA

ALONSO, Rafael, Ph.D.
Assistant Professor
Matsushita Information Technology Laboratory
Princeton, NJ

BERMAN, Helen M., Ph.D.
Professor of Chemistry
Department of Chemistry
Rutgers University
Piscataway, NJ

CANTOR, Charles R., Ph.D.
Principal Scientist of the Department of Energy
Human Genome Project
Lawrence Berkeley Laboratory
Berkeley, CA

DEVEREUX, John R., Ph.D.
President, Genetics Computer Group, Inc.
Madison, WI

KELLY, Thomas J., M.D., Ph.D.
Professor and Director
Department of Molecular Biology and Genetics
The Johns Hopkins University School of Medicine
Baltimore, MD

APPENDIX 8. BIOMEDICAL LIBRARY REVIEW COMMITTEE

The Biomedical Library Review Committee meets three times a year to review applications for grants under the Medical Library Assistance Act.

Members:

JAFFE, Conrade C., M.D. (Chair)
Professor of Diagnostic Radiology and Internal Medicine
Yale University School of Medicine
New Haven, CT

ABARBANEL, R.M., M.D., Ph.D.
Manager, Engineering Computing and Analysis
Boeing Computer Services
Seattle, WA

ABOLA, Enrique E., Ph.D.
Chemist
Department of Chemistry
Brookhaven National Laboratory
Upton, NY

CHANDRASEKARAN, B., Ph.D.
Professor, Department of Computer and Information Science
Ohio State University
Columbus, OH

EZQUERRA, Norberto F., Ph.D.
Associate Professor
College of Computing
Georgia Institute of Technology
Atlanta, GA

FENICHEL, Carol H., Ph.D.
Director of the Library and Professor of Information Science
Hahnemann University
Philadelphia, PA

FIELDS, Christopher A., Ph.D.
Laboratory Director
Comparative Genomic Department
The Institute for Genomic Research
Gaithersburg, MD

FULLER, Sherrilynne, Ph.D.
Director, Health Sciences Library and Information Center
University of Washington
Seattle, WA

HAMBERG, Cheryl J.
Director of the Library
Meharry Medical College Library
Nashville, TN

HAYNES, R. Brian, M.D.
Chief, Health Information Research Unit
McMaster University
Hamilton, Ontario

JACKSON, Sara Jean
Director, Research Medical Library
M.D. Anderson Cancer Center
Houston, TX

LOVE, Erika
Director, Medical Center Library
University of New Mexico
Albuquerque, NM

MUSEN, Mark A., M.D., Ph.D.
Assistant Professor
Depts. of Medicine and Computer Science
Stanford University

MYERS, Eugene W., Ph.D.
Associate Professor
Department of Computer Science
University of Arizona
Tucson, AZ

PEAY, Wayne J.
Director, Eccles Health Science Library
University of Utah
Salt Lake City, UT

REYNOLDS, Robert E., M.D., Dr.P.H.
Senior Associate Vice President for Health Sciences
University of Virginia
Charlottesville, VA

ROE, Bruce A., Ph.D.
Professor of Chemistry and Biochemistry
University of Oklahoma
Norman, OK

SIEVERT, MaryEllen C., Ph.D.
Associate Professor
School of Library and Information Science
University of Missouri-Columbia
Columbia, MO

STORMO, Gary D., Ph.D.
Assistant Professor
Department of Molecular, Cellular and
Developmental Biology
University of Colorado
Boulder, CO

VANNIER, Michael W., M.D.
Professor of Radiology
Washington University
St. Louis, MO

WRIGHT, Barbara A.
Director, Library and Information Services
Fayetteville Area Health Education Center
Fayetteville, NC

APPENDIX 9. LITERATURE SELECTION TECHNICAL REVIEW COMMITTEE

The Literature Selection Technical Review Committee meets three times a year to select journals for indexing in *Index Medicus* and MEDLINE.

Members:

DeBAKEY, Lois E., Ph.D. (Chair)
Professor of Scientific Communication
Baylor College of Medicine
Houston, TX

BERG, Alfred O., M.D., M.P.H.
Associate Professor/Director of Research
Department of Family Medicine
University of Washington
Seattle, WA

BLODI, Frederick C., M.D.
Professor Emeritus
Department of Ophthalmology
The University of Iowa Hospital
Iowa City, IA

FURANO, Anthony V., M.D.
Chief, Section on Genomic Function and Structure
Laboratory of Biochemical Pharmacology
Nat. Inst. of Diabetes and Digestive and Kidney Diseases
National Institutes of Health

GOLDBERG, Herbert S., Ph.D.
Associate Dean for Research and Academic Affairs
University of Missouri
Columbia, MO

GROBE, Susan J., R.N., Ph.D.
Professor, Center for Health Care Research
University of Texas School of Nursing
Austin, TX

PARKER, Curtis L., Ph.D.
Professor and Chairman of Anatomy
Morehouse School of Medicine
Atlanta, GA

PASSAMANI, Eugene R., M.D.
Director, Div. of Heart and Vascular Diseases
National Heart, Lung, and Blood Institute
National Institutes of Health

UTIGER, Robert D., M.D.
Deputy Editor
New England Journal of Medicine
Boston, MA

WALTER, Pat L.
Acting Biomedical Librarian
Louise Darling Biomedical Library, UCLA
Los Angeles, CA

WEAVER, William Lynn, M.D., F.A.C.S.
Chief, Surgical Service
VA Medical Center
Buffalo, NY

WILSON, Frank C., M.D.
Professor and Chief of Orthopaedics
University of North Carolina
School of Medicine
Chapel Hill, NC

DISCRIMINATION PROHIBITED: Under provisions of applicable public laws enacted by Congress since 1964, no person in the United States shall, on the ground of race, color, national origin, sex, or handicap, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any program or activity receiving Federal financial assistance. In addition, Executive Order 11141 prohibits discrimination on the basis of age by contractors and subcontractors in the performance of Federal contracts. Therefore, the National Library of Medicine must be operated in compliance with these laws and executive order.