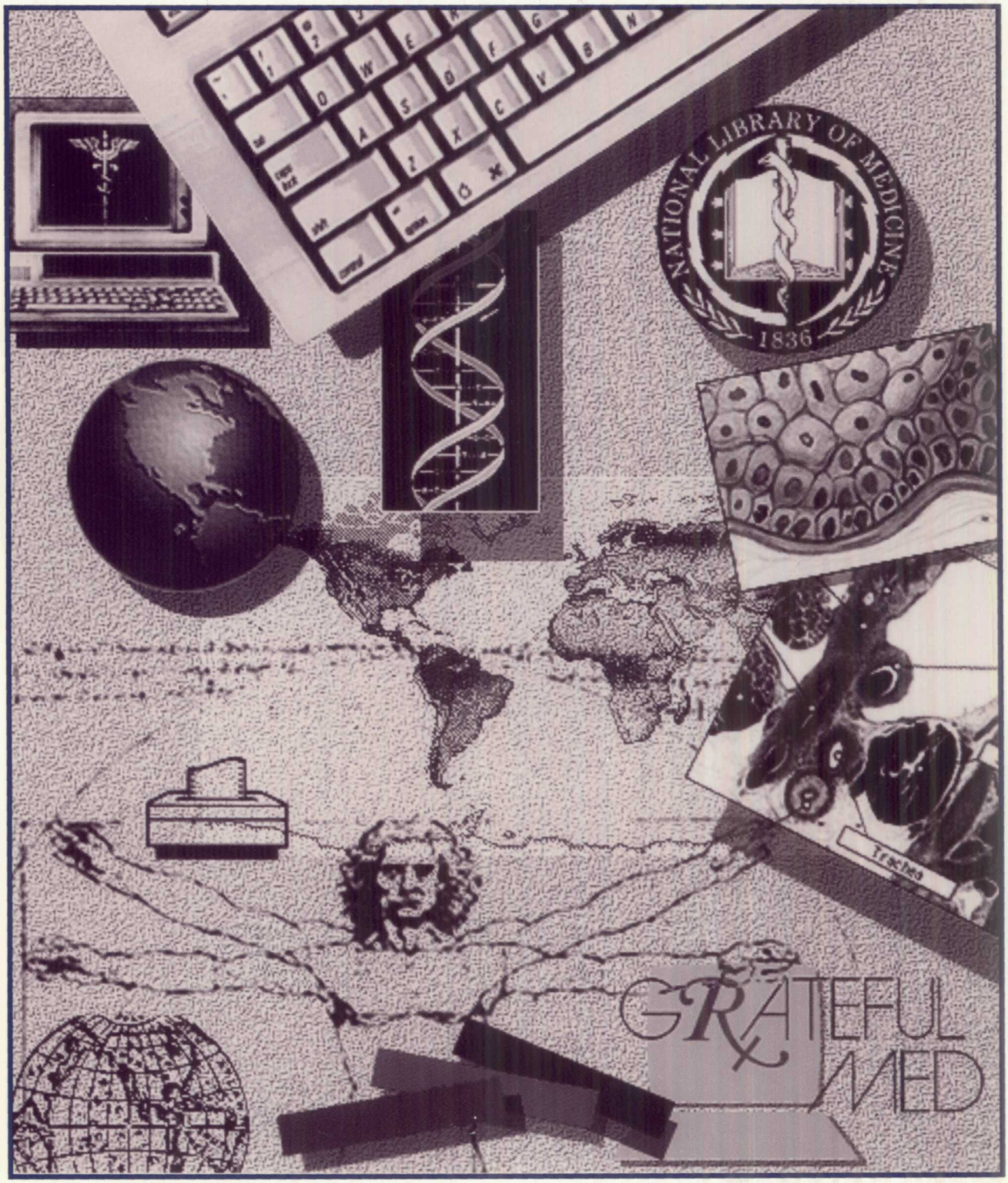
NATIONAL LIBRARY OF MEDICINE

Fiscal Year 1993 Activities



NATIONAL INSTITUTES OF HEALTH • BETHESDA • MARYLAND

The National Library of Medicine was originally established in 1836 as the Library of the Army Surgeon General's Office. In 1862 the Surgeon General moved his offices to the Riggs Bank Building (right), one block from the White House. The building no longer exists. At that time the collection consisted of no more than a few shelves of medical books, although a handwritten catalog of the collection had been prepared in 1840.

In 1866 the Library was moved to Ford's Theater (left), site of President Lincoln's assassination. Located on 10th Street, N.W. in downtown Washington, D.C., Ford's is still used for theatrical productions. One year earlier, in 1865, Dr. John Shaw Billings, a Civil War surgeon, was given charge of the collection. He guided its fortunes over the next 30 years, assembling a substantial medical literature collection of international repute.

NATIONAL
INSTITUTES
OF HEALTH

NATIONAL LIBRARY OF MEDICINE

PROGRAMS & SERVICES
FISCAL YEAR 1993

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

e-Mail: publicinfo@occshost.nlm.nih.gov

CONTENTS

Preface iii

Calendar of Events iv

Board of Regents Biennial Report to Congress 1

Library Operations 5
- Planning and Management 5
- Collection Development 5
- Bibliographic Control 6
- Network Services 7
- Special Onsite Programs 10

Specialized Information Services 16
- TOXNET and Its Files 16
- Databases under ELHILL 18
- Other Programs 19

Lister Hill Center 21
- Computer Science Branch 21
- Information Technology Branch 26
- Communications Engineering Branch 27
- Educational Technology Branch 31
- Audiovisual Program Development Branch 34

National Center for Biotechnology Information 36
- Database Building and Enhancement 36
- Basic Research 38
- Communication 38
- Extramural Programs 39
- Biotechnology Information in the Future 39

Extramural Programs 40
- Overview 40
- Training 41
- Research Grants 42
- Resource Grants 43
- IAIMS Grants 43
- Publication Grants 44
- Bioethics 44
- Small Business Grants 44
- Conferences 45
- Minority Support Activities 45
- Scientific Review 45
- Special Activities 45
- Plans for FY 1994 45

Office of Computer and Communications Systems 47
- Development Branch 47
- Information Management Branch 48
- Systems Branch 50
- Computer Services Branch 50

International Programs 51
Collaboration with Individual Countries 51
International MEDLARS Agreements 51
Collaboration with the W.H.O. 52
Special Foreign Currency Program 52
International Meetings and Visitors 52

Administration 54
Financial Resources 54
Personnel 54
Appointments 54
Awards 54

Appendixes
1. Acronymns, Abbreviations, and Initialisms 56
2. Staff Bibliography 60
3. Extramural Programs-Supported Publications 65
4. Regional Medical Libraries 70
5. Board of Regents 71
6. Board of Scientific Counselors/LHC 72
7. Board of Scientific Counselors/NCBI 73
8. Biomedical Library Review Committee 74
9. Literature Selection Technical Review Committee 76

Tables

Table 1. Growth of Collections 11
Table 2. Acquisitions Statistics 11
Table 3. Cataloging Statistics 12
Table 4. Bibliographic Statistics 12
Table 5. Circulation Statistics 12
Table 6. Online Searches 13
Table 7. Offline Searches 14
Table 8. Reference Services 14
Table 9. History of Medicine Activities 15
Table 10. Extramural Grants and Contracts Program 46
Table 11. International MEDLARS Centers 52
Table 12. Financial Resources and Allocations 54
Table 13. Staff 55

Preface

For the first time, this report includes the biennial report to the Congress of the NLM Board of Regents. The Board's account is an excellent summary of those recent activities that have been of the highest priority to the Library in fulfilling its mission.

Probably no action taken by the Library in Fiscal Year 1993 was more popular than the January 1, 1993, lowering of fees for searching the MEDLARS databases. The 40 percent reduction was no doubt partially responsible for the record rate of online searching in FY 1993—almost 6 million searches. It may also have contributed to the remarkable 1-year increase of 15 percent in the number of MEDLARS access codes—to about 75,000 by the end of the year.

A number of significant events this year involved the Internet:

The introduction of Locator, NLM's online catalog system, to patrons and Internet users;

The making available of selected NLM publications via File Transfer Protocol and Internet;

The extension of Internet access to MEDLINE via Grateful Med to NIH scientists and staff;

The signing of a MEDLARS agreement with Israel whereby that country becomes the first to rely exclusively on the Internet for access; and

The extension of the "connection grants" to link up hospitals.

Concurrent with the emphasis on Internet access is the continued importance of the various aspects of the High Performance Computer and Communications initiative. The NLM solicited applications for health-related HPCC projects in spring 1993, and, just before the fiscal year ended, an award was made to a West Virginia consortium for a statewide project, "Collaborative Technology for Real-Time Treatment of Patients." We are hopeful that more HPCC awards will be made in 1994.

One of the high points of FY 1993 occurred early in the year when thousands of administrators, librarians, and other staff members of hospitals around the country viewed a 2-part satellite broadcast in October and November, 1992. The broadcast, sponsored by the NLM and the Medical Library Association, was an overview of the information resources and services available through hospital libraries and discussions by some 40 experts of contemporary issues affecting the utilization of this information.

Beginning on October 1, 1992, the National Center for Biotechnology Information assumed responsibility for GenBank, the DNA sequence database. GenBank collects all known DNA sequences and is a critical tool in research relating to the analysis and discovery of gene function. As readers of the NCBI chapter will see, usage of GenBank has increased dramatically in the last year, as access is provided via Internet FTP, e-mail servers, and CD-ROM.

I would like to take this opportunity to thank those who serve on NLM's various advisory groups for their much appreciated assistance and also the thousands of members of the National Network of Libraries of Medicine whom we are glad to help serve the information needs of American health professions.

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

Calendar of Events - Fiscal Year 1993

1992
Oct 1: GenBank now provided by NLM
Oct 1-2: Board of Regents
Oct 15-16: Literature Selection Technical Review Committee
Oct 22: Hospital satellite broadcast
Oct 29-30: LHC Board of Scientific Counselors
Nov 5: International MEDLARS Policy Advisory Group
Nov 5: Hospital satellite broadcast
Nov 12-13: Biomedical Library Review Committee
Nov 19: NCBI Board of Scientific Counselors
Nov 24: NLM honor awards ceremony
Dec 14: Dr. C. Everett Koop, former Surgeon General, donates papers to NLM

1993
Jan 1: NLM lowers fees for online MEDLARS access
Feb-Apr: The Proud Profession: Nurses in the Federal Service (exhibit)
Feb 11: HPCC Conference: Improving Medical Care
Feb 11-12: Board of Regents
Feb 22: NLM Locator, online catalog system, installed for patrons
Mar 3-4: Biomedical Library Review Committee
Mar 4-5: Literature Selection Technical Review Committee
Apr 1: NLM/American College of Physicians flat rate agreement extended
Apr 16: 25,000th record added to AVLINE
Apr 16: Israel becomes a MEDLARS Center
May-Aug: The History of Cesarean Section (exhibit)
May 13-14: LHC Board of Scientific Counselors
May 26-27: Board of Regents
Jun: Grateful Med for the Macintosh, Version 2.0, released
Jun 2: NLM-sponsored 1-week medical informatics fellowship program at Woods Hole
Jun 15: EEO Career Day
Jun 17-18: Literature Selection Technical Review Committee
Jun 23-24: Biomedical Library Review Committee
Jun 25: HPCC conference in the Capitol
Jun 28-30: NIH AIDS/HIV Information Services Conference
Sep 7-8: NCBI Board of Scientific Counselors
Sep 20: NIH Research Day
Sep 28-29: Planning Panel on the Education and Training of Health Sciences Librarians
Sep 30-Oct 1: Board of Regents

BOARD OF REGENTS BIENNIAL REPORT TO CONGRESS

Note: This Biennial Report of the National Library of Medicine Board of Regents is the first to be included in the annual NLM Programs and Services.

Introduction

The NIH is a powerful engine creating and discovering new knowledge of human health and disease. Making that knowledge available to health professionals and to all who seek to protect their health is the important final link in the process of discovery, and is fundamentally critical to the prevention and treatment of disease. This is the mission of the National Library of Medicine. This mission is particularly important in the environment of the 1990s, when the health care industry is in the throes of health care reform propelled by serious economic pressures and inadequate health care for a growing number of Americans. Medical education is undergoing a sea change, with its increased emphasis on producing more generalist physicians and the development of new approaches to learning that support problem-solving and lifelong learning. Biomedical research, while under increased scrutiny from both financial and ethical perspectives, continues to produce astonishing medical discoveries—discoveries that generate unprecedented amounts of new biomedical knowledge which must be disseminated to practitioners. Finally, the rapid advances in computer and communications technology create opportunities as well as challenges in harnessing its potential to improve health care delivery, education and research.

These challenges and opportunities require that the nation invest in High Performance Computing and Communications (HPCC) technologies. The HPCC initiative will result in powerful computers to store and analyze biomedical scientific data and will extend high speed computer networks to create a national electronic community of health professionals. Over these high speed digital highways scientists can quickly collaborate and share new findings, and practicing health professionals can instantly access medical databanks and exchange clinical data for the efficient care of patients. The major biomedical advances of the 1990s—structural biology, the human genome, gene therapy, immunology, biotechnology, etc.—depend upon and benefit from this evolving computing and communications infrastructure. The environment of the 1990s and beyond will demand more broad-based and timely communication modes and outreach initiatives that meet the needs of the multiplicity of audiences, including research scientists, health practitioners, educators, students, and a wide variety of health professionals located in both the public and private sectors. To this end, major ventures are underway to improve access to information which can prevent disease and alleviate suffering in minorities, women, and children, to discover the genetic basis of disease, and to ensure that the findings of health services research can be translated into better, more affordable health care.

Summary and Recommendations

The Board of Regents is well pleased with the achievements of the Library in meeting the goals of its own Long Range Plan. Much progress has been made, and, with continued support, new and enhanced information services have a tremendous potential for improving health care and biomedical research. Specific recommendations of the Board include:

1. The Board recommends strong support for basic library services as the necessary underpinning of the Library's national responsibilities to support medical research, practice, and education.

2. The Board applauds NLM's action in eliminating all online charges for searching AIDSLINE, AIDSDRUGS, AIDSTRIALS, and DIRLINE, and points to NLM's activities in HIV/AIDS information as a model for response to high-priority areas of concern.

3. The Board recommends increased support for NLM's outreach programs as a way to improve the health and well-being of the American people through the expeditious communication of new knowledge to underserved populations.

4. The Board recommends increased support for NLM's High Performance Computing and Communications programs to help the biomedical community to benefit from the power and usefulness of new modes of communication.

5. The Board recommends increased support for the biotechnology information databases and tools on which the field of biotechnology depends for its survival and growth.

6. The Board recommends increased support for NLM's toxicology and environmental health information programs.

7. The Board strongly supports the work of the NLM Planning Panel on the Education and Training of Health Science Librarians and underscores the importance of expediting the panel's work with respect to defining the near and long-term responsibilities of the Library.

Basic Library Services

The basic mission of the National Library of Medicine is to acquire, preserve, and provide access to books, journals, and other materials pertinent to medicine. Libraries throughout the nation and databases such as MEDLINE and GenBank depend upon NLM's collection and access services. The success of NLM's outreach efforts has increased the demand for these basic services. Although NLM has made progress in reducing the per transaction cost of many basic services, greater demand increases overall costs. Rising literature prices and an increasing array of material formats that must be acquired and preserved mean that NLM must spend more to maintain a comprehensive back-up collection for current researchers and practitioners and for the scholars of the future. At the same time, escalating journal prices are forcing health science libraries throughout the nation to cancel subscriptions and therefore to increase their dependence on NLM's collection as the "court of last resort." If NLM fails to keep up with acquisition, preservation, and provision of access to the published literature, the result will be a decline in the amount, quality, and timeliness of information services to current health professionals and scholars in the future.

HIV/AIDS Information

The AIDS pandemic is inextricably linked to the ability to obtain needed information that can "make a difference." The National Library of Medicine (NLM) and the Office of AIDS Research (OAR) of the National Institutes of Health (NIH) co-sponsored an invitational conference on June 28-30, 1993 to examine the role that the NIH is playing in providing that information. The meeting involved broad-ranging discussions among members of the various constituent groups who need HIV/AIDS information. The conference report contains recommendations that are intended to reflect the views of the HIV/AIDS community of users. Within the constraints of NIH's mission and resources, it is expected that implementation of these recommendations will result in the provision of new and enhanced HIV/AIDS information services by the various components of the NIH. They are also intended to foster new opportunities for collaboration within the NIH, and with other agencies of the U.S. Public Health Service and non-governmental organizations.

An immediate action taken by NLM was the elimination of all online charges for searching AIDSLINE, AIDSDRUGS, AIDSTRIALS, and DIRLINE, effective January 25, 1994. This is in response to concerns voiced at the HIV/AIDS information services conference that even the existing modest fees were a financial burden that was inhibiting access to AIDS-related information.

Improving Access to Biomedical Information

Improving access to biomedical information, HIV/AIDS-related and otherwise, is a major priority of NLM and its Board of Regents. The NLM outreach program is a cooperative effort with the 3600 member institutions of the National Network of Libraries of Medicine (NN/LM). NLM has initiated more than 200 outreach projects, involving nearly 400 institutions, since the publication of the DeBakey report in 1989. They include extensive efforts to train physicians and other health professionals to use Grateful Med, through supported projects at the Regional Medical Libraries (RMLs) and small-to-medium sized network libraries to improve both local resources and access to online information. There is a special emphasis on libraries in rural and inner-city areas and those that serve minority populations. For example, NLM provided special funding through the RMLs for 13 library improvement projects in the Lower Mississippi Delta to provide support and equipment to libraries in this region that do not have access to NLM databases and that are currently providing less than up-to-date service to their users. NLM also supports an Undergraduate Research Study Program to stimulate undergraduate medical informatics research programs in Historically Black Colleges and Universities (HBCUs), and an initiative in the Toxicology Information Program to strengthen 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.

NLM has also begun a new phase in outreach, concentrating efforts first in six southern states (Arkansas, Louisiana, Mississippi, South Carolina, Tennessee and Texas). The purposes of this effort are (1) to establish linkages with key governmental, health, academic and local community organizations that are concerned with the health needs of pregnant women and children in the southern states, and particularly those states in the Lower Mississippi Delta (a 214-county, 7-state region characterized as the most impoverished area of the nation); and (2) to identify ongoing and planned health programs in which access to NLM's databases could become an integral part of the locally originated programs. A special emphasis is placed on programs that serve minority and rural populations. Meetings have been held in each of the states to bring together key state leaders to collaborate with NLM in planning strategies and activities for the individual states. The Southern Institute on Children and Families provided invaluable assistance in arranging the meeting sites and lists of invitees. If this new approach is successful in building Federal-State linkages, NLM will look toward extending it to other health care areas and to other regions of the country. As a direct result of the "linkage" meeting held in Austin, the Texas Department of Health will

carry out an experimental information access project in the border counties of South Texas. This region has a large Hispanic population that is particularly vulnerable to environmental contaminants. Access to authoritative and timely information by health professionals serving these communities could have a significant impact on identifying, treating, and preventing morbidity and mortality arising from such exposures.

High Performance Computing and Communications

The High Performance Computing and Communications (HPCC) program is creating powerful technologies that have the potential to prevent illness, reduce health care costs, and reduce the economic burden of disease. For example, the cost and inefficiencies of paper-based medical records can be reduced by creation of database and communications technologies for storing, accessing, and transmitting patients' medical records while protecting the accuracy and privacy of those records. HPCC technologies will also allow the creation of high speed digital networks that will enable several health care providers in remote locations to provide real-time treatment to patients. These networks will transmit all kinds of clinical data, including text, images such as x-rays, and signals such as EKG recordings, enabling care providers and researchers to share this varied medical data and imagery almost instantly. Efficient and accurate diagnosis will be assisted by visualization technology for converting two dimensional imagery from X-rays, CAT scans, PET scans, and other diagnostic tools into three dimensional "living images" of internal organs, which can be viewed from any angle on the computer screen. Finally, new database technologies will provide health care providers with access to multiple sources of relevant medical information and literature simultaneously, from systems whose locations and structure need be of no concern to the provider, potentially improving the accuracy and speed with which diagnostic and treatment decisions are made.

Biotechnology Information and Computational Tools

Concerted research by several groups in the National Center for Biotechnology Information (NCBI) has focused on new analysis and search methodologies to cope with the explosive growth in sequence information. One technique is the derivation of a set of some 900 ancient, evolutionary-conserved regions in proteins. This has demonstrated a core set of molecular functions common to diverse biological systems that should prove invaluable in the prediction of function for unknown sequences. Other work has indicated that many highly-variable protein regions with simple, biased compositions and short period repeats are particularly abundant in proteins important in morphogenesis, gene regulations and biological structural integrity. Finally, a new strategy has been developed to recognize three-dimensional structural motifs in protein sequences by calculating empirical energy potentials and 'threading' the sequences through backbone folds found in the structural database. All of these basic research findings have been transferred into flexible, easy-to-use software tools and made available to the biomedical research community.

Most biological data resides in disconnected databases in incompatible forms, prohibiting the discovery of underlying relationships in the data. NCBI has pioneered a new, integrated approach to managing and linking diverse sets of data and presenting them in a cohesive, single system. NCBI has produced a data resource that contains data from seven nucleic acid sequence databases, three protein sequence databases, a 3-D protein structure database, and the MEDLINE literature database. Software tools have been distributed to researchers that brings this integrated functionality to the laboratory. Network-based, retrieval software uses the Internet to bring this integrated information resource to the lab bench.

Toxicology and Environmental Health

The report of a Planning Panel on Toxicology and Environment Health was recently published as part of the Board of Regents' Long Range Plan. Since its inception in 1967, the goals of NLM's Toxicology Information Program have been quite straightforward: to create and maintain automated data banks of information on toxicological subjects and to disseminate that information widely. In the intervening years, "toxicology" has developed meanings and societal importance far beyond those envisioned 24 years ago. In parallel, the demand for access to comprehensive, accurate information about the subject has expanded rapidly. There is growing awareness of the dangers associated with the release of hazardous chemicals into the environment. Dramatic disasters in locales like Bhopal have alerted the world community to the acute toxicity of certain chemicals. More recent reports have highlighted the devastating effects of prolonged exposure of populations in Eastern Europe to industrial contaminants. Public scrutiny is increasingly directed to more subtle hazards to populations exposed to low doses of marginally harmful agents over long periods of time.

In the planning process, the Board found that although modest in resources, the TIP is ambitious in scope and rich in accomplishments. In many cases these accomplishments are directly attributable to a high degree of cooperation with other Federal agencies that brought with it not only expertise and guidance, but funds. The Plan reviews whether the needs of the users of toxicology and environmental health information services are being met, and posits new initiatives needed to respond to changing circumstances and to take advantage of new opportunities.

Implementation of the report's 16 recommendations will help ensure that core information resources in toxicology and environmental health are available and that their contents

meet user needs; that they are used efficiently and effectively; and that the information infrastructure needed for future scientific discovery in the field is developed and put in place.

Planning Panel on the Education and Training of Health Science Librarians

The 1986 NLM Long Range Plan, along with its updates on Outreach (1989), Electronic Imaging (1990), and Toxicology and Environmental Health (1992), recognized the importance of developing a cadre of highly trained health sciences librarians, and other information professionals, to adapt modern information technologies to the needs of the biomedical community. This requirement is also a major underpinning of a key component of the Presidential initiative in High Performance Computing and Communications, which is concerned with the training needs of individuals capable of creating and utilizing emerging computing and networking technologies in the national interest. *Platform for Change*, a recent planning document prepared by the Medical Library Association, lists areas in which health science librarians believe they will need proficiency in the future.

In the era of health care reform, the role of librarians will change dramatically as the institutions in which they work also undergo dramatic change. New health care networks will require a vast array of information, including clinical and epidemiological data, medical literature and other "knowledge-based information," as well as financial and administrative data. The manager of this information must be able to integrate various sources of information from local and distant sites (often via advanced communications and networking technologies), evaluate its usefulness, and provide it to the user in a timely fashion.

The purpose of this panel is to analyze the possible programs and activities of the NLM, of individuals, of professional associations, and of other institutions that might be undertaken over the next ten years in order to assure that our society benefits from the skills of health science librarians. NLM also wishes to assure that persons who choose health science librarianship will be properly educated and trained, and that they have an opportunity to engage in the most important work concerning information and health care. The Panel's report will be released in 1994, and the Board of Regents urges that it be given serious consideration when it is published.

LIBRARY OPERATIONS

Lois Ann Colaianni
Associate Director

NLM's Library Operations Division (LO) is responsible for the basic library services that provide an essential foundation for the Library's Outreach program and for a number of special initiatives such as those in biotechnology, AIDS, and health services research information. LO's core programs center on the world's biomedical literature: its acquisition and preservation; its organization through indexing and cataloging; the dissemination of NLM's authoritative bibliographic data online, in machine-readable products; and document delivery, reference, and research assistance provided as a back-up to services available from other U.S. health sciences libraries. LO also coordinates the National Network of Libraries of Medicine (NN/LM), conducts research and evaluation related to NLM's basic services, maintains an active research program in the history of medicine, and directs or participates actively in a number of NLM-wide initiatives.

The largest of NLM's Divisions, LO employs 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; three smaller units: the Medical Subject Headings (MeSH®) Section, the NN/LM Office, and the new National Information Center on Health Services Research and Health Care Technology; and a small administrative staff in the Office of the Associate Director.

Planning and Management

During FY 1993, LO examined a variety of ways that the Internet might be used to improve its programs and services. LO also continued to review current operations to identify any personnel and other resources that might be reallocated to new or enhanced services. The program adjustments that resulted from these efforts are reported elsewhere in this chapter.

Collection Development

To build NLM's comprehensive collection of biomedical literature, LO establishes and revises literature selection policy, acquires and processes relevant biomedical literature in all formats and languages, and maintains and preserves materials acquired. As of September 30, 1993, the Library owned 2,076,702 printed books, journal volumes, theses and pamphlets and 2,893,009 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.* Early in 1993, NLM published a new edition of the *Manual,* which reflects the comprehensive review and revision of selection criteria completed in 1992. The *Manual* is available in printed form from the National Technical Information Service and in electronic form via Internet ftp (file transfer protocol) from the nlmpubs server. One of the goals of the revision was to reduce overlap among the national libraries by focusing NLM's collecting more sharply on biomedical subjects. Application of the revised guidelines has led to a modest reduction in the number of monographs selected for the NLM collection in FY 1993.

NLM conducts periodic collection assessment studies to determine how successfully it is applying its collection guidelines and to identify subject areas in which the guidelines may need revision or clarification. In FY 1993, NLM Library Associates conducted collection assessments in the fields of herbal medicine and radiation exposure. These studies revealed that NLM's collection is generally excellent in both subject areas, but also identified some aspects of NLM's selection criteria that require clarification. The methodology developed for the herbal medicine study may provide a model for assessing NLM's coverage of other alternative therapies. The radiation exposure study involved automated comparison of data in CATLINE with data obtained from CD-ROM bibliographies such as Books-in-Print. The methodology developed has potential for application to other subject areas.

Acquisitions

The LO staff received and processed 177,751 modern books, serial issues, audiovisuals, and software packages (Table 2) and added 31,801 volumes and 13,406 other items (e.g., audiovisuals, microforms, software, pictures, manuscripts) to the collection in FY 1993 (Table 1). Significant additions to NLM's rich historical collections included: the second edition of Andreas Vesalius's *Zergliederung dess Menschlichen Korpers* (Augsberg 1723) which features the original wood-blocks from the *De Humani Corporis Fabrica* (Basel, 1543); two important works by Paracelsus, *Dreyzehen Bucher* . . . (Basel, 1571) and *Vom ursprung und herkommen des Bad Pfeffers* (Basel, 1576), appropriately acquired during the quincentennial year of their author's birth; four 18th century broadsides from the Hamburg Plague Hospital (1743, 1745, 1746, and 1750) which include engravings of the hospital; a seventeenth century pharmaceutical manuscript, dated 1697, which includes a Latin/German dictionary of ingredients and compounds and a section on experiments and remedies; an extraordinary eighteenth century herbal with more than 200 hand-drawn and colored pictures of plants used in herbal medicines by its author, Andrea Di Petris, a physi-

cian who lived near Padova; and a volume of midwifery lecture notes, circa 1775, from lectures given by Colin Mackensie, a student and assistant of William Smellie, and William Lowder. Both lecturers were in the forefront of eighteenth century obstetrics, but their lectures survive only in manuscript volumes such as this one.

During FY 1993, the LO staff analyzed document delivery data to identify additional infrequently requested *Index Medicus* journals for which second copy subscriptions could be canceled. By canceling additional second copies and modifying some overseas shipping arrangements, NLM was able to limit the growth of its serials budget.

The NLM automated systems that support acquisition, processing, and preservation of literature continue to be modified to reduce processing time and to improve the quality of information available for management of both the literature budget and the collection.

Collection Preservation and Maintenance

To preserve and maintain its collection, NLM binds incoming journal issues, microfilms brittle volumes, provides conservation treatment for rare and unique items, and maintains appropriate storage facilities and conditions for all types of library materials. The Library also promotes the use of permanent paper in new biomedical publications to reduce the magnitude of the future preservation problem and explores the use of new technology in the preservation of library materials.

In FY 1993, NLM microfilmed 2.3 million brittle pages and gave conservation treatment to 254 items from the special collections. Volumes microfilmed included some from *Index Medicus* journals for which commercially available microfilm was found to be unacceptable in quality, as well as early volumes of *Index Medicus* itself. NLM's binding program was temporarily disrupted when the commercial binding facility the Library uses was struck by a tornado.

The Library installed a new book security system and additional compact shelving for older serial volumes. In FY 1993, the decision was made to shelve incoming books by accession number rather than by classified call numbers in NLM's closed stacks. This approach will allow the Library to compress its book collection and should provide adequate space for growth of the book and audiovisual collections through the year 2004. The NLM staff will continue to assign the classification portion of the call number for other libraries which make use of NLM catalog records. In FY 1993, NLM contracted for underground storage in a low-temperature vault for off-site safekeeping of historical motion pictures.

The revised American National Standard NISO/ANSI Z39.48-1992 (Permanence of Paper) was published in early FY 1993, ending a 6-year revision effort directed by NLM staff members. The Library continues to encourage publishers of all journals submitted for consideration for indexing in MEDLINE to use acid-free paper. About 47% of current *Index Medicus* titles are known to be printed on alkaline paper.

Bibliographic Control

To allow users to identify relevant information contained in the world's biomedical literature, NLM maintains the Medical Subject Headings (MeSH®) thesaurus for use in describing the subject content of indexed and cataloged items, develops the NLM Classification for subject arrangement of books on library shelves, and produces authoritative cataloging and indexing records for newly published or acquired items.

Thesaurus

MeSH, NLM's hierarchical thesaurus, now contains 17,678 headings. MeSH's supplementary chemical file includes about 73,500 additional records for substances. In FY 1993, NLM added 716 new MeSH headings and 899 new entry terms and updated the terminology for 263 existing headings to keep pace with developments in biomedicine and changes in the usage of biomedical terms. The additions included new and updated terminology for oncology, virology, and receptors. The MeSH hierarchies for oncology were revised to conform to the current edition of the International Classification of Diseases - Oncology (ICD-0). The virology hierarchies were modified to follow the updated classification from the International Committee on Taxonomy of Viruses.

MeSH is a major component of the UMLS® Metathesaurus, and the MeSH Section staff plays a key role in editing additions to the Metathesaurus. The 1993 edition of the Metathesaurus incorporated all preferred names of diseases from the *International Classification of Diseases, 9th edition Clinical Modification* as well as other significant clinical additions. A long-term goal is to develop a merged MeSH/Metathesaurus editing system that will also meet NLM's diverse requirements for maintaining information about chemical substances. In FY 1993, information from the Metathesaurus was transferred automatically to MeSH for the first time in an important initial step toward the unified maintenance environment. Work also progressed on plans to incorporate data from the CHEMID file into the Metathesaurus and on the development of specifications for the new vocabulary maintenance systems.

Cataloging

NLM catalogs books, journals, and nonprint materials both to describe what is available in the Library's own collection 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 1993, the Library cataloged 21,835 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. Since the number of items cataloged exceeded the number of new items received that require cataloging, NLM's inventory of uncataloged titles was reduced by 1,014. The Cataloging Section also revised procedures to improve cataloging throughput for new materials.

In FY 1993, the History of Medicine Division completed a project to create CATLINE records for NLM's distinguished collection of early Arabic and Persian manuscripts. In another project, a team of catalogers completed a draft of a major revision to the NLM Classification. The draft will be sent to a number of NN/LM librarians for review and comment before its scheduled publication in 1994.

Indexing

To provide effective subject access to recent developments in biomedicine, NLM indexes articles from more than 3,000 journals. LO also indexes gene sequences for incorporation into the National Center for Biotechnology Information's backbone database (see NCBI chapter). If indexed articles are retracted, corrected, or challenged in subsequently published commentaries, NLM updates and annotates the indexed citations to these articles accordingly.

The Literature Selection Technical Review Committee (LSTRC—see Appendix 9 for list of members) advises NLM on the journals that should be indexed in MEDLINE, *Index Medicus*, and other NLM databases. In FY 1993, the LSTRC reviewed 331 journals and rated 55 sufficiently high for NLM to begin indexing them; six currently indexed titles were recommended for deletion. After considering recommendations prepared by professional societies in the fields of nursing, urology, and pediatrics, the LSTRC advised NLM to begin indexing 16 titles and to cease indexing 25 titles in these subject areas. As part of an ongoing interest in improving knowledge of characteristics of the literature NLM indexes, LO staff members conducted a study of the peer-review practices of selected *Index Medicus* journals and directed a Library Associate project to examine the extent to which journals produced in developing countries publish the work of local or regional authors.

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 1993, 376,000 indexed citations were added to MEDLINE. Seventy-five percent of the citations added included English-language abstracts. NLM updated previously indexed citations in the MEDLARS databases to reflect information 19 retractions, 3,504 published error notices, and 29,009 substantive commentaries. Two publication types, Meta-Analysis and Interview, were added to retrospective citations. The Library also worked with the NIH Office of Medical Applications of Research to develop a plan for identifying and tagging retrospective randomized clinical trials. During FY 1993, the Library instituted new procedures designed to improve indexing throughput for journal issues sent out to contract indexers.

The Index Section completed its phased acquisition of new PC workstations for inhouse indexers and tested a new PC-based citation maintenance system. During FY 1993, NLM conducted an initial test of OCR scanning as an alternative to keyboarding citations and abstracts. The results of the experiment showed that this approach is still more costly than keyboarding, but may, with further refinement, be useful for some inhouse keyboarding tasks. The Library also continues to experiment with the use of machine-readable bibliographic data received from publishers as a potential alternative to keyboarding.

Network Services

To promote rapid access to the world's biomedical literature, NLM distributes its indexing and cataloging data in publications, machine-readable formats, and an online retrieval service; provides reference assistance to onsite and remote users; delivers document to onsite users and to remote requesters who need items from the NLM collection to supplement the resources of other U.S. libraries; and directs 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 1993, NLM produced more than 90 individual issues of some 22 recurring indexes, catalogs, and other publications, ranging from *Index Medicus* to specific subject bibliographies. The MeSH publications were changed so that MeSH main headings now print in upper and lower case. Staff members in the Reference Section and the Office of Health Services Research Information produced two special issues of *Current Bibliographies in Medicine* as part of NLM's health services research information program. One includes pre-1992 citations to clinical practice guidelines; the other covers literature on the technique of Meta-analysis. Other *Current Bibliographies* addressed disease prevention, seafood safety, early identification of hearing impairment, and health care for women.

In FY 1993, NLM established the nlmpubs anonymous ftp server as an additional mechanism for distribution of the NLM Fact Sheets and selected publications including *Current Bibliographies in Medicine*, the *AIDS Bibliography*, the *Collection Development Manual of the National Library of Medicine*, and new chapters of the *Online Services Reference Manual*. The number of publications available on the server will increase in FY 1994.

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 which make the data available online or in CD-ROM products. In FY 1993, NLM distributed more than

7,500 tapes of various databases to more than 100 different licensees. At the end of the year, 13 licensees were producing CD-ROM products containing data from 6 different MEDLARS databases.

In FY 1993, the Library implemented a revised pricing structure for tape licensees that eliminates use charges, i.e., tape licensees now pay to lease the data, but there are no additional charges based on how much the data they lease is used. Non-U.S. users still pay a surcharge for use of the data since their tax dollars are not contributing to data collection.

Online Services

NLM offers direct online access to about 50 databases. In FY 1993, online users performed about 6 million online searches of these databases during 324,000 online connect hours (tables 6 and 7). These figures exclude online searching performed on the computer systems of organizations that lease data from NLM.

As a result of special outreach efforts by NLM and health science libraries throughout the NN/LM, the number of users of NLM's online system continues to grow rapidly. At the end of FY 1993, there were 71,693 active codes for use of NLM's online system, an increase of 22 percent from FY 1992. NLM established a new flat-rate Internet access agreement for NIH employees and extended the flat-rate agreement with the American College of Physicians. Additional Internet access flat-rate agreements are planned for FY 1994. Due to a revised policy regarding the categories of costs that must be recovered, NLM's online charges dropped about 49% in January 1993. The charge for an average search is $1.25.

Most individuals who received codes in FY 1993 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 62,707 copies (52,808 IBM PC version; 9,899 Macintosh). Purchasers receive new versions of the software at no extra charge. In FY 1993, NLM issued Version 2 of the Macintosh Grateful Med. New features include the ability to access NLM's databases via the Internet and the Loansome Doc document request capability.

During FY 1993, the Library implemented a number of new retrieval system enhancements, including an ELHILL to TOXNET gateway, larger and fewer backfiles, a shortened ELHILL "stopword" list, improved searching of grant numbers, and better coverage of Chemical Abstracts Service Registry Numbers in MeSH records for chemicals. NLM also beta-tested the ftp distribution of current-awareness search results.

About 80 end users and librarians participated in a five-month beta-test of the new HSTAR (Health Services and Technology Assessment Research) database. HSTAR will become available to all NLM online users in 1994. An initial test version of HSRProj (Health Services Research Projects), a database of ongoing research projects, received alpha-testing. This prototype database was created by the Foundation for Health Services Research with funding from the Pew Charitable Trust. NLM will assume responsibility for distribution and maintenance of HSRProj in 1994. LO staff members also worked with the National Aeronautics and Space Administration to define a NASA/NLM Space Medicine and Life Sciences database and with other NIH components to plan possible new databases covering women's health and NIH-funded clinical trials.

In FY 1993, a voice-mail system was added to the online service desk, cutting the number of abandoned calls in half. A total of 1,511 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. NLM and the RML online training centers examined a number of potential approaches to reducing the cost and improving the effectiveness of the online services training program. In FY 1994, NLM will update searchers on new system features via satellite broadcast.

Reference Services

NLM provides reference service and assistance to onsite users and to remote requesters as a backup to the service available from other U.S. health sciences libraries. In FY 1993, NLM's Reference Section responded to 71,070 requests for reference assistance, from onsite and remote requesters (table 8). The Section established an Internet address specifically for reference inquiries. In addition, 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.

Although LO staff members provided direct literature search support to the first 13 panels established by the Agency for Health Care Policy and Research (AHCPR) to develop clinical practice guidelines, NLM has changed its role to that of consultant to the new guideline panels established in FY 1993. LO staff members wrote the chapter on literature searching in the manual provided to new guideline panels. LO representatives advise each panel on the literature search process and also arrange for the panel to receive direct literature search and document delivery service from other NN/LM members. LO continues to prepare bibliographies in support of the NIH Consensus Development Conferences.

One of the tools available in NLM's Reading Room is MEDSTATS, an NLM expert system that helps users identify items in the NLM collection which can answer specific statistical questions. In FY 1993, the MEDSTATS software was upgraded and its content updated. MEDSTATS will be tested by several other network libraries in 1994.

NLM Locator, a new online public access catalog interface (OPAC), to CATLINE, AVLINE, and SERLINE was implemented in FY 1993 for onsite users and Internet access. A joint project of LO and the Office of Computer and Communications Systems, the new client-server interface has been received enthusiastically by both onsite and remote

users. NLM Locator appears as a menu item on the NIH gopher and in the widely distributed "Libraries of the World" gopher menu option under USA, Maryland.

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 522,472 interlibrary loan and onsite requests for post-1913 documents in FY 1993. The History of Medicine Division received 14,114 requests for items from the historical collections, including orders for copies of pictures. NLM received 307,481 modern interlibrary loan requests and filled 72 percent of them. If requests for which the requester was unable to pay are excluded, NLM's fill rate was 77 percent. Sixty-four 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. Five percent or 15,274 requests were received via telefacimile transfer; of these, 2,037 were needed for clinical emergencies and were processed within two hours of receipt.

NLM's System for Automated Interlibrary Loan (SAIL) filled 7,389 requests from its store of scanned bitmapped pages of selected journal titles. Although SAIL is a technical success, an analysis of data collected during a test period indicated that NLM does not receive a large enough number of interlibrary loan requests for specific articles, or even specific journal titles, for SAIL to be cost-effective method for handling a substantial percentage of NLM's request traffic. The Library will continue to use SAIL to fill requests for some high-demand titles, with publisher permission. SAIL will also assist in delivering some Federal government publications, including AHCPR-sponsored practice guidelines.

The number of libraries using DOCLINE increased to 2,541 in FY 1993, with 12 percent accessing the system via the Internet. Participating libraries entered almost 2.6 million document requests into DOCLINE and filled 86 percent of them. NLM filled an additional 8 percent for an overall fill rate of 94 percent. DOCLINE enhancements implemented in the fall of 1992 (including an increase in the number of libraries allowed in routing tables and the ability to stop routing at a particular cell in a routing table) had the effect of reducing the percentage of DOCLINE requests referred to NLM. Although the number of DOCLINE requests entered by network participants increased 11.6 percent, NLM's share increased only 1.7 percent.

Individual Grateful Med users may employ the Loansome Doc feature to route requests for documents identified in MEDLINE searches to a network library that has agreed to serve them. These requests are routed by DOCLINE if that library is unable to fill them. In FY 1993, Grateful Med users initiated Loansome Doc requests for 103,627 documents. Loansome Doc can now be used by health professionals in the Canadian provinces of Alberta, British Columbia, Manitoba, and Ontario. In FY 1994, access will be extended to all health professionals in Canada.

DOCLINE routes requests to appropriate libraries based on approximately 1.28 million SERHOLD records, which represent the holdings of 3,168 NN/LM member libraries. NLM completed development of an online update mechanism for SERHOLD. SERHOLD participants may now update their holding online, if they choose. The combination of batch updates in SERHOLD or MARC holdings format and the new online update capability should lead to more current and accurate holdings data. In FY 1993, NLM established a mechanism for transferring summary DOCLINE statistics to the Master Serials System to facilitate the analysis of use of serial titles at NLM and throughout the NN/LM.

Onsite users requested 214,991 documents from NLM's closed stacks in FY 1993. Sixty-five percent of the requests were for regular day-time stack service; 25 percent were for the fee-based overnight photocopy service; and 10 percent for special Information Broker Stack Service. Eighty-eight percent of the requests were filled; 94 percent of the requests filled during the day-time were delivered to users within 30 minutes.

National Network of Libraries of Medicine

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 4,020 Network members including health sciences libraries of every size and type located in all parts of the country. NLM's Network Office oversees and coordinates Network programs which are administered by the eight Regional Medical Libraries (RMLs). The Network 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). The RMLs now perform most of the exhibiting and demonstration of NLM's products and services at health professional meetings around the country. NLM staffs exhibits at the meetings in the Washington, D.C. area. In FY 1993, NLM and network services were highlighted at about 120 exhibits throughout the country.

Another related focus of the NLM program is technology transfer. In FY 1993, NLM and the RMLs surveyed all network libraries to determine their satellite downlink capabilities and their level of access to the Internet. The data collected will be used in planning the use of satellite broadcasts in training and outreach and in developing strategies for

increasing the number of health sciences libraries with Internet access.

Special Onsite Programs

In addition to reference and document delivery services, 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 Briefings

NLM is a popular attraction for domestic and international visitors with an interest in biomedical communication, medical librarianship, and information technology. In FY 19932, LO staff members conducted 131 regular daily tours for a total of 396 visitors. The Office of Public Information (Office of the Director) arranged 135_special tours and orientation programs for groups (1764 visitors). NLM staff members also arranged special briefings on library programs and services for many individual visitors.

Historical Programs

In FY 1993, NLM collaborated with the Friends of the NLM, Hahnemann Medical College, and Washington University at St. Louis on a joint observance of the 500th anniversary of the birth of Paracelsus, involving exhibits at all three institutions, a joint exhibit brochure with an extensive essay by Professor Allen DeBus, and a lecture by Dr. DeBus, to be held early in FY 1994.

Stephen Strickland, Ph.D., continued his work on the history of the Regional Medical Programs (RMPs) as a visiting historical scholar. The Library also continued to assist Dr. Morris Allen in his project on the history of medical informatics.

The Library prepared several historical exhibits including "The Proud Profession: Nurses in the Federal Services," "History of Caesarean Section" (with the American College of Obstetricians and Gynecologists), and "Paracelsus and his Impact on Medical Science." NLM also installed visiting exhibits on "Fever, Agues, and Cures: Medical Life in Old Virginia" and "Ellis Island: America's Immigration Cornerstone."

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 collection assessments and analysis of authorship in journals produced in developing countries previously mentioned, the FY 1993 Associates carried out preliminary work on a database of NLM outreach projects, further analysis of structured abstracts in MEDLINE, design of enhanced displays in the Coach UMLS Metathesaurus browser, and a study of non-NLM information sources that are particularly useful in responding to health services research inquiries. 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.

Five Associates completed the 1992/1993 program, including the first international Associate from China. The four U.S. Associates have accepted jobs at the Region 4 RML (University of Nebraska), National Center for Education in Maternal and Child Health (Georgetown University), Union Memorial Hospital in Baltimore, and Lexical Technology, Inc. in Alameda, California. The international Associate returned to Beijing. Five new Associates began the program in September, including an international Associate from Lithuania.

Table 1
Growth of Collections

Collection	*Previous Total (9/30/92)*	*FY 1993*	*New Total*
Book Materials			
Monographs:			
Before 1500	575	1	576
1501-1600	5,760	13	5,773
1601-1700	10,070	25	10,095
1701-1800	24,383	31	24,414
1801-1870	39,951	32	39,983
Americana	2,341	0	2,341
1870-Present	565,976	10,414	576,390
Theses (historical)	281,794	0	281,794
Pamphlets	172,021	0	172,021
Bound serial volumes	977,179	21,285	998,464
Volumes withdrawn	(35,149)	0	(35,149)
Total volumes	2,044,901	31,801	2,076,702
Nonbook Materials			
Microforms:			
Reels of microfilm	(56,753)	(1,731)	(58,484)
Number of microfiche	(265,880)	(1,118)	(266,998)
Total microforms	322,633	2,849	325,482
Audiovisuals	54,000	1,736	55,736
Computer software	578	71	649
Pictures	56,600	0	56,600
Manuscripts	2,445,792	8,750	2,454,542
Total nonbook	2,879,603	13,406	2,893,009
Total book and nonbook	4,924,504	45,207	4,969,711

Table 2
Acquisition Statistics

Acquisitions	*FY 1991*	*FY 1992*	*FY 1993*
Serial titles received	21,181	21,863	22,397
Publications processed:			
Serial pieces	158,939	157,882	154,069
Other	23,344	25,753	23,682
Total	182,283	183,635	177,751
Obligations for:			
Publications	$3,943,338	$4,358,439	$4,129,478
Included for rare books	($184,742)	($193,193)	($149,829)

Table 3
Cataloging Statistics

Item	*FY 1991*	*FY 1992*	*FY 1993*
Completed Cataloging			
Full	12,707	15,278	15,050
Limited	6,480	5,943	6,785
Total	19,187	21,221	21,835

Table 4
Bibliographic Services

Services	*FY 1991*	*FY 1992*	*FY 1993*
Citations published in MEDLINE	363,344	401,562	376,312
For *Index Medicus*	341,874	380,485	358,993
Recurring bibliographies	23	23	22
Journals indexed for *Index Medicus*	3,020	3,048	3,058
Abstracts entered	281,644	295,803	280,599

Table 5
Circulation Statistics

Activity	*FY 1991*	*FY 1992*	*FY 1993*
Requests Received:	494,515	528,288	522,472
Interlibrary Loan	281,606	302,271	307,481
Onsite	212,909	226,017	214,991
Requests Filled:	385,405	401,565	401,162
Interlibrary Loan	207,670	219,831	220,464
Photocopy	193,855	207,685	207,442
Original	12,606	10,726	11,493
Audiovisual	1,209	1,420	1,529
Onsite	177,735	181,734	180,698

Table 6
Online Searches

DATABASES	*FY 1991*	*FY 1992*	*FY 1993*
AIDSDRUGS	310	389	582
AIDSLINE	36,904	38,165	38,485
AIDSTRIALS	646	925	1,377
ALERT	.....	1,923	
AVLINE	15,760	15,435	22,298
BIOETHICS	11,221	13,536	15,450
BIOTECHSEEK	.....	738	781
CANCERLIT	79,511	79,562	83,805
CATLINE	213,376	211,834	279,474
CCRIS 4,862	4,038	4,763	
CHEMID	7,939	9,961	10,782
CHEMLINE	26,878	20,928	18,784
CLINPROT	1,117	.....	.....
DART	4,632	4,168	3,338
DBIR	2,241	1,302	115
DENTALPROJ	279	205	120
DIRLINE	9,482	10,072	11,036
DOCUSER	9,475	10,982	13,082
EMIC	.....	.....	1,082
EMICBACK	2,107	3,794	4,473
ETICBACK	1,776	1,264	1,076
GENETOX	919	1,778	1,496
HEALTH	175,285	172,124	192,083
HISTLINE	5,918	5,697	4,658
HSDB	42,479	36,934	33,239
HSTAR	.....	3,301	
INFORM	306	1,202	.....
IRIS	12,133	20,710	23,244
LOAN STATUS	737	3,821	7,475
MEDLINE	2,731,557	2,788,562	3,063,374
MED86	700,010	580,815	136,374
MED85	.....	.....	539,333
MED83	402,914	327,954	63,882
MED80	247,919	209,780	325,919
MED77	151,620	134,417	33,512
MED75	.....	.....	122,427
MED72	106,882	97,449	20,160
MED66	84,586	76,500	116,844
MESH VOCABULARY FILE	27,079	35,936	38,355
NAME AUTHORITY FILE	3,507	3,644	2,585
PDQ	44,194	25,748	24,342
POPLINE	17,260	18,696	17,328
REFLINE	51,393	47,347	43,301
RTECS 17,684	16,291	15,122	
SDILINE	55,077	52,666	51,733
SERLINE	75,309	81,232	178,945
STORED SEARCH	154	127	248
TOXLINE	80,191	68,634	63,146
TOXLINE65	10,411	7,138	6,125
TOXLIT	17,653	14,047	10,378
TOXLIT65	5,634	3,867	2,546
TRI	35,674	30,207	25,519
TRIFACTS	.....	392	663
USERS	.....	4,581	
YEAR86	18	4	3
Total	**5,533,019**	**5,291,017**	**5,685,067**

Table 7
Offline Searches

DATABASES	*FY 1991*	*FY 1992*	*FY 1993*
AIDSLINE	1,807	1,998	1,971
AVLINE	129	157	82
BIOETHICS	30	23	3
CANCERLIT	3,586	3,408	3,967
CATLINE	555	565	526
CHEMLINE	1	2	0
DENTALPROJ	.....	.....	1
HEALTH	10,654	10,983	10,509
HISTLINE	2	7	4
MEDLINE	5,364	4,419	2,175
MED86	4,751	3,726	596
MED85	.....	.....	1,361
MED83	3,415	2,318	332
MED80	2,513	1,667	1,017
MED77	1,401	876	158
MED75	.....	.....	341
MED72	993	616	94
MED66	610	368	302
MESH VOCABULARY FILE	1	0	1
POPLINE	3,778	4,627	4,684
SDILINE	226,397	227,103	243,713
SERLINE	8	4	1
TOXLINE	5,421	5,610	4,332
TOXLINE65	24	2	2
TOXLIT	4,441	4,191	2,887
TOXLIT65	22	3	0
Total	275,903	272,673	279,059

Table 8
Reference Services

Activity	*FY 1991*	*FY 1992*	*FY 1993*
Reference Section:			
Offsite requests	20,376	23,037	23,489
Onsite requests	46,140	49,511	47,581
Total	66,516	72,548	71,070

Table 9
History of Medicine Activities

Activity	*FY 1991*	*FY 1992*	*FY 1993*
Acquisitions:			
Books	66	152	115
Modern manuscripts	73,388	538,125	0
Prints and photographs	0	0	0
Processing:			
Books cataloged	330	355	297
Modern manuscripts processed	129,000	44,175	31,940
Pictures cataloged	0	0	0
Citations indexed	5,888	5,777	4,801
Public Services:			
Reference questions answered	12,184	9,528	12,352
Onsite requests filled	5,992	6,447	5,751
Pictures supplied	4,683	7,618	4,754

SPECIALIZED INFORMATION SERVICES

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

With a new administration in Washington, governmental, public, and media interest in the environment and health has reached new heights. Vice-President Gore has written *Earth in the Balance*, a best-selling and influential book about the need to take environmental action to restore and preserve our planet. On January 21, 1993, Senator Daniel Patrick Moynihan (D-NY) introduced in the Senate the "Environmental Risk Reduction Act of 1993" (S. 110) to require the EPA to seek advice concerning environmental risks, and for other purposes. A bill has been introduced to Congress to elevate the EPA to cabinet level status. Instead of simply looking at ways of cleaning up the environment, the EPA and other agencies and groups are placing an emphasis on pollution prevention, toxic use reduction, and design for the environment. The continuing evolution of risk assessment is helping to develop methodologies for realistically characterizing environmental health hazards. The NIH, in its 1993 strategic plan entitled *Investment for Humanity*, cites several research initiatives in environmental health which should become high priority areas.

The surge of these activities has been reflected in the past year's work in the Toxicology Information Program (TIP) of the Division of Specialized Information Services. TIP is drafting an implementation plan in response to the recent report of the Board of Regents' Planning Panel on Toxicology and Environmental Health. The three goals of the report were for TIP to consider the following areas with regard to toxicology and environmental health: 1) provide selected core information resources and services, 2) facilitate access to national and international information resources, and 3) provide an information infrastructure for future scientific discovery.

The recent addition of TRI91 (Toxic Chemical Release Inventory - 1991) to the TOXNET system represents the first year that EPA has required industrial reporting of source reduction and recycling activities in compliance with the Pollution Prevention Act of 1990. TRI91 also includes data on quantities of accidental releases. With five years (1987-1991) of TRI data on chemical releases to the environment and transfers to waste sites, TOXNET's inherent computational capabilities now permit the analysis of pollution trends over time. Future years of TRI data will include an expanded list of toxic chemicals and reporting by federal and military installations. The other new file to join TOXNET in 1993 was the EMIC (Environmental Mutagen Information Center) database. This bibliographic file provides up-to-date information on chemical, biological, and physical agents tested for genotoxic activity.

To keep up with the interest in risk assessment activities in 1993, TIP has held several highly successful IRIS (Integrated Risk Information System) workshops on site. These have attracted users whose primary interest is in EPA's IRIS file as implemented on TOXNET, but who also want to learn about other TOXNET files, such as HSDB, for additional information to support risk assessment.

Another 1993 TIP highlight has been making TIP files via ELHILL and TOXNET available on the Internet. The addition of Internet to already existing SPRINTNET, TYMNET, and COMPUSERVE access, throws open TOXNET's doors to a much larger potential audience.

TIP information activities traditionally have reflected fertile areas within the toxicological research community. More and more, TIP is also aligning itself with broader public policy issues pertaining to hazardous chemicals and environmental health, while continuing to take advantage of the latest technological advances to make this information easily and widely available.

TOXNET and Its FILES

TOXNET's networked microprocessor system continued to evolve over the last year to accommodate new system technologies and enhancements. All access to the TOXNET system was upgraded by means of high-speed modems and printers across all terminal servers. Internet access to TOXNET was implemented during FY 1993, and new user-access menus are being developed to simplify Internet traffic. Most important has been the design and implementation of a Windows Workstation for building and updating HSDB records. Testing of the new Workstation's features has been conducted by in-house staff.

The Division of Specialized Information Services was successful in obtaining a new contract for the continuation of support for the TOXNET system. After a long negotiating process, a new competitive contract was awarded to Sentient Systems for one base year and options to renew for the next four years. The new contract award assures continuity in providing TOXNET system support for all of the databases and files mounted on TOXNET.

During FY 1993, two new files were added to the TOXNET system, bringing the total to 15. They were: TRI91 (the Toxic Chemical Release Inventory's 1991 reporting year data) which includes, for the first time, data on Pollution Prevention; and the EMIC (Environmental Mutagen Information Center) database, which is the continuation file for EMICBACK. Both files are searchable on the TOXNET system.

Other system enhancements during FY 1993 include a series of special tabular displays for the TRI91 Pollution Prevention data; and techniques to accelerate the updating process for all files on TOXNET. Updates to all file-building contractors' equipment as well as Scientific Review Panel members' equipment are planned for next year. Online Reference Guides for all TOXNET files were revised and deployed.

The **Hazardous Substances Data Bank** (HSDB) continues to be the most highly used data bank on TOXNET, averaging over 600 hours of online access each month. The decision by the Agency for Toxic Substances and Disease Registry (ATSDR) to withdraw support for HSDB had an adverse impact on the number of records enhanced, updated, and peer-reviewed. During this period, 17 new records were added to the file, bringing the total to 4,372 records, and 312 records were enhanced or updated. A total of 1,284 records went through Public System Update, which includes source updates for Threshold Limit Values.

The **Toxic Chemical Release Inventory** (TRI) series of files, including TRI87, TRI88, TRI89, TRI90, and TRI91 remains 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. Starting with the TRI91 file, which was released to the TOXNET user community in May 1993, the reporting facilities were required to report source reduction and recycling activities and quantities in addition to the environmental releases. These additional reporting requirements were mandated by the Pollution Prevention Act of 1990 and have considerably increased the size and complexity of the TRI91 data. The EPA has announced other future increases to the TRI files, with the addition of another 200 reportable chemicals by 1995, and the inclusion of submissions from Federal facilities and military installations.

The **TRIFACTS** file continues to supplement the environmental release data on chemicals listed in TRI, with information related to the health and ecological effects, and safety and handling of these chemicals. These data are 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. As new chemicals are included in TRI, new TRIFACTS records will need to be added.

The **Chemical Carcinogenesis Research Information System** (CCRIS) continues to be built and maintained directly on TOXNET by the National Cancer Institute. The data bank has grown during this report period by over 1000 chemical records, bringing the total to 4,624. The chemical-specific data falls into the areas of carcinogenicity, tumor promotion, mutagenicity and tumor inhibition.

The **Integrated Risk Information System** (IRIS), EPA's health risk assessment file, continued to make significant growth in usage during the past year and it is now the second most heavily used file on TOXNET. The size of the file has increased to 647 chemical records. Concept Menu and Grateful Med form screens will increase its usefulness even more for novice and occasional users. Over 60 users have been trained in the use of IRIS and related TOXNET files in a series of workshops conducted by in-house staff during FY 1993.

The **Developmental and Reproductive Toxicology** (DART) now contains over 17,000 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 Registry Numbers 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. Records from DART are also included in the TOXLINE database and are added on a quarterly basis. DART is a continuation of the Environmental Teratology Information Center Backfile (ETICBACK) database on TOXNET. ETICBACK contains almost 50,000 citations to literature published from 1950-1989. ETICBACK citations are also found in TOXLINE.

The **Environmental Mutagen Information Center** (EMIC) database contains citations to literature on agents that have been tested for genotoxic activity. It is produced by the Department of Energy's Oak Ridge National Laboratory and managed by NLM. In June 1993, the new EMIC database, containing over 5,000 citations to literature published since 1991 was made publicly available on the TOXNET system. Like DART, many of the citations for EMIC are derived from MEDLINE. All of the records contain bibliographic citations, abstracts (when available), EMIC special keywords, chemical names, and Registry Numbers. ORNL is also responsible for locating and adding citations to relevant publications not found in MEDLINE. A backfile for EMIC (EMICBACK), with 70,000 citations, has been publicly available through TOXNET since June 1989. Records from EMICBACK are included in TOXLINE. Plans are under way to add the records from the new EMIC database to TOXLINE, as well.

These four bibliographic databases on TOXNET (EMIC, EMICBACK, DART, ETICBACK) are funded by the Environmental Protection Agency and the National Institute of Environmental Health Sciences and are managed by NLM.

The **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. New

categories of international data were added in FY 1993. In addition, SIS continues to assist in the location of CAS Registry Numbers for records in RTECS so that these important identifiers can be present on records to help users locate information across the toxicology files.

The GENE-TOX file continues to be built and updated directly on TOXNET by EPA scientific staff. GENE-TOX contains peer-reviewed genetic toxicology (mutagenic) test data on about 3000 chemicals. This is one of six files on TOXNET that can be searched simultaneously via the CrossFile command.

Databases under ELHILL

ChemID (Chemical Identification File) is an online chemical dictionary that contains over 267,000 records, primarily describing chemicals of biomedical and regulatory importance. ChemID allows users to search by utilizing a variety of chemical and biological identifiers; to formulate strategies for searching other MEDLARS files; and to locate other files that contain more information about the chemical in question.

ChemID contains an important set of regulatory data, collectively known as SUPERLIST. Over 9,000 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, and the ATSDR Priority List. These data allow users to determine if a certain chemical is mentioned on a given list and, it so, under what name; they also allow searching for chemical classes on these lists. In FY 1993, ChemID was augmented with Locators pointing to the Canadian Domestic Substances List (DSL), and SUPERLIST was enhanced with additional state lists of regulated chemicals, including the Pennsylvania Right to Know list.

CHEMLINE (Chemical Dictionary Online) is an online chemical dictionary and directory file that 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 that 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 from its Registry System. 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 1993, the scope of coverage of the CHEMLINE file continued to increase. CHEMLINE now covers substances in the TRI91 and EMIC files on TOXNET. For seekers of regulatory information, the data from the TSCA Inventory were updated, as well as data for the European Inventory of Existing Commercial Chemical Substances (EINECS). Locators pointing to the Canadian Domestic Substances list (DSL) were added.

TOXLINE (Toxicology Information Online) is an online bibliographic retrieval service, produced by merging "toxicology" subsets from 17 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 over 3,600,000 records. Approximately 20,000 records are added with each monthly update.

During FY 1993, the TOXLINE files were rebuilt to add current MeSH indexing vocabulary to the subset derived from MEDLINE, the Biological Abstracts subfile records, and the DART (Developmental and Reproductive Toxicology) subfile. Also this year, work was begun to add the most current mutagenicity records to TOXLINE. This subfile, EMIC (Environmental Mutagen Information Center), already has older records in TOXLINE, and is now also available as a separate database on the TOXNET system. Additionally, work was done this year to add more FEDRIP (Federal Research in Progress) subfile records to TOXLINE, covering additional Federal agencies supporting research related to toxicology.

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.

The first year of NLM's responsibility for the general biomedical subfile of DIRLINE was completed successfully. The original Library of Congress component was reviewed and records covering subjects outside of NLM's scope were culled from the database. A contract was initiated to maintain the subfile, including reviewing and revising existing records and adding new ones.

The NIH Research Resources subfile of DIRLINE was completely updated with descriptions of several hundred new resources added to the database. The oversight for this subfile is provided by the NIH Office of Extramural Programs.

The Directory of Biotechnology Information Resources (DBIR) was removed from the TOXNET system and is available only as a subfile of DIRLINE. A complex conversion routine was created to improve the format of this data in DIRLINE prior to its removal from TOXNET. The DBIR subfile now contains over 2,000 records describing databases and other information services, organizations, collections, repositories, nomenclature committees and publications, all related to biotechnology and molecular biology.

AIDS

NLM continues to collaborate with the National Institute of Allergy and Infectious Diseases, the Food and Drug Administration, and the Centers for Disease Control and Prevention in support of the PHS AIDS Clinical Trials Information Service. The **AIDSTRIALS** and **AIDSDRUGS** databases are produced as part of this effort and made available by the Division for online searching.

NLM hosted and co-sponsored (with the NIH Office of AIDS Research) the NIH HIV/AIDS Information Services Conference in June 1993. The purpose of the conference was to evaluate NIH's existing information services and recommend changes and additions. Representatives from five major groups: clinical researchers; medical, dental, and nursing care providers; allied health care providers; media and the general public; and patients and the affected community, were brought together to examine some of the major issues facing users of NIH's HIV/AIDS information services. A conference report with recommendations is being written and used by the NIH for planning purposes.

Other Programs

Relational Technology

In the *Relational Technology* project, SIS is investigating the use of a Relational Database Management System (RDBMS) for building, maintaining, and delivering AIDS information. Working with Aspen Systems, the contractor for the PHS AIDS Clinical Trials Information Service, SIS is developing a prototype that will allow the delivery of integrated information, e.g., data on clinical trials and drugs, as well as bibliographic and directory information. This represents a wide variety of information types including large blocks of text, bibliographic citations, and numerical data.

Outreach

SIS continued its support of the Toxicology Information Outreach Project. The objective of this initiative is to strengthen the capacity of Historically Black Colleges and Universities (HBCUs) to train medical and other health professionals in the use of NLM's toxicological, environmental, occupational health and hazardous wastes information resources. In addition to providing workstations, training, and free online access to nine HBCUs participating in a pilot training development project, NLM has collaborated with the Agency for Toxic Substances and Disease Registry to train representatives from 27 additional schools in the use of NLM's valuable online resources. During the past Fiscal Year, one of the training classes was hosted by Howard University and was coordinated by Howard's participant in a "Train-The Trainers" class designed and conducted by the Oak Ridge Institute for Science and Education (ORISE) to prepare trainees for conducting classes at their respective institution. The three instructors for this introductory-level training class were from Howard, ORISE and NLM. A preliminary evaluation report was given at the spring meeting of the Toxicology Information Outreach Panel which indicated successful implementation of this project at each institution participating in the pilot.

User Support Activities (Microcomputer-Based Training Section)

The Division completed development of multimedia, microcomputer-based tutorials for the TOXLINE and MEDLINE databases. These tutorials are written in TOOLBOOK for delivery in a Windows environment. They take full advantage of the Windows' graphical user interface incorporating color graphics, photography, animation, and audio. They also include an "ELHILL Simulator" with mini-databases for practice searching. These tutorials provide an alternative or augmentation to formal classroom training. Development of similar tutorials for the CHEMLINE/ChemID databases as well as the ELHILL software is under way. The Library plans to release the tutorials as a set in the Spring of 1994. Additionally, as part of a collaborative project with the International Center for Cancer Information, the Division has directed the development of a microcomputer-based tutorial for the CANCERLIT database. It is written in QUEST for delivery in a DOS environment and is expected to be completed in Fall 1993. The Library also plans to repackage its DOS versions of TOXLEARN, CHEMLEARN, MEDTUTOR, and ELHILL LEARN under QUEST using the CANCERLIT LEARN as a model.

Alternatives to Animal Testing

SIS staff has continued the compiling of the quarterly annotated bibliographies, "Alternatives to the Use of Live Vertebrates in Biomedical Research and Toxicology Testing." Copies are distributed free to requesters and are announced as being available in the *ILAR NEWS*, a publication of the National Research Council. The NRC Institute of Laboratory Animal Resources (ILAR) also publishes an annual version of these quarterly bibliographies.

Long Range Planning Panel on Toxicology and the Environment

The report of the NLM Long Range Planning Panel on Toxicology and Environmental Health, developed at its several meetings (October 1991, January and March 1992), provides advice on directions to be taken in the domain of toxicology and environmental health. It was presented to the NLM Board of Regents in October 1992 and reviewed with the National Research Council's Toxicology Information Program Committee (TIPCOM) on November 17-18, 1992. SIS staff is developing a plan to support the implementation of components of the Panel's report. Appropriate scenarios were discussed with TIPCOM at the meeting which convened at NLM on July 15-16, 1993.

Information Services to Other Agencies

SIS provided support to the ATSDR, NIOSH, NCI, NIEHS and EPA in the building and deployment of files of special or mandated interest to those organizations. SIS also provided training to staff of those organizations in the use of all information resources accessible through NLM. The Division also provided support for the Subcommittee on Information Coordination (SIC) of the DHHS Committee to Coordinate Environmental Health and Related Programs (CCEHRP) in the Office of the Assistant Secretary of Health. In this regard, SIS continued its efforts in developing directories of DHHS-sponsored projects in risk assessment and epidemiology.

Because of the Division's commitment to providing literature support for animal alternatives in research and testing, NLM is now represented on the Interagency Coordinating Committee for the Use of Animals in Research (ICCUAR), which is mandated under the NIH Revitalization Act of 1993. ICCUAR is responsible for an annual inventory of activities relevant to the use of animals in biomedical research and toxicological testing.

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 Center's intramural research programs. The Board is composed of scientific and technical experts (see Appendix 5 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.

In FY 1993 the Lister Hill Center selected its second group of students for the Undergraduate Research Study Program (URSP). This program provides two-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 Lister Hill Center's staff and facilities to provide experience in biomedical image processing and computer visualization. In 1993 two students from Morgan State University in Maryland and one student from Southern University in Louisiana completed the two year program.

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, intelligent database 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 (UMLS) 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 eight-week NIH "Medical Informatics" elective for third-year and fourth-year medical students each spring.

Expert Systems Program

Expert systems are computer programs combining 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 1993 have been the Coach Expert Search Refinement System and Coach Metathesaurus Browser, the AI/RHEUM consultant system in rheumatology, the CTX "criteria engine" shell and its family of automated testing and validation tools, the medical expert systems evaluation project, and the Rheumatology Image Library videodisc and its associated browser program.

The Coach Expert Search Refinement System brings to bear the UMLS Metathesaurus and other knowledge sources to help users of NLM's Grateful Med program improve

retrieval when searching MEDLINE. The system works interactively with the user, with Grateful Med, and with the ELHILL system on NLM's mainframe. 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. The Coach Metathesaurus Browser provides users with multiple views of the UMLS Metathesaurus, an extremely rich and increasingly comprehensive resource. Concept definitions, semantic types, synonyms, lexical variants, related terms, child and sibling terms, co-occurring terms and other information can be displayed, helping users to visualize search terms in MeSH tree context and to choose appropriate additional terms to incorporate into their queries. The Coach Metathesaurus Browser was distributed on CD-ROM as part of the 1993 release of the UMLS Knowledge Sources.

The Coach Expert Search Refinement System analyzes user searches, applying MeSH and its own special knowledge sources in performing such functions as detecting occupational specialty headings and synonyms of topical subheadings and mapping 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. The Coach system was alpha tested late in 1992 by 36 colleagues inside NLM. It is now in beta test at multiple sites across the country, with distribution anticipated early in 1994.

The knowledge base of the AI/RHEUM consultant system was substantially expanded in FY 1992 and its data entry process streamlined. The system was field tested for the first time, with collaborating clinicians in family practice settings in five sites in Missouri and in internal medicine practices in five sites in Utah. Early in FY 1993, focus group sessions were held with the clinicians who had participated in this phase of the study. The focus group discussions provided important feedback from these busy physicians on the way a user interaction must flow if the system is to be helpful in real-world medicine. Suggestions on additional access points to auxiliary knowledge sources such as the Rheumatology Image Library video segments and on augmented disease coverage in the knowledge base were received. In consultation with expert rheumatologists in Utah, Expert Systems Program staff made system changes directly responsive to the clinicians' comments. The new version of AI/RHEUM was retested in the Utah clinics with very positive reaction from the clinical users. This final phase of the study concluded the data acquisition portion of the expert systems evaluation project.

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, has been tested with projects in several new subject domains. Discussions with the Agency for Health Care Policy and Research (AHCPR) have explored the use of CTX to implement developmental versions of the AHCPR's Clinical Practice Guidelines. The CTX shell, with its explicit and unusual multimedia links to knowledge sources in different forms in different places, even on different machines, is designed to serve as a useful resource to developers in multiple subject areas. 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.

Dr. Kingsland of the Expert Systems Program served again in FY 1993 as coordinator for the eight-week NIH "Medical Informatics" elective for third-year and fourth-year medical students. Eleven students from medical schools across the U.S. completed the elective. The course included a seminar series of more than 30 90-minute lectures, independent research projects under the direction of NIH preceptors, and oral and written presentations of research results. Several of these extremely bright, highly motivated students themselves made important contributions to Expert Systems Program projects.

Natural Language Systems Program

The Natural Language Systems (NLS) Program investigates the contributions 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. SPECIALIST includes a parsing system supported by a large lexicon, modules that access the extensive UMLS Knowledge Sources, and a retrieval module for experimentation in information access and retrieval. The SPECIALIST system has been under development for several years. In FY 1993 several of its major components were tested and prepared for release to interested researchers. Feedback from these researchers will result in continued improvements to the system.

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

words in the language.

The SPECIALIST lexicon currently contains over 60,000 lexical records with more than 118,000 lexical forms. Lexical entries are created with a lexicon building tool that incorporates a grammar of all the allowable slots and their values. The lexical grammar serves to constrain the possible choices that must be considered when entering an item and to provide an automatic check of the correct form of completed lexical records. Beginning in 1994, the lexicon will be released on a regular basis with the UMLS Knowledge Sources. Consistent with the other knowledge sources, the lexicon will be provided both in unit record and in relational table formats. ASN.1 format is being considered for future releases. During FY 1993 the lexicon was sent to beta-test sites at several institutions including the University of Utah, Columbia University, Oregon Health Sciences University, and the University of Twente in The Netherlands.

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. During FY 1993 the SPECIALIST morphology module was prepared for release with the UMLS Knowledge Sources. The programs are currently being used to detect lexical variants in the set of terms scheduled for inclusion in Meta-1.4, the 1994 release of the UMLS Metathesaurus. Morphological analysis is most successful when interleaved with look-up in a lexicon. This ensures that part-of-speech and other crucial information is available for the morphologic processing. The morphology module in conjunction with the SPECIALIST lexicon should comprise a powerful tool for those conducting research in biomedical language processing and information retrieval.

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 has reliably determined the syntactic structure which supports the semantic interpretation of a particular sentence. However, it typically also produces numerous additional parses which are syntactically correct but do not contribute to the final interpretation. During FY 1993, extensive work focused on reducing these unwanted parses. This involved a modification to the parsing system that exploits the syntactic information available in the lexicon and generates a single underspecified syntactic representation for each linguistic structure encountered. This representation then depends on flexible semantic interpretation and robust domain knowledge processing to produce a final interpretation, or conceptual structure. At the heart of this approach is a mapping to concepts in the UMLS Metathesaurus and the UMLS Semantic Network.

During FY 1993, NLS staff developed Meta, a Unix-based retrieval system for browsing, navigating, and extracting information from the UMLS Metathesaurus and the Semantic Network. The system can be used both for batch processing and interactively to answer individual user questions. It can be used to query any of the information in the Metathesaurus related to a particular concept. In addition, global searches can be executed to report all concepts having a particular feature or set of features. An initial client-server version of Meta was implemented in FY 1993 and released for beta-testing. The system currently uses a simple client-server protocol which will eventually be enhanced to be compliant with the Z39.50 protocol. The first version of the client software is already being used in the SPECIALIST system. It is expected to be ready for general distribution with the 1994 release of the UMLS Knowledge Sources.

The problem of providing users with the information they seek can be viewed as the problem of mapping the language of the user to the language of the database. Users formulate queries in ways that reflect their knowledge and understanding of the topic, expecting this to be sufficient to retrieve relevant information from the database. However, because of the richness and diversity of natural language, mapping between requests and information is rarely straightforward. An important goal of the NLS research program is to establish a more precise understanding of the relationship between user queries and the information that may be relevant to those queries. Providing structured representations of both queries and documents is an essential first step in this mapping process. Mapping then becomes a matter of matching conceptual structures. The type of semantic processing currently being explored in the context of the SPECIALIST system has shown some early promise for enhancing existing information retrieval methods.

NLS program staff participates actively in the medical informatics community. Program staff have published several research papers during FY 1993. Dr. Alexa McCray, NLS Program leader, is co-editor of the International Medical Informatics Association (IMIA) Yearbook of Medical Informatics.

Machine Learning Project

Since 1989, the Lister Hill Center has been exploring the application of machine learning technology to biomedical problems. Machine Learning encompasses a wide 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 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 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 it, analyze it and send a report 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 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 which 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 searchers 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 sequence database of the National Center for Biotechnology Information), 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 blocks 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 model 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 personnel 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 in the design of artificial intelligence 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.

Applications of machine learning technology developed at the Lister Hill Center have recently led to significant biomedical results. The bone disease Osteogenesis Imperfecta

(OI) is caused by mutation in collagen, although the exact mechanism remains a mystery. Recent developments in gene sequencing technology have led to the identification of about 70 specific mutations that result in OI, ranging from nearly asymptomatic to neonatally lethal. The relationship between mutation and disease severity has been under intense scrutiny for several years. Published expert hypotheses about the relationship were contradicted by additional data, and traditional statistical and computational methods were unable to draw conclusions about the complex molecule given such a small amount of data. However, a method developed in the machine learning project called the focus-induce-extract discovery strategy was used successfully to find patterns in mutation lethality in the data. These patterns are now being used to direct more detailed molecular modeling studies, and may shed light on open questions about the structure of collagen in addition to the disease mechanisms.

MedIndEx Project

The MedIndEx (Medical Indexing Expert) 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 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 Lucid Common Lisp/Sun 4.1 and runs on Sun SPARCstation 2 workstations under the Solaris 1 operating system. Domain-independent project software includes a Lisp-based experimental frame language. The interface uses public-domain window system software X Windows (X11 Release 5) and CLUE (Common Lisp User Interface Environment). MedIndEx is designed to run similar indexing and KB Manager applications in other domains. As of late FY 1993, the KB contained 5800 frames (MeSH concepts).

In FY 1993, the indexer interface was enhanced to provide access to MeSH scope notes and Chemtool indexing terms. When a term is highlighted using a mouse pointing device, its scope note is displayed in a pop-up window. Chemtool enhances the coverage of chemicals by providing access to the MeSH Supplementary Chemical Records, which contain chemical indexing terms not available in the regular MeSH. Records in this file prescribe mappings from these Chemtool terms to regular MeSH terms and suggest associated regular MeSH pharmacologic action terms for consider-

ation as additional indexing terms. The MedIndEx implementation of Chemtool includes a searchable, browsable, mousable menu of preferred Chemtool terms with their aliases. When a Chemtool term is selected or keyboarded, the system automatically reminds the indexer to consider adding the associated pharmacologic action term.

Also in FY 1993, an evaluation project was completed resulting in an experimental design to evaluate MedIndEx. The design establishes a new methodology for evaluating indexing systems, including specifications for an indexing experiment comparing MedIndEx with NLM's Automated Indexing Management System (AIMS), procedures for establishing an *a priori* indexing standard for test documents, scoring procedures for indexing quality and inter-indexer consistency, procedures for selecting test documents, general recommendations for collecting time-stamped data, and a questionnaire to elicit opinions about the usability and performance of MedIndEx. The design will be distributed by the National Technical Information Service in a report titled "Design for a Study to Evaluate the MedIndEx Approach to Subject Indexing," which includes an executive summary of the evaluation project by NLM. In addition, NLM has developed software for computing indexing quality and consistency scores as specified in the design.

Work planned for 1994 includes developing an evaluation project to test MedIndEx using the experimental design produced in FY 1993; enhancing the indexer interface to more fully encompass the indexing task, improve efficiency, and improve displays and help messages; enhancing interfaces for assignment of documents to indexers and generating indexing score reports; and updating the MedIndEx system design document and user manuals.

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 helps 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 pgs.), and the database of 92 NIH Consensus Development Conference Reports.

The initial version of IRx2 still employed the IRx1 Search Engine. During 1993, IRx2 Search Engine was completely rewritten in C++ allowing for much greater modularity and performance improvements. Present IRx2 developments are targeted towards the addition of non-text objects, enhanced retrieval functions and user interfaces, user annotations, hypertext, portability to other-than-Unix computer systems, improved communication protocols in support of the client/server paradigm, and online text maintenance.

CD-ROM Program

NLM has a growing need to disseminate large full-text databases, digitized images, and digitized audio in a number of program areas and across multiple platforms such as MSDOS, 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 per disk in large quantities. The effective utilization of CD-ROM, however, is encumbered by an access time much slower than magnetic hard disks.

In 1990, ITB established a laboratory for CD-ROM developments and began to establish technical expertise in CD-ROM design and premastering. The primary laboratory tools include CD-ROM Premastering and Simulation Work-

stations and a CD-ROM Recordable (CD-R) write-once mastering unit. The workstations allow the formatting of tapes for mastering and the simulation of CD-ROM applications even prior to mastering. The CD-R 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.

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), and the NLM ChemID and Toxline databases. 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 or compound terms with the entry of a small number of characters. The search engines for the fielded-data databases have been have been implemented, to date, under Unix and Microsoft Windows.

Communications Engineering Branch

Projects in the Communications Engineering Branch focus on R&D in image engineering: 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 NLM's responsibilities for document delivery, archiving, and preservation. In addition, research into imaging techniques that support medical educational packages employing digitized radiographic, dermatological, and other imagery is also being pursued.

DocView

Following a pilot project proving the technical feasibility of automated document image delivery by mail and fax, a system called DocView is being developed to allow remote users access to document images over Internet. Not only does DocView overcome some of the inherent limitations of fax in terms of image resolution and transmission speed, it represents a state-of-the-art approach to empowering an end user to seek and use electronic documents stored remotely. The components of the prototype DocView system are a Unix-based server holding bitmapped document images, and a PC-based Microsoft Windows client for the end user. The inhouse-developed software for the system amounts to 300 kilobytes of C code. The hardware for the user workstation is a 486 machine, a Kofax image processing board, an Ethernet card and a laser printer, all available off-the-shelf.

The client software running under Windows 3.1 on a 486 platform allows a user to select a server located anywhere on the Internet, the current options being a Sun 690MP in the lab, a 3-processor C3830 Convex supercomputer on the NIH campus, and a Sparcstation at the University of Arizona. After connecting to any one of these servers, the user receives information on the number of documents available and document citations. The user may key on a citation to retrieve a document, and choose to preview just the first page of the desired document on the screen before deciding to download the entire document. If after previewing, the document is not of interest, the user may delete it. The user may navigate through a document selecting pages of interest by invoking an *electronic bookmark* feature, rotate pages to view those in landscape mode more conveniently, zoom in for better legibility, copy sections of interest into a *notebook* and add text from a wordprocessing program, and either print or store electronically items of interest.

The current version of DocView relies on a store of document images at a Unix server. Work is proceeding to extend this so that the DocView client will also be able to receive images from public domain servers as well as from Ariel. Ariel is a commercial system that, in a manner analogous to conventional fax transmission over telephone lines, transmits document images over Internet, and is increasingly being used by medical and other libraries. Another step will be the alpha and beta testing of DocView to investigate performance and design issues. This evaluation of the client software is planned in terms of ease of use and utility of image manipulation functions. Also an evaluation of both client and server performance will be done in terms of document transfer rates from jukebox to server, error rates, and delivery throughput over the Internet.

System for Automated Interlibrary Loan (SAIL) Program

This R&D program seeks to investigate the technical feasibility and role of automated document delivery to meet the requirements of the NLM's interlibrary loan (ILL) service. This program is motivated by the increasing burden faced by the Library in servicing the interlibrary loan requests in the traditional manual way. The research staff designed and built a prototype system consisting of a networked complex of PC-based workstations. This system automatically retrieved ILL requests from the NLM mainframe computer, parsed these, and used fielded data contained in them to retrieve document images from optical disks and automatically fax or print them for mailing. Operators used document capture systems developed inhouse to scan and store biomedical journals selected according to criteria that predicted high use.

The system is operated in a pilot test mode to investigate performance and cost issues. Cost on a unit basis (per article delivered) turned out to be comparable to the cost

of delivering documents the conventional (manual) way, even considering the disparity in volume (SAIL handled 5% of the total ILL load). The prime component in the cost figure is the labor necessary to convert paper documents to bitmapped electronic images. In terms of performance, delivery is in minutes and hours rather than days or weeks, but there is variation due to ambiguities in the requests, or not having the disk containing the requested article currently mounted in a drive, and other factors. Solutions were found for these performance problems.

By early FY 1993 it was found that the 64 titles preselected for this pilot project delivered 5% of the total ILL requests to the NLM, proportionately a high figure considering the size of the journal collection at the Library, but of the articles stored only one-third were accessed to serve the ILL service. This has motivated a second look at the way articles are entered into the system, prompting an investigation of a point-of-request or delivery-on-demand system. The design of such a system is based on a 486 platform running under Microsoft Windows 3.1. The idea is that the only human operation involved is scanning the requested document. All other operations, e.g., faxing, printing, transmitting over Internet, extracting information from the ILL requests, and updating DOCLINE as to status, are to occur automatically in the multitasking environment of Windows. The design of this integrated system is under way.

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

ANN to Reduce Ambiguities. Work begun in the previous year (and described in the *1992 NLM Programs and Services*) on the application of Artificial Neural Network (ANN) to classification problems encountered in SAIL operation continued. One such problem centered on the ambiguities in ILL requests resulting from remarks made by users in the unstructured comments field in the requests. A system consisting of a parser and an ANN of the back error propagation type was developed. Evaluation of the system showed that it correctly determined that 57% should be filled, correctly determined that 30% should not be filled, left 13% in the uncertain category, and made 0.3% errors. The conclusion is that 87% of the ILL requests which have unstructured comments can be handled automatically, and that the remaining 13% may be referred to a human operator for a decision. By significantly reducing the operator intervention required, this research promises to yield time and cost savings in future operational systems for automated document delivery. This work appears in the literature: Hauser SH, Hsu W, Thoma GR: Request Routing with a Back Error Propagation Network. *Proc. SPIE Conference on Intelligent Information Systems*, 1993, Vol. 1965, pp. 689-95.

ANN for Journal Identification. Image analysis is the subject of another ANN project (also described last year) intended to aid scanning operators by automatically identifying a journal, and thereby reduce operator error in the selection of journal titles while scanning. This is accomplished by processing the image characteristics of a journal's cover page. This research is reported in: Hauser SH, Cookson TJ, Thoma GR: Using Back Error Propagation Networks for Automatic Document Image Classification. *Proc. SPIE Conference on Intelligent Information Systems*, 1993, Vol. 1965, pp. 142-50.

Simulation Studies. To predict a migration path for a scaled-up SAIL system, a discrete event simulation language, GPSS/H, is being used to model the image retrieval subsystem. This model allows a representation of varying numbers of fax servers, optical disk drives, magnetic disk drives and jukeboxes. The results of the simulation will establish theoretical bounds on the number of system components and the overall system architecture for different levels of service.

Automated Portrait/Landscape Mode Detection. As part of research into automated document imaging, an algorithm was developed to detect the orientation (portrait vs. landscape) of a binary page image. Detecting page orientation is a necessary preprocessing stage for optical character recognition, skew detection or skew correction. In addition, page orientation is crucial for automated document entry in which the contents of a printed document is segmented into such regions as headlines, text columns, graphics or footnotes.

The algorithm developed is based on an analysis of projection profiles, vertical and horizontal variances on a page, and a technique to reduce the impact of nontextual data (blanks, graphics, forms, line art, large fonts and dithered images). Using a sample of several thousand images of medical journal pages, the algorithm was found capable of detecting page orientation at an accuracy rate of 99.92%. This work is the subject of a patent filing, and has been reported in the literature: Le DX, Thoma GR: Automated Portrait/Landscape Mode Detection on a Binary Image. *Proc. SPIE Visual Information Processing II*, 1993, Vol. 1961, pp. 202-12.

Automated Document Skew Detection. Rescanning of documents is a time-consuming and costly step, but often necessary in document conversion. Errors are detected at the quality control (QC) stage. A multistage technique was designed to automatically detect page skew. The principal components of this algorithm are component labeling, a procedure to reduce the amount of data to be processed, a technique to minimize the effect of nontextual data (graphics, forms, line art, large fonts and dithered images), and the Hough transform. The algorithm was tested with several hundred images of medical journal pages, and found to detect skew with an accuracy of about 0.5 degrees. This work appears in the literature: Le DX, Thoma GR: Document Skew Angle Detection Algorithm. *Proc. SPIE Visual Information Processing II*, 1993, Vol. 1961, pp. 251-62.

Digital X-ray Prototype Network (DXPNET) Program

This program, whose acronym also stands for *Digital X-ray Prototype workstations linked via InterNET*, aims to investigate the technical feasibility and design issues in

developing, maintaining and operating an archive of digitized radiographs. It 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 impetus for the program is to support 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.

At the inception of this program, which was described in last year's annual report, 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 the University of California, Los Angeles. 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 used for quality checking the images received from the scanning contractor on 5 1/4 inch WORM-type optical disks.

While the PC-based system remains operational at NCHS for the purpose of being able to access data from the part of the NHANES collection which has been written to WORM, in FY 1993, the QC process at NCHS was augmented by the acquisition of a Sun SPARCclassic workstation and eraseable single-platter optical drive. The Branch provided direct hardware and software system integration support to create a new Sun-based QC workstation. The new workstation which runs *Imview*, an inhouse-developed software package, has the advantage of enabling a technician to view the images much faster than with the PC-based system; in addition, images may be read directly from the erasable optical platters on which they are now being written by the scanning contractor.

The Branch organized and hosted a "Digitized Radiographic Images: Challenges and Opportunities" workshop in June 1993 in collaboration with NIAMS and NCHS. The workshop convened a wide range of experts in radiology, epidemiology, image compression, user interface and image presentation issues, and statistics. Design issues related to user interface, image presentation and compression were addressed at the workshop.

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. Our long-term goal is to 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), the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference (September 1992), and the SPIE Medical Imaging Conference (February 1993).

Image Compression Studies. Since image compression is needed to maximize storage capacity and minimize transmission time, key to making such projects as DXPNET practical, compression studies have been undertaken inhouse and with a *collaboratory* of external investigators. The participants in this collaboratory are researchers at Stanford University, Monash University (Australia), the Canada-France-Hawaii Telescope Consortium, NASA Ames Research Center, IBM Almaden Research Center, and NIH. While inhouse studies are focusing on the use of the international JPEG (Joint Photographic Experts Group) standard employing the Discrete Cosine Transform (a "lossy" technique), our external collaborators are experimenting with both lossy and lossless techniques using DXPNET images downloaded over Internet. These methods include pyramidal/wavelet technique, vector quantization, context-based DPCM, and bitplane partitioning.

Experiments with lossless techniques conducted on a sample of the DXPNET images have yielded a compression ratio averaging 2. The compression times on a Sun 690MP are 20 seconds for cervical images and 35 seconds for lumbar images. Lossy JPEG compression applied to a sample of the images showed no perceptible difference for compression ratios up to 40, after which point unacceptable artifacts appeared. Compression and decompression times averaged 25 seconds. This work has been reported in the literature: Berman LE, Long R, Pillemer SR: Effects of Quantization Table Manipulation on JPEG Compression of Cervical Radiographs. *Digest of Technical Papers, Society for Information Display International Symposium*, J. Morreale, ed., Society for Information Display, Playa del Rey, CA. 1993, 43: 937-41.

JPEG Evaluation Tool (JET). The degree of compression attained in lossy JPEG compression technique largely depends on the specific quantization table used. JET was developed by Branch researchers to study the effects of this selection on the radiographs in the DXPNET collection. It is a tool that enables a researcher to modify individual elements of the quantization table or use a predetermined table. It offers a graphical user interface to perform this manipulation, and also computes quality-related factors such as root mean square error, normalized mean square error, compression

ratio and image entropy. (A description of this tool was presented at the *1993 International Symposium of the Society for Information Display*, and appears in the proceedings: Berman LE, Nouri B, Roy G, Neve L: Interactive Selection of JPEG Quantization Tables for Digital X-ray Image Compression.) Experiments under way with researchers at the NASA Ames Research Center appear especially promising in this area: their tools allowing the manipulation of the JPEG quantization table incorporate aspects of the human visual system, properties of the cosine basis functions, and the physics of the monitor.

Electronic Document Delivery System Program

The approach in this automated document delivery system is motivated by the hypothesis that there exists a class of users who need *direct access* to an electronic archive of document images. Three increasingly functional generations of a prototype system have been built and tested.

The final version of the prototype system, built and demonstrated in FY 1992 and reported in the literature in FY 1993, allows a remote user equipped with a Document Request Workstation (DRW) to perform a search of MEDLINE via Grateful Med, and then directly access an electronic store of document images and receive the document images through fax, mail or local pickup. The remote user's DRW consists of affordable, off-the-shelf components such as an IBM-compatible PC clone and a standard fax machine. It also requires special inhouse-developed software. After demonstrating the third-generation EDDS system, in FY 1993, its functions were described in a publication: Thoma GR, Walker FL: *Essential Functions in an Electronic Document Delivery System*. Chapter in: High-Performance Medical Libraries, Advances in Information Management for the Virtual Era (N. Broering, ed.), Meckler Publishing, Westport CT, 1993, pp. 77-88.

Machine-Readable Archives in Biomedicine Program

While bitmapped images are essential for the document delivery and other projects, machine-readable text coupled with graphics can also have a role in such areas as indexing and cataloging. Converting bitmapped document images to machine-readable ASCII characters involves optical character recognition (OCR). However, the reliability of OCR is compromised by "noise" in the image and the presence of unwanted page edge effects. The near-term goal of this program is to conduct research into techniques to enhance bitmapped images to a quality level commensurate with OCR requirements by reducing noise in the bitmapped images and to automatically remove unwanted borders (page edge effects). The long-term goal is to develop a prototype system that will implement the conversion from paper scanned by a bitmapping engine to a database of segmented text/image files.

Recent work has been in image enhancement studies on document images to improve quality commensurate with OCR requirements. The Branch developed an image enhancement technique employing first order statistics (projection profiles) for improving bitmapped image quality by automatically eliminating page edge effects, necessary for low error OCR conversion. This method was used to remove borders from a sample of bitmapped document images available in the lab with partial success.

The image enhancement studies continue with morphological image processing and artificial neural network (ANN) techniques for automated border removal. Morphologic operations such as erosion, dilation, opening and closing will be combined to (a) remove noise, i.e., small bitonal spots on the image; and (b) reduce the image to relatively large blocks. This processed image will be divided into rectangles whose projection profiles will be calculated. The projection profiles and the xy coordinates of the rectangles will be used as the input vector to an ANN of the back error propagation type. Each input will be paired with the correct classification of each rectangle, i.e., (unwanted) border, margin, or (desired) image, and this data will be used to train the ANN. The trained network will be tested with samples of bitmapped images from the lab. It is intended to experiment with different rectangle sizes and positions, and to combine the ANN output with other known information, to reliably classify border regions in the bitmapped image.

Biomedical Digital Image Processing Program

The information content of images, both monochrome and color, is a key to understanding the disease conditions in many medical disciplines. A prerequisite for "understanding" this content, or its extraction or analysis, is a set of image processing activities, including: image capture, segmentation, compression, image manipulation, image file format conversion and related areas. This requires the digital capture of images at full 24 bits/pixel color and the subsequent processing to pare the portions of the image down to 8 bits/pixel and 1 bit/pixel in those areas where there is gray material and two-tone text respectively. The fundamental goal of this program is to further the existing inhouse capability to acquire, compress, retrieve, manipulate, segment, analyze, display and transmit digitized biomedical images.

Recent activities include: (a) The development of an Internet file server to provide access to digitized x-rays for collaborators at Stanford, Yale, Monash Universities and the IBM Almaden Research Center; (b) An analysis of selected image compression techniques (Hadamard, DCT) in removing the redundancy in a set of digitized dermatologic images.

Current work includes research in automated extraction and classification of features in digital x-ray images. This work is motivated by the possibility of using features (e.g., texture, shapes, edges) as descriptive elements in images for image retrieval. The principal tool to be used is mathematical morphology. The approach is to determine if texture or shape can be used to discriminate between, say, vertebrae and other areas of the image. If so, then the task

will be to devise morphological operations (e.g., erosion, dilation, etc.) and kernels as structuring elements to identify the vertebrae and to experiment with a sample of digital spinal x-ray images to define the optimum set of operations and structuring elements.

Engineering Laboratories

The R&D conducted in the Communications Engineering Branch relies on several laboratories that are designed, equipped, and maintained by the staff. These are described below.

Signal Processing Laboratory. Housed in this laboratory are 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.

Networks include both Ethernet and Token Ring at backbone rates of 10 Mbps connected to other local area networks throughout the building and to the Internet.

Image Processing Laboratory. This laboratory supports the investigation of image processing techniques for both grayscale 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 Solbourne workstations, and IBM PC-compatible computers, all linked via an Ethernet local area network.

The Sun/Unix machines include a Sun 690MP SPARCserver with a large magnetic disk storage capacity (6.4 Gigabytes). This machine operates as a file server for Unix machines both within and outside the Image Processing Laboratory. Additional storage for the 690MP is provided by a 9-track tape drive and an 8mm tape drive.

Large-volume storage is provided by a jukebox containing 144 5.25-inch rewriteable optical platters, each formatted to contain files in the format required by the Unix file system. Each platter has a storage capacity of 586 Megabytes, for a total jukebox storage capacity of 81.5 Gigabytes. The host computer for the jukebox is a Sun 670MP computer with 2.4 Gigabytes of magnetic disk storage, a CD-ROM player, and an 8 mm 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 to provide the user interface for DXPNET. A Sun Sparc 10 workstation is also used in the lab for the DXPNET. This machine has 400 Megabytes of magnetic disk storage.

The lab also contains Sun-compatible Solbourne workstations and a PC which are used for DXPNET project development. In addition there are three PCs in the lab attached to the Ethernet. They are used for display and grayscale 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 grayscale or color images. Other capabilities include the capture of 8 bit grayscale, and 8 or 24 bit color, from both paper and 35mm slides. Using the Howtek VSCANIT software, the Howtek Scanmaster flatbed scanner captures the images and transfers them across an IEEE 488 interface for storage. The images may be displayed on a 1024 x 856 Multisync color monitor, which is controlled by a Truevision ATVista board in one of the PCs.

Educational Technology Branch

The Educational Technology Branch (ETB) conducts research and development in computer and multimedia technologies, disseminates information about these technologies to NLM's various constituencies, and supports their application in health professions education. Branch activities include developing new technologies and liaison to health professions schools and professional societies for field testing and other collaboration. Educational systems developed both internally and elsewhere are demonstrated and information about the technology is disseminated through publications, seminars, and workshops. These systems typically combine microcomputer and optical disc technologies and use analog videodisc-based images and varied digital image formats. Recent developments in client-server distributed computing have important implications for health professions education.

Branch staff identify and demonstrate technologies and applications that meet needs in health science instruction and information transfer; they evaluate the effectiveness of educational and information technologies; and they assist health professions educators and others in improving the use of technology in health sciences education through training, demonstration, and consultation activities. Branch staff participate in the national and international health professions education, computer and information

science, and medical informatics communities.

Branch programs involve managing a Learning Center for computer-based technologies, and carrying out research in educational and information technologies.

The Learning Center for Interactive Technology

A focal point of ETB's activities is the Learning Center for Interactive Technology. The Center is the setting for the use, investigation, and demonstration of new and effective applications to faculties and staff of health sciences institutions. The establishment of the Center reflects a commitment by the NLM to develop and support innovative, computer-based approaches to training health care professionals. Computer-based technology offers these professionals a means of improving their own knowledge of the discipline as well as an opportunity to incorporate this technology in the curriculum for the benefit of their students. The Center is a facility where health care professionals can:

- gain an awareness of new technologies and of the types of teaching programs they make possible;
- become motivated to experiment with these technologies at their institutions;
- develop skills in evaluating technical components and interactive instructional programs; and
- develop the knowledge and skills required to select and use interactive technologies effectively and to measure their educational value.

The Center has three components: 1) a demonstration area; 2) an R&D area; and 3) an interactive training facility. The demonstration area has 15 carrels that allow Center staff to demonstrate a wide range of interactive multi-media programs. Large-group demonstrations are presented from a carrel configured for video projection, and are also conducted in the training facility.

The Center has many visitors. Since its inception in 1984, the Center has been visited by some 7,000 individuals, including physicians, nurses, scientists, librarians, and other health care professionals. During FY 1993 the Center had several visitors of note, including Senator Robert Kerrey, Dr. David Kessler, and Dr. Harold Varmus. International visitors came from Australia, Taiwan, India, Egypt, Guyana, Guam, and from many European countries.

During FY 1993 a major remodeling of the Learning Center was undertaken. This resulted in a number of functional improvements to the existing space. In September 1993, the Center hosted its first Open House for NLM staff.

Research in Educational Technologies

During FY 1993 two ETB research projects were successfully completed, the Computer-based Curriculum Delivery Systems (CCDS) project and the Dermatology Visual Database Project. The primary goal of the CCDS project has been to design, produce, and test experimental technology-deliverable curricula for the health professions. The project produced two videodiscs during the 1981-83 period and offered them to medical schools for testing in mid-1983. Based on data collected from students and faculty, the presentation format was revised and eight additional videodiscs were produced and offered to schools for beta-testing. The programs have been well received by students and faculty and are now in use in 47 U.S., 8 Canadian, 2 Caribbean, 1 European (U.K.), and 2 Philippine health professions schools. Some schools have used the videodiscs to replace lecture and laboratory sessions, but the majority have used them as enrichment or remediation tools. During the 1991-93 period new (second) editions of the first two programs were created to make them conform to the format adopted after the first phase of testing. These new editions have been made available to the test sites in time for the 1993-94 academic year.

In addition to the 10 videodiscs, the project has produced computer programs that control presentation of the material in self-instructional mode; 11 dBase databases that describe the contents of each videodisc (the eleventh database is a concatenation of the first 10 that can be used to prepare exams); 11 question databases for use with each lesson and a final exam; editors that permit local faculty to tailor the question databases to fit local needs and view and change the status of each student using the programs; an authoring system that enables faculty members to construct entirely new question databases using ASCII files generated by word processors; and a data extractor that enables faculty to examine the data and student comment files produced by the program.

CCDS has collected performance data and evaluation forms from over 10,000 student and faculty users over a period of 10 years. One final videodisc was produced by the end of FY 1993 that contains all of the pathology images (5932) and all of the microscopic anatomy images (5546). Such a videodisc, combined with the databases, editors, and authoring tools described above, provides health professions faculty with the tools required to produce self-instructional and examination materials. CCDS is actively seeking a private sector entity interested in future distribution of the videodiscs and computer programs.

The Dermatology Visual Database Project has had three primary objectives: to investigate image quality requirements for skin lesion recognition; to investigate archiving and retrieval techniques for a skin lesion image collection; and to investigate application of any new technologies developed to dermatology education. These objectives have substantially been attained. A method for objective assessment of observer performance with medical images has been developed. A method has been described for producing electronic images from color slide transparencies which results in dermatologists' performance in morphology recognition without significantly greater errors than from the original film. This method is now achievable at reasonable cost with commercially available equipment. The project has resulted in three test databases of color electronic images stored on three different optical media: NTSC

videodisc, CD-ROM, and Photo CD. In collaboration with the American Academy of Dermatology, project staff have developed an integrated index to the CD-ROM collection allowing retrieval by source, diagnosis, and etiology. Such image collections can be accessed locally or in client-server mode via network technology.

The American Academy of Dermatology has formed a nonprofit subsidiary, the Sulzberger Institute for Dermatologic Education, which will apply the technologies investigated by this project for the improvement of dermatology education. The Dermatology Visual Database Project has resulted in 10 peer-reviewed publications, with 3 additional manuscripts currently in preparation by the principal investigator. These publications as well as numerous other presentations on these topics have stimulated benefits in medical image processing in areas other than dermatology.

AuthorBase, a database of authoring system software created in 1991, was updated in FY 1993 and an agreement was entered into with the U.S. Army Research Institute to jointly maintain it. AuthorBase includes information on over 100 authoring tools for developing courseware and other materials for classroom use. Users can browse the database for information about currently available software and can identify those that work with different platforms, operating systems, and multimedia.

ETB's Interactive Technology Sampler was initially produced in FY 1991. The sampler is an interactive videodisc with vignettes portraying 20 interactive technology programs in the health sciences. Individual vignettes may be accessed via a main menu or index. A wide range of specialty areas is represented, including general medicine, dentistry, nursing, anatomy, pathology, psychiatry, physiology, cardiology, and radiology. Varied educational applications of interactive technology are also shown, such as visual databases, clinical simulations, hypermedia, and expert systems. During FY 1993 a videotape was produced that will ultimately result in an update to the sampler. The programs portrayed in the sampler and its update may all be seen in their entirety in the Learning Center. The sampler was developed to make information about these programs and technologies more generally accessible and can be used to illustrate lectures or can be used individually in kiosks and exhibits. Since the videodisc was designed for use with multiple control systems (barcode, videodisc player remote controllers, or Macintosh and IBM compatible computers), it demonstrates several different ways to use interactive video. The sampler is available to libraries and media centers supporting health professions education programs.

In September 1992, the Department of Biological Structure at the University of Washington entered into a 1-year research collaboration with the NLM. The *Anatomy Browser* client software developed at the University was installed in ETB's Learning Center. The Macintosh client software remotely accesses a structural information server running on a NeXt machine in Seattle. The server supports images, graphics, and a knowledge base of anatomic names and their relationships. The purpose of the collaboration is to carry out a system performance analysis of this distributed client-server based structural biology information system. The analysis has measured performance at both the client and server computers and across the Internet.

During FY 1993 several research reports were produced by ETB staff. These include the following new monographs: *An Interactive Multimedia Technology Primer*, *Guidelines for Designing Effective and Healthy Learning Environments for Interactive Technologies*, *Interactive Instruction in Nursing and other Health Sciences*, and *Computer-based Technology in the Health Sciences*.

Branch staff made presentations at several major educational technology conferences and organized and delivered a number of workshops during the year.
The workshops involved interactive multimedia, repurposing of existing videodisc systems, authoring tools, nursing leadership development, and network technologies, including ETB's Educational Technology Network (E.T.Net). In November 1992, Branch staff participated once again in the annual Radiologic Society of North America meeting by offering demos, theater presentations, and workshops. ETB hosted and participated in the Health Science Education Association's competition for the Interactive Technology Media Award in December 1992.

Research in Information Technologies

The Electronic Imaging Document System (EIDOS) project is a new ETB project. It involves research and development in electronic document imaging, indexing, and retrieval methods. The project seeks to develop methods to facilitate computer-assisted access to information in textual and image form. The project investigates document management and retrieval methods based on the client-server, wide-area network paradigm. During FY 1993 a large document collection was prepared for inclusion in the EIDOS prototype system. The collection consists of Regional Medical Programs (RMP) materials donated to NLM's History of Medicine Division. Approximately 1500 documents were cataloged and indexed in preparation for subsequent scanning into the EIDOS system.

Project staff have worked with the National Coordination Office staff of the High Performance Computing and Communications initiative to develop an in-house HPCC library. Many of these documents have also been included in the Gopher server produced by the EIDOS group. The server was successfully demonstrated at the Medical Informatics training seminar held in Woods Hole, Mass., in June 1993. Project staff are currently working on a Mosaic interface to the Gopher server. This interface provides hypertext links between important documents in the collection.

Branch staff participate in NLM's Unified Medical Language System (UMLS) project. During FY 1993 staff members worked with other UMLS participants in the preparation of the 1993 release of the UMLS Knowledge

Sources. Progress was made on a hypertext version of the UMLS documentation. This is scheduled to be released with the 1994 version of the Knowledge Sources. A preliminary project plan involving the vocabulary of computerized patient records was developed. The project will investigate the potential of the UMLS Metathesaurus and Semantic Network for representing clinical concepts and their relationships as they are represented in patient records. Acting Branch Chief, Dr. Alexa McCray, co-authored an extensive review paper on the UMLS project that appeared in the July 1993 issue of *Methods of Information in Medicine.*

Audiovisual Program Development Branch

The Audiovisual Program Development Branch (APDB) 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 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 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 is an area being explored that represents the future for improved electronic image quality. Multimedia techniques are being pursued for the educational and cost advantages 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, the central core expertise remains the creative, artistic, and communication skills of the staff.

Cardiac Embryology

APDB continues to play a leading role in a multi-institutional demonstration project in which 3-dimensional computerized representations of the embryonic development of the human heart are being created and tested in a medical education setting. Beginning in January 1993, a beta test version of the program was mounted on a Macintosh Quadra computer system in the Instructional Resources Center in the PreClinical Teaching Building of the Johns Hopkins University School of Medicine. Supported by the School's Medical Informatics Department, the program is being used and tested by medical students, whose comments and suggestions are helping to further refine the project. Utilizing the latest computer graphics systems and high-speed communications networking, the final product will be a widely accessible, desktop-computer-based, interactive hypermedia presentation, available in videodisc or CD-ROM format. Collaborating institutions include the Center for Advanced Instructional Design, Yale University School of Medicine; several departments of the Johns Hopkins University School of Medicine; the Armed Forces Institute of Pathology; the National Museum of Health and Medicine; and the Walter Reed Army Medical Center.

Cervical Cancer: Success in Sight

During the previous fiscal year, in collaboration with the Early Detection Branch of the National Cancer Institute, APDB developed an award-winning educational program on the early detection of cervical cancer. An interactive laser videodisc was initially designed for the IBM M-Motion system. Subsequently, the same videodisc was programmed for operation on the IBM Infowindow system. During FY 1993, the Branch redesigned the subject material and developed a Compact Disc-Interactive (CD-I) version of the program in order to make use of the audio, video, and programmatic capabilities of that all-digital platform. NCI's International Cancer Information Center expressed great satisfaction with the resultant program, confirming that they will assume responsibility for distributing the CD-I program, as well as any possible future collaborative interactive programs that might be developed.

Basic Medical Pathology Series

The first two programs in this interactive videodisc series were re-edited so as to match the technical and educational formats of the remaining programs. In addition, modifications and improvements were effected in the visual image data banks contained at the end of each program. Final premaster videotapes were then converted to laser disc format and made available to the Educational Technology Branch for distribution.

Educational and Information Program Support

At the request of the National Coordination Office of the High Performance Computing and Communications Program, the Branch produced a short (6.5-minute) videotape program, showing how the National Research and Education Network will provide high-speed, digital communications links to better serve the nation's research, education and information dissemination needs. The program was titled "Toward a National Information Infrastructure."

Two existing programs, "Proud Profession" and "Caesarian Birth," were augmented by graphic materials, edited, and converted to DRAW (Direct Read After Write) videodisc format and used in History of Medicine exhibits mounted in the main lobby of the Library.

Level II laser videodiscs were made from 16 short videotapes on major Lister Hill Center research and development projects (produced by APDB in FY 1992). One was set up in the NLM Visitors' Center; another was installed in the Lister Hill Learning Center. Both a Level II disc and an Alpha disc with chapter stops were provided to the Lister Hill Center Director, for use in outreach activities.

A Library-sponsored nationwide teleconference, "Information STAT: Rx for Hospital Quality," containing several Branch-produced remote location videotapes, was conducted in two-hour segments on October 22 and November 5, 1992. Tapes of this teleconference were later edited into an "Overview" and several other programs for specialized NLM-selected audiences.

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; and,
- Coordinate efforts to gather biotechnology information worldwide.

There are presently 50 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 and information science, and mathematics.

NCBI programs are divided into three areas: 1)creation and distribution of sequence databases, primarily GenBank; 2)basic research in computational molecular biology; and, 3)dissemination and support of molecular biology databases, software, and services. Within each of these areas, NCBI has established a network of national and international collaborations and also closely coordinates its activities with other NLM divisions. NCBI integrates data from NLM databases such as MEDLINE into specialized data resources for the molecular biology community.

Database Building-the NIH GenBank Sequence Database

Beginning in October 1992, NCBI assumed responsibility for all phases of GenBank production, support, and distribution. GenBank is the NIH genetic sequence database, an international database that collects all known DNA sequences and a critical research tool in the analysis and discovery of gene function. NCBI is directly involved in the creation and revision of sequence records and the distribution of GenBank via Internet FTP, e-mail servers, and CD-ROM. A group of indexers in Library Operations with specialty training in molecular biology creates new records from the journal literature and annotates direct submissions from scientists worldwide. NCBI staff and on-site contract personnel are involved in the record building and revision process, and NCBI scientists review the records for accuracy of biological information. Every effort to improve the biological accuracy of submitted data and to correct existing entries is being made by the GenBank team. New releases of GenBank have been made by NCBI every two months since October 1, 1992; daily updates are made available via the Internet as well.

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 3,600 journals are scanned for sequence data, and the NLM has expanded its journal coverage to include all journals that regularly contain sequence data even 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.

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.

An international collaboration with the EMBL Data Library in Germany and the DNA Database of Japan ensures comprehensive collections of data worldwide. Cooperative arrangements are used to augment the in-house data capture operation. 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. Further enhancement of GenBank with specialized sequence data, such as vector sequences, repetitive DNA, and protein sequences from the older journal literature, is provided by R & D contracts in place through a Broad Agency Announcement.

NCBI staff also are active in creating smaller-scale, special-purpose databases, including transcription factors, and an integrated view of E. Coli genetic map and sequence data. The Expressed Sequence Tags database (dbEST) initiated last year collects the growing number of gene fragments obtained through cDNA sequencing and now contains over 25,000 entries. Participants submit data electronically to NCBI and are provided with access to the BLAST network service, plus software tools to assist in their cDNA analyses.

NCBI is collaborating with an international working group to implement a merged taxonomic tree produced at NCBI that contains ASN-formatted representations of taxonomic data associated with sequence databases. The tree was produced using TaxMan, a taxonomy database management tool developed at NCBI. TaxMan was used to build representations of each of the sequence database taxonomies, and to merge these taxonomies into a single tree. An international working group of taxonomists is now enhancing and revising

the NCBI taxonomy.

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. NCBI also established a directory of diagnostic DNA laboratories in conjunction with Dr. Roberta Pagon, University of Washington, to assist the medical community in locating testing laboratories for genetic diseases.

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) tool is an International Standards Organization data description language that provides 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 that 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. A large number of commercial and academic developers are experimenting with the tools, and a major commercial sequence analysis package has incorporated the NCBI code. In order to disseminate information about NCBI software and database projects, and provide technical support, NCBI established this year a Professional Software Developers Forum. This forum is intended for professional software developers in the biotechnology and pharmaceutical industries.

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 DNA and protein sequence databases and the protein structure databases. ASN.1 is also being used as a distribution format for NCBI sequence and bibliographic data. The Entrez: Sequences CD-ROM and Internet client/server delivers sequence and MEDLINE data in the ASN.1 format to provide a structured version of the databases that is 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 with techniques that are fast and easy to use. 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. New releases of the database are produced every two months. A continuing effort has been made to improve the ease of use of the software, and paid subscriptions have increased from 710 (Release 1.0) to 2,325 (Release 6.0).

GenBank CD-ROMs in two other formats, the traditional GenBank flat file format and NCBI's own ASN.1 format, are also produced bimonthly. Beginning in June 1993, a companion disc called Entrez: References was distributed with Entrez: Sequences; this disc contains 200,000 MEDLINE references to the literature of molecular sequence data.

The software also runs in a network environment and versions are available that allow a user (client) anywhere on the Internet to formulate queries on a local workstation and send off requests to a server at NCBI. After several months of testing, Network Entrez was officially announced in August 1993. It has been well received, and more than 250 sites were registered by October. Approximately 150 individual users were using the system 400 times per weekday at that time.

NCBI is also using a server-client approach in building systems for network sequence searching. The BLAST sequence searching server has been implemented as a network-based retrieval system. More than 100 major sequencing centers and research institutions around the country use this software 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 2,000 search requests are processed by network BLAST each day with a response time of under 15 seconds for many queries. 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 research.

NCBI's electronic mail servers, RETRIEVE and BLAST, are heavily used by the scientific community. 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 search the nucleic acid and protein sequence databases for matches to a user's sequence using the BLAST algorithm to identify sequence similarity. In the past year, a public key encryption option has been implemented to guarantee the confidentiality of user data as it traverses the public networks. A third server was made available in May 1993 for retrieving reports on cDNA sequences from the Expressed Sequences Tags database (dbEST). Any user in the world with e-mail access can submit a query to the servers and have an answer returned within minutes. More than 2,000 queries are handled daily by the three electronic mail servers.

GenBank is also distributed over the networks through the standard File Transfer Protocol (FTP) program. More than 1,000 requests are processed daily for downloading files from NCBI's public data repository, including 150 of the daily GenBank update files. The entire database of over 425 megabytes as well as daily updates are available to network users. Thirty additional databases of interest to the molecular biology community are distributed through the NCBI Data Repository. These databases have been provided to over 500 requesters on a CD-ROM as well.

In FY 1993, NCBI substantially upgraded its networking and computing infrastructure. NCBI's most powerful multiprocessor servers and uniprocessor workstations were moved from an Ethernet-based network to an optical fiber-based FDDI Network, increasing the network bandwidth by more than a factor of ten. Other desktop systems were migrated from Ethernet to 10BaseT twisted-pair Ethernet, a more modular and manageable network architecture. In addition, one new RISC-based multiprocessor Sun SPARCecenter 2000 computer was acquired and two older model Sun servers were upgraded to SPARC center 2000 systems. These powerful computers, which are expandable to 20 processors and terabytes of storage, support NCBI's e-mail and network client-server applications and production relational databases. A number of smaller high-performance workstations were acquired from Sun and Silicon Graphics for use as servers for smaller databases or for molecular modeling applications, respectively.

Basic Research

Basic research is at the core of NCBI's mission. The Basic Research and Information Engineering Branches at the NCBI are made up of a multidisciplinary group of scientists who carry out research on fundamental biomedical questions at the molecular level by developing and utilizing mathematical, statistical and other computational methods. The approach is both theoretical and applied. These two lines of research are mutually reinforcing and complementary. The basic research has led to new practical methods and the application of these methods has opened new areas of research.

Several analytical algorithms and methods have been developed to complement the investigation of biomedical problems. Analytical methods and algorithms have focused on sequence similarity, structural modeling and genome analysis. Pattern matching and sampling methods for biopolymer sequence data, which include the statistics of sequence comparisons, have been applied to investigate such protein families as the ras-line GTPases, steroid receptors, cold shock domain proteins, HMG-1 box-containing proteins, and NTPases, as well as other analyses of different sequence motifs. Biomolecular structural investigations include such projects as the analysis of packing contacts in protein crystals, 2D lattice models of proteins, 3D modeling of ribonucleic acids, protein threading and the energy distribution of compact states of a peptide. Genome analyses have been a significant part in both the methods development and the scientific analysis of several different genomes including that of E. coli, RNA viruses and humans.

In order to attempt to unify biotechnology information, explicit descriptions for biosequence objects have been defined using an international standard data description language, ASN.1. Innovative approaches have been taken for database design to permit the integration and linkage of vast amounts of sequence-related data. ASN.1 software libraries have enabled the rapid development of software tools for the retrieval and analysis of information from these databases.

Text information retrieval methods, document analysis, and evaluation of retrieval methods have been important components of the basic research. New techniques in use and under evaluation for the analysis of large amounts of text include Gibbs sampling approaches, Bayesian methods and the building of neighboring lists.

The intramural group is engaged in 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, more than 50 in FY 1993, and the requests for outside collaborations.

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 of computer science and biology. NCBI has also extended its outreach to the library science community by invited presentations and workshops on biotechnology information topics.

The Visitors Program continues to be 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. This program, administered in conjunction with Oak Ridge Associated Universities, facilitated over 60 visits from senior

researchers in its third year.

NCBI continues to work with Dr. Elvin Kabat to update and maintain his database of protein sequences of immunological interest, an important resources for immunologic researchers.

As NCBI's involvement in GenBank has expanded, the need for increased user support has grown. Contract staff has been added to complement existing NCBI staff in the Information Resources Branch who provide responses to telephone, e-mail, letter and FAX queries for information and assistance. The three main areas of user support include: submission of new GenBank data and revision of existing records; requests and support for software that assists authors with the creation of an input record to GenBank; and technical assistance with *Entrez* and other data retrieval services. Most responses are immediate and nearly all answers or information are provided within 24 hours of receipt of a message. Likewise, authors who submit their sequences to GenBank are furnished with accession numbers for publication within 24 hours.

To increase awareness of the Center and its programs, NCBI staff participate in exhibits, seminars, workshops, and courses, both nationally and internationally. NCBI staffed exhibits at scientific society meetings, including FASEB-Experimental Biology '93, American Society for Biochemistry & Molecular Biology, and the American Society for Cell Biology. In addition, senior NCBI members participated as faculty at courses sponsored by the American Association for Cancer Research, the United Nations, and Johns Hopkins School of Medicine/The Jackson Laboratory. Several staff demonstrated individual projects at the Waterville Valley meeting on Macromulecules, Genes, and Computers; and NCBI was the co-sponsor of a workshop on "Computer Dabatases and Software for Molecular Biomedical Research" at the NIH Research Festival. NCBI hosted a workshop on taxonomy, where NCBI's merged taxonomic tree was reviewed and revised by a panel of expert systematists. Three issues of a newsletter were distributed to a mailing list of 18,000 biologists and institutions, and new Fact Sheets on programs and services were distributed at all public forums where NCBI was represented. On the NIH campus, more than 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. Within the NIH, the NCBI coordinates with other institutes and particularly with the National Center for Human Genome Research on databases and informatics programs that 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. The scope 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 explosive growth in the fields of genetics and molecular biology reinforces the need to build and maintain a strong infrastructure of information support. NCBI will continue to develop and employ new methods for disseminating knowledge to the biomedical community. Based on a core of advanced intramural research in several areas of computational biology, NCBI can rapidly address the evolving informatics needs of researchers by developing state-of-the-art software and databases that function as integral tools in the research process.

EXTRAMURAL PROGRAMS

Milton Corn, M.D.
Acting Associate Director

Introduction

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 the Medical Library Assistance Act was awaiting reauthorization by the Congress, EP's MLAA budget was kept at FY 1992 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. An exception is the Connections program, a new program supported with HPCC funds and designed to promote connections of health science sites to Internet; it is described below in the Resource Grants section.

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.

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 1993 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 (NCI) 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 the National Heart, Lung, and Blood Institute and NCI to fund projects originally assigned to NLM, which had high priority scores and informatics themes of interest to the other Institute.

In FY 1993 NLM joined a number of other Institutes and Federal agencies in the Human Brain Project, a major effort to apply informatics to neuroscience under the lead of the NIMH.

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, judicious administrative site visits 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. Applying information science to modern health care and research poses complex problems whose solution will depend on well-trained specialists. NLM supports both institutional training programs and fellowship programs.

The ten five-year institutional training grant awards made in 1992 entered their second year in FY 1993. Of special interest in FY 1993 was the relatively heavy response to the newly created fellowship in applied informatics, which is open to almost all health care professionals, and is designed to promote training of experts in applying informatics to the clinical, educational, research and administrative problems of health centers. Although only three could be awarded, the large interest in the program is gratifying.

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 can be realized by linkage of the various information systems present in the library, research sites, administration offices, medical education operations, hospitals, and outpatient areas.

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.) During FY 1993, eight planning grant applications, and five implementation grant applications were reviewed, the latter by site visit. Decisions about awards have been deferred to FY 1994.

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 all 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, inner city or other underserved areas, to establish facile contact with the national biomedical library system.

A revised program announcement describing Resource Grants was issued in 1993 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 on Library Operations. The Special Foreign Currency Program, administered by the Extramural Programs' International Programs Branch, is described elsewhere in the annual report under International Programs.

Training (MLAA)

Research in informatics demands highly trained, creative individuals who are able to apply information management, computers, and telecommunications to biomedicine. There is an acute and growing need in academic medicine for medical informaticians. Through its training program, NLM provides grants for career training in this field.

In 1993, ten institutional awards were made. These ten awards provide training at 15 separate institutions. Each site offers an excellent setting for didactic instruction, involvement in computer science studies, and opportunities for work in advanced information science research. Because of their training, these investigators will have the ability to study the role of knowledge in professional life, analyze the social structures for managing knowledge, and advance the frontiers of the computer sciences for organizing, retrieving, and utilizing health knowledge.

Sixty-five postdoctoral and 35 predoctoral trainees were supported in FY 1993. In addition, funds were made available to support 13 students in health professional schools who wish to spend one or more summers or semesters working in laboratories associated with this program. In this way, such individuals will become exposed to informatics research and issues, and may return to this field for additional training.

The ten awards, some with co-principal investigators, and the fifteen training sites:

Robert Greenes
Harvard Med. Sch., & Brigham Women's Hosp.

Peter Szolovits
Massachusetts Inst. of Technology

Octo Barnett
Massachusetts General Hospital

Steve Pauker
New England Medical Center

Ed Hammond
Duke University

Charles Friedman
University of North Carolina

Perry Miller
Yale University

Paul Clayton
Columbia University

Joyce Mitchell
University of Missouri

Kent Spackman
Oregon Health Sciences University

Ted Shortliffe
Stanford University Medical Center

G. Anthony Gorry

Rice University

Randy Miller
University of Pittsburgh

Lael Gatewood
University of Minnesota Health Sciences Center

In addition to the institutional training grants supported by NLM, an individual fellowship program is also available. The postdoctoral research fellowship is patterned after the institutional training program, and is intended to support individuals who are planning a research career in the field of medical informatics. Such awards allow training to occur at qualified centers of excellence other than where Institutional training grants are located. In FY 1993, two new fellowships were awarded at the University of Utah and the University of Washington.

A new fellowship program, initiated in FY 1992, is in applied informatics. There is a growing need in all areas of the health sciences to have people sufficiently trained to apply the concepts and technologies of informatics to their particular area of expertise. Applicants to this program need not possess a doctoral degree. Thus it is of interest to people in the fields of librarianship and nursing where doctoral degrees are not the normal professional degree, as well as to those holding M.D. or Ph.D. degrees. This program, while new, has generated heavy response. Fellowships were awarded at the University of Washington and Duke University in FY 1993.

Research (PHS 301)

Through a variety of research grants, the Library seeks new understandings of medical knowledge and new ways of using knowledge more effectively for health care, research, and education. For many years, NLM has emphasized research in different aspects of informatics. Informatics is usually understood as the profession, or investigation, of managing information in all its variety, with special reference to the use of computer and telecommunications technology. The computer science problems of representing knowledge offer numerous challenges, but because information has personal and social dimensions, informatics also includes large elements of social and behavioral science. In NLM's research programs there are further distinctions among medical informatics, biotechnology informatics, and health library information science. Within these areas are further distinctions between research on fundamental questions and on more immediately applicable work.

Medical Informatics

In addition to continuing support for previously initiated project grants, NLM awarded six new, competitively reviewed grants this year. Only one of these awards was for more than $200,000. In every case, a budget reduced from the original request was either negotiated by staff or was offered by the principal investigator. There is necessarily some reduction as well in the scale of effort, but it is clear that the essential research objectives will be undertaken and that the validity of the research design is not compromised. Although the new awards are few, the range of research activities is large, a fitting match to the large program scope.

At the University of Pittsburgh, Dr. Gregory Cooper will investigate and evaluate techniques for automating Bayesian networks for large clinical databases. This work is fundamental and aims forward to eventual improvement of computerized diagnostic aids.

At the University of Illinois, Professor Arthur Elstein studies the effects of decision support systems on clinical reasoning through research on cognition and information related behavior.

At Yale University, Dr. Conrade C. Jaffe won another award to continue his research into novel, nontextual methods for indexing and retrieving from large clinical image databases.

At the University of Washington, Professor Ira Kalet was given another grant to continue his investigation of computer algorithms for radiation therapy planning. He seeks to overcome dosage calculation problems through artificial intelligence methodology.

At George Washington University, Dr. William Knauss is evaluating the interaction between health care staff and his intensive-care information system called APACHE. The evaluation plan is based on well accepted statistical methods. APACHE is one of the few sophisticated decision support systems in clinical use. Assessing its effectiveness and utility has implications for the entire medical informatics field.

Biotechnology Informatics

In the area of biotechnology informatics, research is supported to investigate effective methodologies for organizing and analyzing data related to molecular control of life processes. Relevant problems include designing and managing databases, retrieving information from multiple factual databases, and pattern-matching algorithms for biological sequences. Recognizing that scientific advances in molecular biology will depend in large part on successful resolution of knowledge and information issues, NLM devotes about $4 million to this program.

In addition to support of 11 previously initiated grants, a new award in FY 1993 was made to MIT for "A Clinical Geneticist's Workstation."

In FY 1993 the biotechnology program also supported a successful conference on the analytical aspects of biology. This conference was held in the Lister Hill Auditorium at NLM and was over-subscribed. Proceedings of the papers presented have been published.

Health-Science Library and Information Science

This program area concerns information issues and problems which relate directly to medical bibliography and to

the functions or interests of libraries. Because this work almost always involves computer science as well, the boundary between this area and informatics is often indistinct. For example, one research topic placed here is information-related behavior. Another is the use of the literature, in reports of controlled trials, to derive conclusions through a pooling of data in meta-analyses. These topics obviously concern other fields as well as librarianship.

This year, available funds supported two research projects. Professor MaryEllen Sievert of the University of Columbia-Missouri will continue her investigations of retrieval from various kinds of databases, including full text databases. Full text is becoming increasingly available, but earlier work shows that the usual searching strategies for indexed literature do not work as effectively on full text. In this project, Sievert will explore new, more effective searching tools and strategies.

At Sloan Kettering Institute, Professor Colin Begg is studying aspects of publication bias in clinical trials. The goal is to improve methods for meta-analyses. In this way, bibliography can serve to confirm clinical trial results more quickly and reliably.

Resource Grants (MLAA)

Access and Systems Grants

The Information Access Grant (directed towards small and medium-size libraries) and the Information System Grant (for larger ones) continued to further the use of computer and communications technology for access to health science information. In FY 1993 all of the grants awarded represented projects related to networking, connectivity, and integration either locally, regionwide, or statewide.

Four Information Access Grants were awarded to promote Grateful Med access in health science library consortia in North Central and Western Kansas (University of Kansas-Wichita), Western North Carolina (Mountain AHEC in Asheville), South Alabama (University of South Alabama in Mobile), and Northwest Ohio (The Toledo Hospital). Another Information Access Grant was awarded to the Northeastern Ohio Universities College of Medicine (NEOUCOM) in Rootstown for the purpose of installing a communications network among the teaching hospitals in order to share information resources such as an online integrated library system and MEDLINE. An Information Access Grant to the Southwest Georgia Area Health Education Center will enable 12 sites to connect to the Georgia Interactive Network for Medical Information (GaIN) which offers library and information resources and services.

An Information Systems Grant was awarded to the Bowman Gray School of Medicine (Winston-Salem, NC) to equip students with portable computers, software, modems, and printers to enable them to access the library and other medical databases during their preceptorships at remote sites. Three grants were awarded to automate library technical service functions: the libraries of Millard Fillmore and Buffalo General Hospitals will install a joint integrated library system; seven Denver area hospital libraries through the "Shared Automated Library System (SALS) Enhancement Project" will connect to the statewide CARL library system; and the Ellis Hospital in Schenectady, New York, will install an online public access catalog to be accessed through a local area network.

NLM grant assistance also benefited minority populations. An Information Access Grant to D.C. General Hospital on behalf of the libraries of the Hospital Council of the National (D.C.) Capital Area will enhance the interlibrary loan network for the rapid delivery of materials in emergency situations. An Information Access Grant to the Salish Kootenai (Tribal) College on the Flathead Indian Reservation in Pablo, Montana, will enable the library to become a member of the Western Montana Medical Libraries Access Network which is implementing a Grateful Med project.

IAIMS Grants

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, the IAIMS Program was extensively revised with important changes in scope, objectives, and funding levels in response to NLM's 10 years of experience with the program, and the changes that have occurred during that time in biomedical culture and technology (details of the revision were given in last year's annual report). Hundreds of inquiries have been received in response to the change, and the first grant applications under the revised program were received and reviewed in FY 1993. Thirteen applications were submitted: eight applications requested support for IAIMS Planning, and five requested IAIMS Operational support. Two of the Planning applications were from hospitals, which was one of the objectives of the revised grant. Reviews of the applications were undertaken during the summer: all five Operational applications received priority scores and were recommended for further consideration as were four of the Planning applications; the remaining four Planning applications were not recommended for further consideration. Decisions about grant awards for these applications will be made early in FY 1994.

During FY 1993, continuation grant awards were made to Baylor College of Medicine, Columbia University, Duke University, and Georgetown University for IAIMS implementation. A contract to the Oregon Health Sciences University for IAIMS implementation also remained active. The first year of a two-year IAIMS Planning grant was

awarded to Vanderbilt University. In FY 1992 NLM awarded three High Performance Computing and Communication (HPCC) grants as a one-time special initiative to certain IAIMS institutions; continuation grants for these projects were made in FY 1993 to Dr. Edward L. Chaney at the University of North Carolina, Dr. Walter B. Panko at the University of Michigan, and Dr. Gordon K. Springer at the University of Missouri-Columbia.

Internet Connection Grants

For the second consecutive year NLM entered into an interagency agreement with the National Science Foundation to offer grants to health science organizations to connect to the National Research and Education Network (NREN). Grants for this purpose are available for an initial "hook-up" for $30,000 or for an extension of an extant connection to additional sites for $50,000. In FY 1993 57 applications were submitted resulting in 16 awards totaling $522,185 (NLM contributing $400,000 and NSF $122,185) to: George Washington University Medical Center (Washington, D.C.), Georgetown University Medical Center (Washington, D.C.), Mc Lean Hospital (Belmont, MA), Methodist Hospitals of Memphis, Methodist Medical Center of Illinois (Peoria, IL), National Jewish Center for Immunology and Respiratory Medicine (Denver, CO), National League of Nursing (New York, NY), New York Academy of Medicine (New York, NY), Northwestern Memorial Hospital (Chicago, IL), Rochester General Hospital (Rochester, NY), St. Louis University Medical Center, Tufts University School of Medicine (Boston, MA), University of Illinois-Carle Foundation Hospital & Clinics (Urbana, IL), University of South Dakota (Vermillion, SD), University of Virginia Medical Center (Charlottesville, VA), and University of Southern California (Los Angeles, CA).

Publication Grants (MLAA)

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 chapter on International Programs.)

During FY 1993, NLM awarded nine Publication Grants totaling $269,000. 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 1993, including both direct and indirect costs, was under $30,000. A list of grant-supported publications that appeared in FY 1993 is in Appendix 3.

In addition to five grants receiving continuing support, four new publication awards were made in FY 1993:

- Borst, Charlotte; University of Alabama: Midwives, Physicians, and the Professionalization of Childbirth
- Jones, Kathleen; Virginia Poly. Inst.: Development of American Child Psychiatry
- Vilensky, Joel; Indiana University: Evaluation of the Denny-Brown Research Collection
- Tomes, Nancy; SUNY-Stony Brook: Spreading the Germ Theory of Disease

Bioethics (MLAA)

NLM continued to support a National Reference Center for Bioethics Literature at Georgetown University. This Center has developed what may be the most significant collection and reference service in the world for bioethics. There are two aspects to the Center's operations. The first concerns the collection and organization of material, assistance to users, and publication of various bibliographic aids. The second concerns the organization of a bibliographic database which NLM makes available, with some modification, as BIOETHICSLINE, one of the MEDLARS group of medical databases. The first is supported by an NLM Specialized Center Grant, the second by an NLM contract. The Specialized Center Grant makes possible a unique collection drawn from a variety of literatures (e.g. medicine, biology, jurisprudence, political science, and moral philosophy). It consists of 19,000 monographs, 100,000 separate articles or documents, and 250 current periodicals.

The services are offered not only through the online database, but also through a toll free telephone line. The Center also publishes a number of bibliographic aids which have been highly praised. In view of the growing number of bioethical issues, and consequent social concern, the Reference Center's services and availability to the national health care community have steadily grown in value and importance.

Small Business Innovation Research Grants (PHS 301)

These grants assist small businesses to bring new ideas to commercial practicability by supporting research and development. The grants are mandated by law for all executive branches of the Government where there is a research authority. The funding is based on an assessment on each program's research appropriation and cannot be used for any other purpose.(Unutilized funds must be contributed to an agency pool for use elsewhere.) NLM's funding level this year was $105,000.

Three grants were awarded in FY 1993. One will

support development of a CD/ROM disk for current population surveys, another will support development of an interactive, animated computer learning program for embryology, and a third will supplement the evaluation of an innovative literature service for internists.

Conference Support (MLAA, PHS 301)

Encouraging disparate disciplines or specialties to combine forces in the solution of what are major informatics problems is a significant program goal for the Extramural Programs. Such work requires collaboration in the fullest sense. The research problems must be problems of serious interest to both, and resulting papers should be co-authored and should appear, as appropriate, in the refereed literature of both fields. Workshops, symposia, and conferences can help initiate this interdisciplinary collaboration. After sponsoring a 1992 workshop on "Creating an Infrastructure for Intelligent Systems In Molecular Biology," NLM, in 1993, awarded a conference grant to one of the workshop organizers, Professor Jude Shavlik of the University of Wisconsin. The workshop and the following grant-supported conference offer a good model, one worth pursuing in other areas, such as informatics issues in the development of large scale, interactively shared image databases, information retrieval, and others. EP plans to continue supporting such workshops from time to time.

Minority Support Activities (MLAA, PHS 301)

EP provides additional funds to resource grant principal investigators who wish to train minority students. EP also strongly encourages resource grant applications from health institutions which serve minority populations, both rural and inner-city. When funds permit, EP provides additional funds to R01 principal investigators who wish to add a minority scientist to the project; one such was funded in FY 1993.

Scientific Review

The Biomedical Library Review Committee (BLRC) is NLM's initial review group. The Chairman of the full committee during FY 1993 was Dr. Carl Jaffe. Ms. Sherrilynne Fuller was appointed Chairperson of the BLRC for FY 1994. This Committee reviews about 90% of all grant applications which are assigned primarily to the Library. A roster of BLRC members is in Appendix 8.

The Biomedical Library Review Committee met three times in FY 1993 and reviewed 102 applications; 90 were recommended for further consideration. The Committee operates as a "flexible" review group; i.e, it is composed of three standing subcommittees: 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. (A list of the Regents is in Appendix 5.) 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.

Over the years, the BLRC has been concerned with the evaluation components of grant applications. In an attempt to explicate the various types of evaluation schemes that are applicable at various stage of project development, some committee members have recently published a paper on the topic. (Stead WW, Haynes RB, Fuller S, et al. Designing medical informatics research and library resource projects to increase what is learned. J Am Med Informatics Assoc; 1994; 1:28-33.) The paper attempts to provide some guidance to potential applicants, and to assist the review committee in their deliberations of specific applications. Dr. William Stead, the primary author on this paper, was formerly a member and chairman of the committee.

In a similar vein, Virginia Bowden and other members of the Committee published a paper (Bowden, VM. National Library of Medicine resource grants: application and review. Bull Med Libr Assoc 80 (2) April 1992, 157-168.) seeking to be of assistance to potential applicants to the resource grant program. The paper details and explains the criteria used by the committee when reviewing resource applications.

Special Activities

An extensive program of monitoring legislation pertinent to EP and the NLM continued, as did a host of activities designed to inform law-makers, other executive branch agencies, and relevant national organizations of NLM's missions and programs.

Plans for FY 1994

Internal reorganization of EP is nearing a final draft and will be presented to the Director, NLM, in the fall of 1993.

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.

An RFA for Biotechnology Databases will be issued in the fall of 1993 to include applications for the maintenance of protein databases in the field of molecular biology. Data sources may be sequence data, NMR structural data or other. It is expected that two awards will be made.

In preparation is a cooperative agreement RFA dealing with the general topic of the electronic patient record, one of the issues specifically mentioned by Congress in language relevant to HPCC.

Table 10

Extramural Grant and Contract Program
(dollars in thousands)

Category	*FY 1991*		*FY 1992*		*FY 1993*	
	No.	*$*	*No.*	*$*	*No.*	*$*
Research	49	11,231	51	11,863	51	11,674
Resource projects	17	4,721	14	4,802	16	4,289
Resource access	17	684	7	493	9	589
Training	7	2,714	10	3,520	10	3,878
Fellowships	9	306	9	298	6	260
Regional Medical Libraries	8	5,500	8	5,482	8	5,500
Publications	13	365	10	283	9	269
(IAIMS projects)	(9)	(3,693)	(3)	(3,938)	(5)	(3,150)
(Med. info. research)	(26)	(6,066)	(29)	(6,894)	(31)	(6,899)
(Biotech. research)	(16)	(4,192)	(13)	(3,759)	(13)	(3,915)
Totals:	**120**	**$25,521**	**109**	**$26,741**	**109**	**$26,459**

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, online, 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 thousand 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 Information Man agement Branch;
- executed on computers under operating system control of the Systems Branch; and
- provided as an around-the-clock service by the Computer Services Branch.

To improve access to NLM data and information, OCCS implemented a number of initiatives in FY 1993. Chief among these are: 1) the development of an Internet server to permit the medical community to obtain access to NLM publications. The use of this facility, called NLM PUBS, is discussed more fully in the Library Operations chapter; 2) the provision of monthly awareness MEDLINE search results on-line over the Internet. This capability permits users of NLM's SDI services to obtain their bibliographies more rapidly and in machine-readable form; and 3) the development of a computer-to-computer link between the NLM 3090 MEDLARS system and the SIS TOXNET system. This project required extensive development of computer software and LAN network capabilities and permits a MEDLINE user to switch automatically to the use of a TOXNET file (and back) with a simple "File" command. Switching of a TOXNET user to MEDLINE is also possible.

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 (LAN) services, developing Information Systems Laboratory (ISL) projects, and developing a new NLM Locator system.

The Technical Services System (TESS), which will integrate various functions of the Technical Services Division, is being developed as a distributed processing system that integrates mainframe computer, personal computer, LAN, and database technologies.

TESS development was initiated in 1989. The first implementation under TESS provided for creating and maintaining cataloging data. 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 creating and maintaining the name authority file.

The next major release of TESS (scheduled for November 1993) will include the conversion of all 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, principally the distribution of catalog records in MARC format to MARC subscribers.

During FY 1993, Grateful Med continued to be the main access point for NLM's many databases and services. Over 3 million Grateful Med searches were performed representing over 70% of the total ELHILL use by individual user codes. Almost 15% of the Grateful Med searches were performed using the Internet (TCP/IP). This access method is becoming more widely used as additional TCP/IP scripts are implemented. Version 6.0 of Grateful Med currently supports Wollongong, NCSA, LAN Workplace for DOS and PC/TCP protocols.

A test version of Grateful Med Version 6.0 containing the HSTAR (Health Services and Technology Assessment Research database) form screen access was developed this year. This version was distributed to a set of testers for evaluation. As a result of this testing, HSTAR will be included in the 1994 MeSH update package which will be distributed to all PC Grateful Med Users.

A new Apple Macintosh release of Grateful Med, Version 2.0, was completed and distributed this year. This version contains a number of new features and enhancements which include the document ordering capabilities of Loansome

Doc, Internet access, and new MeSH window features with "expanded" MeSH tree display. Almost 8,000 automatic updates were mailed out when Version 2.0 became available.

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 or DOCLINE library. Loansome Doc, which became fully operational in 1992, has been integrated into Grateful Med Version 6.0 for the PC and into the Apple Macintosh Version 2.0 and hs now been now distributed to all Grateful Med users. The reports from those health professionals who have ordered documents using Loansome Doc continue to be positive.

A collection of Local Area Networks forms the internal communications systems of NLM. Together, these various LANs support approximately 700 NLM staff, contractors, and 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 LAN systems were upgraded to incorporate newer technology in 1993. NLM installed a high-speed fiber optics backbone network that interconnects medium-speed Ethernet subnets and thus provides a single, consolidated architectural approach to its LAN needs. The high-speed backbone also provides connections to other high-speed networks such as the NIH RESNet and eventually NREN.

The Information Systems Laboratory (ISL) was created within the Development Branch in 1991. During FY 1993, the ISL has supported the development and implementation of Locator, the NLM public access catalog; NLMPUBS, the anonymous FTP service providing online copies of NLM technical publications; and Implement, a meta-DBMS toolkit designed to address the special problems of bibliographic data storage and retrieval. The ISL also supported development for remote cataloging and indexing activities.

The ISL is continuing to introduce 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 facilitate the provision of remote NLM services both domestically and internationally. An additional goal of these new systems is to provide Internet-based remote use and ultimately to provide better support for remote cataloging and indexing activities.

In 1991 OCCS developed a prototype Online Public Access Catalog (OPAC) that became a full-scale development effort in FY 1992. The operational system, named NLM Locator, was introduced to the Reading Room on February 22, 1993. In May the system was made available over the Internet and usage continues to grow. NLM Locator provides direct access to the NLM collections through the ELHILL databases CATLINE (monographs), AVLINE (audiovisuals), and SERLINE (serials).

This project was OCCS's initial Unix-based client/server development effort and the development team was challenged to acquire skills in new technologies. The complete project included mastering new skills in new computer hardware, system software, networking, programming languages, and a number of utility tools.

One of the important concepts of the client/server architecture is the ability to adapt to change without reengineering the application. NLM Locator utilizes a workstation client communicating with function servers that in turn communicate with a data server. The client workstations in the Reading Room are DOS PCs, while Internet users log in as VT-100 terminal sessions to a Unix client process executing in the function servers. The function servers are Sun and IBM computer systems running the Unix operating system. Having multiple function servers provides greater reliability as well as additional capacity, should usage demand it. If necessary, another function server could be added without software modification. The data server is the NLM mainframe computer utilizing the ELHILL retrieval system. No changes to the legacy systems were required to implement NLM Locator.

The system continues to function well and comments received via the electronic mail feature have been most complimentary. Reading Room and NLM staff usage has stabilized. Internet usage continues to grow with an average of 20 first-time NLM Locator users per day.

The second phase of the project is to provide circulation control and collection management features. This effort is well under way and will provide online registration, availability data, status information to the patron, and a great number of management reports to the Public Services Division.

Information Management Branch

The Information Management Branch (IMB) of OCCS supports the various NLM programs and serves as the nucleus of all automated programming support services. The following is a summary of the major accomplishments.

The DOCLINE Interlibrary Loan (ILL) system, NLM's online facility for requesting library materials was converted to a Virtual Telecommunications Access Facility (VTAM) environment on the powerful IBM/3090 mainframe computer in FY 1992. In FY 1993, a number of major enhancements were made to the DOCLINE system to facilitate interlibrary loan activity in the national and international medical community.

IMB installed a set of enhancements requested by the medical library community. The most significant was an increase in the number of libraries in a routing table from 80

to 180 organized in 9 hierarchical cells instead of 7. This allows finer tuning and a larger selection of possible lending libraries before resorting to NLM to fill a request. IMB also provided the capability to begin and end routing in a specific cell, allowing greater user control of requests, and the messages provided to client libraries about transactions were improved. Program enhancements were made to allow foreign libraries (specifically Canadian) to participate in DOCLINE either directly or through Loansome Doc and a function will be installed to receive loans in OSI/EDIFACT format for libraries wishing to process loans automatically in a standard machine-readable format.

To facilitate Internet client access the Branch introduced a command to download sets of responses, either to a local printer or as a dataset on the mainframe for later high-speed FTP file transfer. IMB also completed a new DOCLINE interface for Internet clients and NLM personnel with 3270 emulation capability. It builds on the VTAM panel capability to provide full screen interface with more efficient response saving features. Work has been completed for the BORROW function and is in process for the remaining DOCLINE functions. These features should significantly reduce communications costs and efforts to input and update interlibrary loans and provide a client/server framework for future platforms.

Modifications were made to allow Loansome Doc users to order documents from the HSTAR file. Software enhancements were developed that allow NLM to track and bill requests from Federal libraries based on a number of various agreements. New software and databases were prepared to provide information about journal articles requested and filled by NLM and throughout the DOCLINE network. These reports are essential to NLM management planning for anticipating collection access requests as well as understanding information delivery in the network. New statistical reports about Loansome Doc usage were provided to NLM and the regional libraries. Also, reports about daily communications line usage were provided to NLM to encourage Internet use.

Major enhancements to NLM's publications systems included: 1) refinements to the *Bibliography of Bioethics* to improve sorting and cross reference algorithms as well as a new capability to insert italicized text within a field; 2) enhancements to the MeSH publications to improve accuracy of medical terms and the use of various fonts and character sets; and 3) a single character for left and right quote to conform to ANSI standards.

New generations of software for subsystems of the Automated Indexing Management System (AIMS) became operational. AIMS is an IBM mainframe computer application which runs under the Customer Information Control System (CICS). It provides access to the Inquire Data Base Management System (DBMS) for storage/retrieval of new records or old records to be maintained. The bibliographic data entered, verified, and validated become part of the MEDLARS databases and associated publications. Subsystems affected were Journal Control, Indexing, Checkin, Bibliographic Processing, Binding, and Gapping. Enhancements include a capability to view master serials, binding, and gapping data interactively while being attached to other data entry subsystems such as the Checkin subsystem. This helps to ensure that incoming data will be handled properly. The online interactive full-screen processing mode of the various AIMS input subsystems greatly facilitates timely capturing and accuracy of NLM bibliographic data.

Additionally, new generations of software for the In Process (INPROC) and Literature Selection (LSTRC) CICS systems were created. The INPROC system provides an efficient mechanism to capture and control INPROCess information related to books and monographs that will become part of the NLM collection. Likewise, the LSTRC system provides the capability of collecting and accumulating information about journals. The information is evaluated and processed during different states of the literature selection process. Like AIMS, these data entry verification/validation systems are an IBM mainframe computer application that runs under the Customer Information Control System (CICS).

Other major software enhancements were made to the Model 204 DBMS based MeSH system. One such enhancement was the ability to process Unified Medical Language System data. This system provides for interactive timely, efficient, and accurate data entry/verification and validation of MeSH. Appropriate data are extracted daily from the Model 204 MeSH database and updated to the MEDLARS information retrieval MeSH database. These controlled thesaurus data are used to formulate searches of the MEDLARS online databases.

A beta version of the Medline Citation Maintenance System (MCMS) was made available in September 1993. This system provides for individual citation maintenance and new record creation of MEDLINE bibliographic data. The system is a client/server application, written in the "C" programming language and uses Novell LAN and Btreive DBMS facilities. The system architecture nucleus was designed to allow creation of other similar file maintenance systems in a timely and efficient manner.

More than 10% of the NLM bibliographic data of more than 7 million records were class maintained this year. Class maintenance is the adding of new terms, deleting old terms and replacing terms with preferred ones in the MEDLARS database records. Moreover, new data fields are introduced to the records as required. Major software enhancements to support the class maintenance effort were: processing of additional "Publication Type" data; removing the MeSH "ZN Tree Number" and replacing it with "country" values; and identifying inconsistent data between MeSH chemical records and data carried in the MEDLINE family of files.

A pilot study was performed to determine the cost effectiveness of using Optical Character Recognition (OCR) technology as an alternative to keyboarding bibliographic citations. It was determined that the technology can be used

effectively for certain journals, and a limited production implementation is being planned for FY 1994.

Staff of the Information Management Branch is collaborating with John Wiley (publishers) and writing conversion software to convert their journal citation data encoded in the Standard Generalized Markup Language (SGML) to MEDLARS II format as an alternative to keyboarding the data. SGML is part of the International Standards Organization (ISO) initiative.

Systems Branch

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

During FY 1993 hardware and system software were provided and enhanced to support the NLM's services. Mainframe computer reliability and availability were improved by adding components, enhancing procedures and implementing automated operations. The following is a summary of accomplishments.

Storage was expanded on the IBM 3090-300J from 128/256 million characters to 256/512 million characters, thus improving overall system performance by reducing paging and physical input/output operations. The bibliographic retrieval system in particular benefited from this.

Eight IBM 3480 magnetic tape drives were added to the computer system. They have reduced processing time for jobs that back up direct access storage.

Support to contract indexers was increased by adding Renex Protocol Converters with 62 additional ports.

Direct access storage capacity was increased by 22 billion characters. This storage is RAID (Redundant Arrays of Inexpensive Desk) technology and provides the capability to increase bibliographic retrieval system file sizes.

System software tools and enhancements were implemented to improve computer availability and reliability, simplify logon procedures, improve security, and provide additional performance and capacity measurements.

Support was provided for more than 100 software products used by programmers, users and system support staff.

Office automation support for personal computers and the PROFS calendaring and messaging system was provided. More than 1,700 calls for service were satisfied.

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. Staff support is provided around the clock, six days a week. 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 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 12 million pages or 600 million lines printed locally as well as on remote printers. Fan fold, cut sheet laser printers, and high-speed impact printers are attached to the mainframe system. Much of this printout is for subscribers who pay for it under terms of their agreements with NLM.

During the past fiscal year the Computer Services Branch created and mailed out 7,300 magnetic tapes of MEDLARS and TOXNET database information and files to both domestic and international subscribers.

In addition, the Computer Services Branch maintains an IBM 9370 in support of the Library's PROFS calendaring and message system.

INTERNATIONAL PROGRAMS

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

In FY 1993, NLM marked another year of active 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 medical information services in Egypt through a National Library of Medicine. 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.

A new project with the Department of State is to provide scientists in the Former Soviet Union with access to American colleagues and computerized scientific information. The goal of this project is to improve the state of medical information in Belarus, Ukraine, and Kazakhstan (among other countries) through major medical libraries. A prerequisite to attaining this is the implementation of e-mail for communication by biomedical scientists. With e-mail, it is possible to search MEDLARS databases and exchange correspondence with American colleagues.

International MEDLARS Agreements

The Library has MEDLARS agreements with partners in 17 foreign countries and with two international organizations (Table 11).

The Berman National Medical Library at the Hebrew University in Israel is the site of the newest International MEDLARS Center that is now providing MEDLARS services to health professionals in that country. On April 16, 1993, a formal agreement establishing a MEDLARS Center in Israel was signed by the Director of NLM and the Dean of the Medical Faculty at the Ministry of Health in Jerusalem, Israel. The Center is experimenting with the use of Grateful Med (GM) through the Internet for access from Israel.

The Australian MEDLARS Center has changed from being a tape leasing center is now online center.

Three other countries are experimenting with the use of Grateful Med through the Internet for access. They are: Jordan, Spain, and Hong Kong. Jordan sent three professional staff to NLM to participate in a week-long search training course.

NLM has a MEDLARS agreement with the Pan American Health Organization (PAHO), an 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 collaborated with PAHO and the University of Chile to improve a gateway system named BITNIS which allowed health professionals to conduct MEDLINE searches from Argentina, Chile, Costa Rica, Mexico, and Venezuela. In 1992, the BITNIS gateway software was ported to a Sun Spartan Workstation with many new features. A Beta test was conducted from June to October 1992 by 12 institutions. Today there are more than 150 codes with a daily average of 25 searches being logged onto the system.

To execute BITNIS, a MEDLINE search is initiated using Grateful Med. The search commands created by Grateful Med are transmitted to NLM through the Internet network. The retrieval obtained from the NLM computer are transmitted back to the originator through Internet; and Grateful Med is used again to edit and present search results. The objective of the BITNIS project is to provide NLM MEDLINE to health professionals in all Latin American countries where the high cost of international communication services inhibits access to the NLM databases. The first BITNIS Regional Workshop was held in Santiago, Chile on May 3-5, 1993 with over 60 participants from ten countries.

NLM accepted an International Associate into its 1993-94 Associate Program. A medical librarian with computer experience from the Lithuanian National Library of Medicine in Vilnius, Lithuania, was selected from a group of many qualified applicants.

Table 11

International MEDLARS Centers

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

*Combined online/tapes

Collaboration with the World Health Organization

The National Library of Medicine 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 continued to expand their collaboration. In 1993, WHO became an Internet host. A copy of Grateful Med is loaded onto a PC in the WHO medical library for online access through the Internet. NLM and WHO also continued their 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.

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 1993, two projects were in book production at the publisher in India: Bernhard Naunyn's *Memories, Thoughts and Convictions*, edited by Dr. David L. Cowen; and Jacques Tenon's *Memoirs on Paris Hospitals*, edited by Dr. Dora B. Weiner. Naunyn's book was originally published in Germany in 1925, and is considered a major medical autobiography. His memoirs centered on the development of scientific medicine with tangential accounts of personal life, customs, current events and society of the times. Tenon's book, originally published in 1788 in France, is a classic text of the French Enlightenment, providing the model for the public hospital.

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, commercial and governmental, from a number of countries. Common interests include the economics of primary and secondary publications, transborder flow of information, electronic publication, standardization and the information needs of developing countries. At the 1993

Council Meeting held May 16-19, 1993 in Williamsburg, Virginia, NLM was responsible for providing a technical symposium entitled "High Performance Computing and Communications In A Biotechnology Age." The session featured the Director of NLM as the keynote speaker. The Deputy Director of the Library continued his term as ICSTI President.

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. Among the visitors in 1993 were: the Humphrey Fellows (USAID) in Public Health, local and state health officers from China, and pharmacy students from Korea. Visitors came from the following countries:

Argentina, Australia, Bulgaria, Chile, China, Colombia, Croatia, Czechoslovakia, Egypt, France, Germany, Greece, India, Japan, Jordan, Kenya, Korea, Kuwait, Lithuania, Mexico, Nigeria, Norway, Pakistan, Romania, Russia, Singapore, Spain, Switzerland, Thailand, Ukraine, the United Kingdom, and Zimbabwe.

ADMINISTRATION

Kenneth G. Carney
Executive Officer

Financial Resources

In FY 1993, the Library had a total appropriation of $103,613,000. Table 12 displays the FY 1993 budget authority plus reimbursements from other agencies, and the allocation of these resources by program activity.

Table 12
Financial Resources and Allocations, FY 1993
(In Thousands of Dollars)

Budget Authority:	
Appropriation, NLM	$103,613
Plus: Reimbursements	11,984
Total	**$115,597**
Budget Allocation:	
Extramural Programs	26,459
Intramural Programs	81,042
Library Operations	(49,824)
Lister Hill National Center Biomedical Communications	(14,286)
National Center for Biotechnology Information	(10,285)
Toxicology Information	(6,647)
Research Management and Support	8,096
Total	**$115,597**

Personnel

The NLM closed FY 1993 with 596 full time equivalents (FTE). In anticipation of a reduction in the FTE ceiling for FY 1994, the Library began operating under a modified employment freeze during the final weeks of FY 1993. Table 13 displays FTE by utilization by program.

The NLM Director announced the creation of a new office within the Lister Hill National Center for Biomedical Communications (LHNCBC): the Office of High Performance Computing and Communications (OHPCC). The office is responsible primarily for coordinating high performance computing and communications planning, research, and development activities with federal, industrial, academic, and commercial organizations.

The NLM staff were saddened by the death on August 4 of Mrs. Patricia Southcomb, the NLM's Personnel Officer since 1989. Prior to coming to the Library, Pat worked as a Personnel Management Specialist in the Office of the Director, NIH.

Appointments

Ms. Betsy Humphreys, in addition to her role as Deputy Associate Director for Library Operations, was appointed NLM Assistant Director for Health Services Research Information. In this additional capacity, Ms. Humphreys will be responsible primarily for implementing the provisions of the "NIH Revitalization Act" pertaining to the establishment of a National Information Center on Health Services Research and Health Care Technology.

Lawrence C. Kingsland III, Ph.D., Kingsland, in addition to his responsibilities as Chief, Computer Sciences Branch, Lister Hill National Center for Biomedical Communications, was appointed NLM Assistant Director for Applied Informatics. His appointment responds to an essential need to establish a focal point within NLM for coordinating the varied activities in medical informatics research and the appropriate implementation of these advanced techniques in an operating environment.

Sally E. Howe, Ph.D., joined the NLM as a mathematician with the new High Performance Computing and Communications Office this year. Prior to her appointment at the Library, Dr. Howe was with the National Institute of Standards and Technology where she managed research and technical studies in advanced architectures, scientific visualization, and advanced data communications.

Roman Tatusov, Ph.D., was appointed as a Visiting Fellow with the National Center for Biotechnology Information. Dr. Tatusov is a graduate of Moscow State University where he specialized in mathematics, physics, and mathematical methods in biology. He is expert in computer methods, algorithms, and advanced programming techniques for analysis of sequences and structures of biological molecules.

The National Center for Biotechnology Information has recruited three Postdoctoral Fellows and two Predoctoral Fellows through the Intramural Research Training Award (IRTA) Program. The IRTA program is designed to provide an opportunity for training and practical research experience in NIH laboratories to U.S. postdoctoral students, recently graduated postdoctoral scientists, and predoctoral biomedical students.

Awards

The NLM Board of Regents Award for Scholarship or Technical Achievement was awarded to Joseph Hutchins, Office of Computer and Communications Systems, for technical achievement in developing the design concept and software for the NLM Locator which provides convenient and easy access to multiple, heterogeneous NLM information resources.

The PHS Special Recognition Award was presented to Melvin L. Spann, Ph.D., Specialized Information Services,

for his outstanding leadership in developing and enhancing toxicological information services at the Library, and ensuring their wider availability to underserved populations.

The PHS Meritorious Service Medal was awarded to David J. Lipman, M.D., National Center for Biotechnology Information, for his exceptional leadership and creativity at the National Center for Biotechnology Information in conducting computational biology research and providing information services to the biological community.

The PHS Outstanding Service Medal was awarded to Daniel R. Masys, M.D., Lister Hill National Center for Biomedical Communications, for his continuing effective leadership of a multi-technology based research and development program that gives realistic expectation of achieving the "virtual library."

The NIH Director's Award was presented to Dennis E. Black, Ph.D., Office of Administration, in recognition of his exceptional contributions to the NIH and the NLM acquisition activities, and the professionalism he brings to his career field.

The NLM Director's Honor Award, presented in recognition of exceptional contributions to the NLM mission, was awarded to two employees this year. Lawrence C. Kingsland, III, Ph.D., Lister Hill National Center for Biomedical Communications, and Patricia Carson, Office of the Director. Dr. Kingsland was recognized for sustained excellence in developing expert systems applications for health care and information retrieval. Ms. Carson was recognized for her foresight, commitment, and achievement in the advancement of the High Performance Computing and Communications programs.

* * * * *

Table 13
Staff, FY 1993 Full-Time Equivalents

Program	Full-Time Permanent	Other
Office of the Director	21	0
Office of Public Information	5	1
Office of Administration	50	6
Office of Computer and Communications Systems	59	9
Extramural Programs	14	4
Lister Hill National Center for Biomedical Communications	71	8
National Center for Biotechnology Information	19	14
Specialized Information Services	34	3
Library Operations	256	22
TOTAL	529	67
TOTAL FTEs	596	

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 MELARS 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 Re-

	sources 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 Academic 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
TOXNET	TOXicology Data NETwork
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

Abremski K, Sirotkin K, Lapedes A. Application of neural networks and information theory to the identification of E. coli transcriptional promoters. In: Avula XJR, ed. Mathematical Modelling and Scientific Computing, Vol 2. St. Louis: Principia Scientia, 1993;636-41.

Ackerman MJ. Multimedia medical education projects at the National Library of Medicine. In: Broering NC, ed. High-performance medical libraries: advances in information management for the virtual era. Westport CT: Meckler, 1993;177-84.

Ackerman MJ. Searching on-line data bases. Radiographics 1993 Jul;13(4):939-41.

Altschul SF. A protein alignment scoring system sensitive at all evolutionary distances. J Mol Evol 1993;36:290-300.

Baxevanis AD, Vinson CR. Interaction of coiled coils in transcription factors: where is the specificity? Curr Opin Genet Dev 1993;3:278-85.

Behr C, Sarthou JL, Rogier C, Trape JF, Dat MH, Michel JC, Aribot G, Dieye A, Claverie JM, Druihle P, Dubois P. Antibodies and reactive T cells against the malaria heat-shock protein Pf72/Hsp70-1 and derived peptides in individuals continuously exposed to Plasmodium falciparum. J Immunol 1992;149(10):3321-30.

Benson D, Lipman DJ, Ostell J. GenBank. Nucleic Acids Res 1993;21(13):2963-5.

Berman LE, Long R, Pillemer SR. Effects of quantization table manipulation on JPEG compression of cervical radiographs. Society for Information Display International Symposium Digest of Technical Papers 1993;24(43):937-941.

Berman LE, Nouri B, Roy G, Neve L. Interactive selection of JPEG quantization tables for digital x-ray image compression. IS&T/SPIE conference on imaging, San Jose, CA, Feb. 1-4,1993.

Blaisdell BE, Rudd KE, Matin A, Karlin S. Significant dispersed recurrent DNA sequences in the Escherichia coli genome: several new groups. J Mol Biol 1993;229:833-48.

Boguski MS, Lowe TMJ, Tolstoshev CM. dbEST—database for "expressed sequence tags". Nature Genet 1993;4:332-3.

Bork P, Koonin EV. An expanding family of helicases within the 'DEAD/H' superfamily. Nucleic Acids Res 1993;21:751-2.

Bouffard G, Ostell J, Rudd KE. GeneScape: a relational database of Escherichia coli genomic map data for Macintosh computers. Comput Appl Biosci 1992;8(6):563-7.

Bryant SH, Lawrence C. An empirical energy function for threading protein sequence through folding motif. Proteins Struct Func Genet 1993;16:92-112.

Claverie JM, Hardelin JP, Legouis R, Levilliers J, Bougueleret L, Mattei MG, Petit C. Characterization and chromosomal assignment of a human cDNA encoding a protein related to the murine 102-kDa cadherin-associated protein. Genomics 1993;15(1):13-20.

Claverie JM, States D. Information enhancement methods for large scale sequence analysis. Comput Chem 1993;17:191-201.

Claverie JM. A database of ancient protein sequences [letter]. Nature 1993;364:19-20.

Claverie JM. L'ordinateur a la chasse au genes. Sci Vie 1992;181:104-13.

Claverie JM. Sequence signals: artifact or reality? Comput Chem 1992;16:89.

Farber R, Lapedes A, Sirotkin K. Determination of eukaryotic protein coding regions using neural networks and information theory. J Mol Biol 1992;226:471-9.

Fetrow JS, Bryant SH. New programs for protein tertiary structure prediction. Biotechnology 1993;11:479-83.

Gautheret D, Major F, Cedergren R. Modeling the three-dimensional structure of RNA using discrete nucleotide conformational sets. J Mol Biol 1993;229:1049-64.

Ghosh D. Status of the transcription factors database (TFD). Nucleic Acids Res 1993;21(3):3117-8.

Gish W, States DJ. Identification of protein coding regions by database similarity search. Nature Genet 1993;3:266-72.

Godeny EK, Chen L, Kumar SN, Methren SL, Koonin EV, Brinton MA. Complete genomic sequence and phylogenetic analysis of the lactate dehydrogenase-elevating virus (LDV). Virology 1993;194:585-96.

Gorbalenya AE, Koonin EV. Comparative analysis of the amino acid sequences of the key enzymes of the replication and expression of positive-strand RNA viruses. Validity of the approach and functional and evolutionary implications. Sov Sci Rev D Physicochem Biol 1993;11:1-84.

Gorbalenya AE, Koonin EV. Helicases: amino acid sequence comparisons and structure-function relationships. Curr Opin Struct Biol 1993;3:419-29.

Green P, Lipman D, Hillier L, Waterston R, States D, Claverie JM. Ancient conserved regions in new gene sequences and the protein databases. Science 1993;259:1711-6.

Gregory PE, Gutmann DH, A. M, Park S, Boguski M, Jacks T, Wood DL, Jove R, Collins FS. Neurofibromatosis type 1 gene product (neurofibromin) associates with microtubules. Somat Cell Mol Genet 1993;19:265-74.

Gutmann DH, Boguski M, Marchuk D, Wigler M, Collins FS, Ballester R. Analysis of the neurofibromatosis type 1 (NF1) GAP-related domain by site-directed mutagenesis. Oncogene 1993;8:761-9.

Hauser SE, Cookson TJ, Thoma GR. Using back error propagation networks for automatic document image classification. SPIE Proc 1993;1965:142-50.

Hauser SE, Hsu W, Thoma GR. Request routing with a back error propagation network. SPIE Proc 1993;1965:689-95.

Humphrey SM. Evolution toward knowledge-based indexing for information retrieval. In: Kwasnik BH, Fidel R, eds. Advances in classification research, Proceedings of the 2nd ASIS Sig/CR Classification Research Workshop, Washington, DC, October 27, 1991. Medford NJ: Learned Information, 1992;53-65.

Humphrey SM. Indexing biomedical documents: from thesaural to knowledge-based retrieval systems. Artif Intell Med 1992 Oct;4(5):343-71.

Humphrey SM. The MedIndEx prototype for computer assisted MEDLINE database indexing. In: Mulvany NC, ed. Indexing, providing access to information: looking back, looking ahead, American Society of Indexers 25th Annual Meeting, Proceedings, Alexandria, VA, May 20-22,1993. Port Arkansas, TX: American Society of Indexers, 1993;45-54.

Humphrey SM. Use and management of classification systems for knowledge-based indexing. In: Williamson NJ, Hudon M, eds. Classification research for knowledge representation and organization, Proceedings of the 5th International Study Conference on Classification Research (ISCCR '91), Toronto, June 24-28, 1991. New York: Elsevier, 1992;89-100.

Humphreys BL, Lindberg DA. The UMLS project: making the conceptual connection between users and the information they need. Bull Med Libr Assoc 1993 Apr;81(2):170-7.

Humphreys BL, Schuyler PL. The Unified Medical Language System: moving beyond the vocabulary of bibliographic retrieval. In Broering NC (ed): High-performance medical libraries: advances in information management for the virtual era. Westport CT: Meckler, 1993;1-44.

Iris FJ, Bougueleret L, Prieur S, Caterina D, Primas G, Perrot V, Jurka J, Rodriguez-Tome P, Claverie JM, Dausset J, Cohen D. Dense Alu clustering and a potential new member of the NF kappa B family within a 90 kilobase HLA class III segment. Nature Genet 1993;3:137-45.

Jones KM, McPherson MJ, Baron MJ, Mattaj IW, Riordan CL, Wootton JC. The gdhA1 point mutation in Escherichia coli K12 CLR207 alters a key lysine residue in glutamate dehydrogenase. Mol Gen Genet 1993;240:286-89.

Karlin S, Altschul SF. Applications and statistics for multiple high-scoring segments in molecular sequences. Proc Natl Acad Sci USA 1993;90:5873-7.

Keister L. AIDS posters in the National Library of Medicine. In: 1993 Medical and health annual. Chicago: Encyclopedia Britannica, 1992; 124-5.

Kingsland LC 3d, Harbourt AM, Syed EJ, Schuyler PL. Coach: applying UMLS knowledge sources in an expert searcher environment. Bull Med Libr Assoc 1993 Apr;81(2):178-83.

Koonin EV, Deutscher MP. RNase T shares conserved sequence motifs with DNA proofreading exonucleases. Nucleic Acid Res 1993;21:2521-2.

Koonin EV, Dolja VV. Evolution and taxonomy of positive-strand RNA viruses: implications of comparative analysis of amino acid sequences. Crit Rev Biochem Mol Biol 1993;28:375-430.

Koonin EV, Gorbalenya AE. The superfamily of UvrA-related ATPases includes three more subunits of putative ATP-dependent nucleases. Protein Seq Data Anal 1992;5:43-5.

Koonin EV, Ilyina TV. Computer-assisted analysis of the relationships between protein and DNA sequences involved in rolling circle DNA replication. Biosystems 1993;30:241-68.

Koonin EV, Ilyina TV. Geminivirus replication proteins are related to prokaryotic plasmid rolling circle DNA replication initiator proteins. J Gen Virol 1992;73:2763-6.

Koonin EV, Senkevich TG, Chernos VI. Gene A32 product of vaccinia virus may be an ATPase involved in viral DNA packaging as indicated by sequence comparisons with other putative viral ATPases. Virus Genes 1992;7:89-94.

Koonin EV, Senkevich TG. Fowlpox virus encodes a protein related to human deoxycytidine kinase: further evidence for independent acquisition of genes for enzymes of nucleotide metabolism by different viruses. Virus Genes 1993;7:289-95.

Koonin EV, Woods DF, Bryant PJ. dlg-R proteins: modified guanylate kinases [letter]. Nature Genet 1992;2:256-7.

Koonin EV. A common set of conserved motifs in a vast variety of putative nucleic acid-dependent ATPases including MCM proteins involved in the initiation of eukaryotic DNA replication. Nucleic Acids Res 1993;21:2541-7.

Koonin EV. A highly conserved sequence motif defining the family of MutT-related proteins from eubacteria, eukaryotes and viruses. Nucleic Acid Res 1993;21:4847.

Koonin EV. A new group of putative RNA helicases. Trends Biochem Sci 1992;17:495-7.

Koonin EV. A superfamily of ATPases with diverse functions containing either classical or deviant ATP-binding motif. J Mol Biol 1993;229:1165-74.

Koonin EV. Bacterial and bacteriophage protein phosphatases [letter]. Mol Microbiol 1993;10:785-6.

Koonin EV. Comparative analysis of biopolymer sequences: reflections on the validity of the methodology and underlying general principles. In: Gindikin S, ed. Mathematical Methods of Analysis of Biopolymer Se-

quences. Providence, RI: American Mathematical Society, 1992;63-74. (DIMACS Series in Discrete Mathematics and Theoretical Computer Science; vol 8).

Koonin EV. Computer-assisted identification of a putative methyltransferase domain in NS5 protein of flaviviruses and lambda 2 protein of reovirus. J Gen Virol 1993;74:733-40.

Koonin EV. Escherichia coli dinG gene encodes a putative DNA helicase related to a group of eukaryotic helicases including Rad3 protein. Nucleic Acids Res 1993;21(6):1497.

Koonin EV. Evolution of double-stranded RNA viruses: a case for polyphyletic origin from different groups of positive-stranded RNA viruses. Semin Virol 1992;3(5):327-39.

Koonin EV. Genome sequencing: a glimpse of the area at the brink of the virus era. Biosystems 1993;30:269-76.

Koonin EV. The second cholera toxin, Zot, and its plasmid-encoded and phage-encoded homologues constitute a group of putative ATPases with an altered purine NTP-binding motif. FEBS Lett 1992;312:3-6.

Lacroix E-M. The international role of the National Library of Medicine. Bibliotheca Medica Canadiana 1993;14(3):123-32.

Lai CC, Boguski M, Broek D, Powers S. Influence of guanine nucleotides on complex formation between Ras and CDC25 proteins. Mol Cell Biol 1993;13(3):1345-52.

Landsman D. A signature for the HMG-1 box DNA-binding proteins. Bioessays 1993;15(8):539-46.

Landsman D. No HMG-1 box signature [letter]. Nature 1993;363:590.

Lawrence CE, Altschul SF, Boguski MS, Liu JS, Neuwald AF, Wootton JC. Detecting subtle sequence signals: a Gibbs sampling strategy for multiple alignment. Science 1993;262:208-14.

Layne SP, Merges MJ, Spouge JL, Gelderblom HR, Nara PL. Extensive and fundamental investigations of how humoral immunity blocks HIV infection. Ninth International Conference on AIDS, Jun 6-11. Berlin: The Fourth STD World Congress, 1993:17.

Le DX, Thoma GR. Automated portrait/landscape mode detection on a Binary Image. Proceedings of SPIE on Visual Information Processing 1993 April;1961:251-262.

Le DX, Thoma GR. Document skew angle detection algorithm. Proceedings of SPIE on Visual Information Processing 1993 April;1961:202-12.

Lindberg DA. Information: an interconnected world. Acad Med 1993 Aug;68(8):606-7.

Lindberg DA. NLM online network logs user increase. US Med 1993 Jan;29(1-2):48-9.

Lindberg DA. The National Library of Medicine and its role. Bull Med Libr Assoc 1993 Jan;81(1):71-2.

Lindberg DAB, Humphreys BL, McCray AT. The Unified Medical Language System. Meth Inform Med 1993;32:281-91.

Lindberg DAB, Siegel ER, Rapp BA, Wallingford KT, Wilson SR. Use of MEDLINE by physicians for clinical problem solving. JAMA 1993 Jun 23/30;269(24):3124-9.

Lindberg DAB. Global information infrastructure. Int J Biomed Comput 1994 Jan;34(1-4):13-9.

Lindberg DAB. Opinion: Health care and HPCC. Sharing medical imagery requires massive power and speed. Washington Technol 1993 Mar 25:13.

Locatis C, Park O. Some uneasy inquiries into ID expert systems. Educ Tech Res and Dev 1992; 40(30):87-94.

Locatis C. An interactive multimedia technology primer. LHNCBC Monograph 93-2. Bethesda MD: National Library of Medicine, 1993;23.

Majlessi L, Roth C, Benaroch P, Claverie JM, Kourilsky P, Bordenave G. Identification of T cell peptides from C gamma 2ab involved in IgG2ab T cell- induced allotypic suppression. Relationship with the MHC haplotype of the suppressor T cell-donor Igha mice. J Immunol 1993;151(3):1263-72.

Major F, Gautheret D, Cedergren R. Reproducing the three-dimensional structure of a tRNA molecule from structural constraints. Proc Natl Acad Sci U S A 1993;90:9408-12.

Masys DR. The evaluation of the source selection elements of the prototype UMLS Information Sources Map. In: Frisse ME, ed. 16th Annual Symposium on Computer Applications of Medical Care: a conference of the American Medical Informatics Association; 1992 Nov 8-11; Baltimore. New York: McGraw Hill, 1993;295-8.

McArthur J, Bolles J, Locatis C, Henderson J, Burke D, Rizzolo. A survey of videodisc and interactive videodisc projects in North America; part II. J Audiov Media Med 1993;16:25-33.

McCray AT, Aronson AR, Browne AC, Rindflesch TC, Razi A, Srinivasan S. UMLS knowledge for biomedical language processing. Bull Med Libr Assoc 1993 Apr;81(2):184-94.

McCray AT. Intelligent information retrieval. In: Mars NJ, ed. Informatiewetenschap 1992: Proceedings of 2nd annual STINFON conference. Enschede: Stinfon, 1992;21-28.

McCray AT. Representing biomedical knowledge in the UMLS Semantic Network. In: Broering NC, ed. High-performance medical libraries: advances in information management for the virtual era. Westport CT: Meckler, 1993;45-55.

Mehnert B. Outreach: a continuing commitment of the National Library of Medicine. NFAIS Newsl 1993 Aug;35(8):85-6.

Mehnert RB. National Library of Medicine. Bowker Annual 1993; 38:145-8.

Mehnert RB. NLM director to head National

Office of High-Performance Computing and Communications. Bull Med Libr Assoc 1993 Jan;81(1):82.

Mushegian AR, Koonin EV. Cell-to-cell movement of plant viruses. Insights from amino acid sequence comparisons and from analogies with cellular transport systems. Arch Virol 1993;133:239-57.

Mushegian AR, Koonin EV. The proposed plant homologue of connexin is protein kinase-like protein. Plant Cell 1993;5:998-9.

Oppenheim AB, Rudd KE, Mendelson I, Teff D. Integration host factor binds to a unique class of complex repetitive extragenic DNA sequences in Escherichia coli. Mol Microbiol 1993;10(1):113-22.

Prabhu V, Claverie JM. Correlations in intronless DNA. Nature 1992;359:782.

Prabhu VV. Symmetry observations in long nucleotide sequences. Nucleic Acids Res 1993;12:2797-800.

Rudd KE. Maps, genes, sequences, and computers: an Escherichia coli case study. ASM News 1993;59(7):335-41.

Salamon P, Wootton JC, Konopka AK, Hansen LK. On the robustness of maximum entropy relationships for complexity distributions of nucleotide sequences. Comput Chem 1993;17:135-43.

Schuyler PL, Hole WT, Tuttle MS, Sherertz DD. The UMLS Metathesaurus: representing different views of biomedical concepts. Bull Med Libr Assoc 1993 Apr;81(2):217-22.

Shamma SA, Fleshman JW, Wiser PR, Versnel H. Organization of response areas in ferret primary auditory cortex. J Neurophysiol 1993;69(2):367-83.

Soehner CB, Wray ST, Richards DT. The landmark citation method: analysis of a citation pattern as a collection assessment method. Bull Med Libr Assoc 1992 Oct;80(4):361-6.

Sparks S, Kuenz M. Interactive instruction in nursing and other health sciences: review of evaluation instruments. Lister Hill Monograph 93-1. Bethesda MD: National Library of Medicine. 1993;334.

Sparks S, Skiba D. The role of telecommunications in healthcare reform [interview]. Nurs Educ Microworld 1993 Apr-May;7(4):26.

Sparks S. National Library of Medicine offers nursing resources: E.T.Net's AVLINE conference. Reflections 1993;19(1):7.

Sparks SM. E.T.Net for health science education. J Biocommun 1993;20(2):2-13.

Sparks SM. ETB interactive technology resources for nurse educators. Deans Notes 1992 Nov;14(2):1-3.

Sparks SM. Exploring electronic support groups. Am J Nurs 1992 Dec;93(6):62-5.

Sparks SM. National Library of Medicine offers nursing resources. Reflections 1993 Spring;19(1):7.

Takao M, Abramic M, Moos M, Otrin VR, Wootton JC, McLenigan M, Levine AS, Protic M. A 127 kDa component of a UV-damaged DNA-binding complex, which is defective in some xeroderma pigmentosum group E patients, is homologous to a slime mold protein. Nucleic Acids Res 1993;21(17):4111-8.

Thoma G.R. Automated document delivery. Q Bull Int Assoc Agric Inf Spec 1992;37(1-2):84-8.

Thoma GR, Walker FL. Essential functions in an electronic document delivery system. In: Broering NC, ed. High-performance medical libraries: advances in information management for the virtual era. Westport CT: Meckler, 1993;77-88.

Thoma GR. Document preservation by electronic imaging: status and directions. In: Beyer C, ed. Preservation research and development, Round table proceedings, September 1992; Washington, DC: Library of Congress, 1993;70-2.

Toledano MB, Ghosh D, Trinh F, Leonard WJ. N-terminal DNA-binding domains contribute to differential DNA binding specificities of NF-kappa B p50 and p65. Mol Cell Biol 1993;13:852-60.

Tugendreich S, Boguski MS, Seldin MS, Hieter P. Linking yeast genetics to mammalian genomes: Identification and mapping of the human homolog of CDC27 via the expressed sequence tag (EST) data base. Proc Natl Acad Sci U S A 1993;90:10031-5.

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

van Bemmel JH, McCray AT. Editorial: advances in an interdisciplinary science. Methods Inf Med 1992;31(4):235-46.

Vasta BM, Wexler PI. Computer database helps douse toxic flames. Fire Chief 1993 Jul;7:42-51.

Vasta BM. TOXNET—National Library of Medicine [profile]. In: Katz Ls, Orrick S, Honig R, eds. Environmental profiles: a global guide to projects and people. New York, NY: Garland, 1993;572-3.

Wainwright P, Sparks SM. E.T.Net—an electronic conferencing system. Inf Technol Nurs 1993;5(3):10-2.

Waldherr M, Ragnini A, Schweyen RJ, Boguski MS. MRS6—yeast homologue of the choroideraemia gene [letter]. Nature Genet 1993;3(3):193-4.

Weisberg M. Guidelines for designing effective and healthy learning environments for interactive technologies. Interpers Comput Technol Electron J 21st Century 1993 Apr;1(2). (Available from the IPCT-L, a listserve operated by Georgetown University, listserv@guvm.georgetown.edu.)

Weisberg M. Guidelines for designing effective and healthy learning environments for interactive technologies. In: Proceedings of the 34th international conference of the Association for the Development of Computer-Based Instructional Systems. 1992;133-152.

Wilbur WJ. A retrieval system based on automatic relevance weighting of search terms. In: Shaw D, ed. Proceedings of the 55th American Society of Information

Science Annual Meeting. Pittsburgh, PA: Learned Information, Inc., 1992:216-20.

Wilbur WJ. Retrieval testing in hypergeometric document models. J Am Soc Inf Sci 1993;44(6):340-51.

Woodsmall RM, Benson DA. Information resources at the National Center for Biotechnology Information. Bull Med Libr Assoc 1993;81(3):282-4.

Wootton JC, Federhen S. Statistics of local complexity in amino acid sequences and sequence databases. Comput Chem 1993;17:149-63.

Appendix 3: Extramural Programs-Supported Publications

Atallah MJ. A faster parallel algorithm for a matrix searching problem. Algorithmica. 1993;9:156-67. (R01 LM 5118).

Parker WC, George DG. The Protein Identification Resource (PIR). CODATA Bulletin. 1991;23:59-61.(P41 LM 5206).

Barker WC, George DG. Mewes H-W, Tsugita A. The PIR-international protein sequence database. Nucleic Acids Res. 1992;20 Suppl:2023-6. (P41 LM 5206).

Bayard DS, Milman MH. Schumitzky A. Design of dosage regimens:a multiple model stochastic control approach. Los Angeles, CA 90033:Lab. of Applied Pharmacokinetics, Univ. of Southern California; 1992;92-1. 14p. (R01 LM 5401).

Bengtson JR, Kaplan AJ. Peiper KS, Wildermann NM, Mark DB, Pryor DB, Phillips HR, Califf RM. Prognosis in cardiogenic shock after acute myocardial infarction in the interventional era. J Am Coll Cardiol. 1992 Dec;20(7):1482-8. (G08 LM 04613).

Berner ES. The problem with software reviews of decision support systems. MD Comput. 1993;10(1):8-12. (R01 LM 5125).

Birkmeyer JD, Goodnough LT, AuBuchon JP, Noordsu PG, Littenberg B. The cost-effectiveness of preoperative autologous blood donation for total hip and knee replacement. Transfusion. 1993;33(7):544-51. (T15 LM 7044).

Birney E, Kumar S, Krainer AR. A putative homolog of U2AF in S.cerevisiae. Nucleic Acids Res. 1992;20(17):4663. (R01 LM 4971).

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; Wash., DC. New York: McGraw-Hill;[1992]:460-4. (R01 LM 04925).

Brinkley JF. Structural informatics and its applications in medicine and biology. Acad Med. 1991 Oct;66(10):589-91. (R01 LM 04925).

Broering NC, Bagdoyan H, Fagan D, Hylton J, Guttmann R. Medical facts file:a self service database of reference information. In: Frisse ME, ed. Supporting collaboration. Sixteenth Annual Symposium on Computer Applications in Medical Care; a conference of the American Medical Informatics Association; 1992 Nov 8-11; Baltimore. New York: McGraw-Hill;[1993]:886-8. (G08 LM 4392).

Brutlag DL, Galper AR, Millis DH. Knowledge-based simulation of DNA metabolism: prediction of enzyme action. Comput Appl Biosci; 1991;7(1):9-19. (R01 LM 4957).

Buchanan BG, Moore J, Forsythe D, Banks G, Ohlsson S. Involving patients in health care: explanation in the clinical setting. In: Frisse ME, ed. Supporting collaboration. Sixteenth Annual Symposium on Computer Applications in Medical Care; a conference of the American Medical Informatics Association; 1992 Nov 8-11; Baltimore. New York: McGraw-Hill;[1993]:510-4. (R01 LM 5299).

Burrows S. The community-based database prototype of the Southeast Florida AIDS Information Network. Bull Med Libr Assoc. 1993 Apr;81(2):223-6. (G08 LM 5168).

Chao KM, Hardison RC, Miller W. Constrained sequence alignment. Bull Math Biol. 1993;55(3):503-24. (R01 LM 5110).

Charpiat B, Breant V, Dumarest C, Maire P, Jelliffe R. Prediction of future serum concentrations with Bayesian fitted pharmacokinetic models:results with data collected by nurses versus trained pharmacy residents. Los Angeles, CA 90033:Univ. of Southern California School of Medicine;1993;93-1. 29p. (R01 LM 5401).

Conley DM, Kastella KG, Sundsten JW, Rauschning W, Rosse C. Computer-generated three-dimensional reconstruction of the mediastinum correlated with sectional and radiological anatomy. Clin Anat. 1992;5:185-202. (R01 LM 04925).

Crabo LG, Conley DM, Graney DO, Freeny PC. Venous anatomy of the pancreatic head:normal CT appearance in cadavers and patients. AJR Am J Roentgenol. 1993 May;160:1039-45. (R01 LM 4925).

Daniels N. Caring for the uninsured and underinsured; is the Oregon rationing plan fair? JAMA. 1991 May 1;265(17):2232-5. (R01 LM 5005).

Daniels N. Duty to treat or right to refuse? Hastings Center Report. 1991 Mar:36-46. (R01 LM 5005).

Daniels N. Growth hormone therapy for short stature: can we support the treatment/enhancement distinction? Growth & Growth Hormone. 1992 May;8 Suppl 1:46-8. (R01 LM 5005).

Daniels N. HIV-infected professionals: public threat or public sacrifice? Milbank Quarterly. 1992;70(1):3-42. (R01 LM 5005).

Daniels, N. HIV-infected professionals, patient rights, and the 'switching dilemma.' JAMA. 1992 Mar 11;267(10):1368-71. (R01 LM 5005).

Daniels N. Insurability and the HIV epidemic: ethical issues in underwriting. Milbank Quarterly. 1990;68(4):497-525. (R01 LM 5005).

Daniels N. Justice and health care rationing. In: Beehler R, Copp D, Szabados B., eds. On the track of reason: essays in honor of Kai Nielsen. Boulder, CO:

Westview Press;1992:57-74. (R01 LM 5005).

Daniels N. Rationing fairly: programmatic considerations. Bioethics. 1993;7(2/3):224-233. (R01 LM 5005).

de Bliek R, Martz JM, Reich GM, Friedman CP, Wildemuth BM. Domain knowledge and information retrieval in bacteriology: an information science perspective. Acad Med. 1992 Oct;67(10):S54-6. (R01 LM 4843).

DeCillis AP, Giuse DA, Thayer EH, Giuse NB, Miller RA. Using speech recognition in medical knowledge acquisition. In: Lun KC et al., eds. Proceedings of MEDINFO 92;1992 Sep; Geneva, Switzerland. North-Holland:Elsevier:1323-9. (T15 LM 07059).

DeCillis AP, Giuse DA, Thayer EH, Giuse NB, Miller RA. Using speech recognition in medical knowledge acquisition. In: Lun KC et al., eds. MEDINFO 92. North-Holland: Elsevier;1992:1323-9. (T15 LM 7059).

Elstein AS. Beyond multiple-choice questions and essays: the need for a new way to assess clinical competence. Acad Med. 1993 Apr;68(4):244-9. (R01 LM 4583).

Elstein AS. Paradigms for research on clinical reasoning: a researcher's commentary. Teaching and Learning in Medicine. 1992;4(3):147-9. (R01 LM 4583).

Elstein AS, Holzman GB, Belzer LJ, Ellis RD. Hormonal replacement therapy: analysis of clinical strategies used by residents. Med Decis Making. 1992;12:265-73. (R01 LM 4583).

Elstein AS, Kleinmuntz B, Rabinowitz M, McAuley R, Murakami J, Heckerling PS, Dod JM. Diagnostic reasoning of high- and low-domain-knowledge clinicians: a reanalysis. Med Decis Making. 1993;13:21-9. (R01 LM 4583).

Elstein AS, Kleinmuntz B, Rabinowitz M, McAuley R, Murakami J, Heckerling PS, Dod JM. Diagnostic reasoning of high- and low-domain-knowledge clinicians:a reanalysis. Med Decis Making. 1993 Jan;13(1):21-9. (R01 LM 4583).

Elstein AS, Rabinowitz M. Medical cognition: research and evaluation. In: Rabinowitz M, ed. Cognitive science foundations of instruction. Hillsdale, NJ:Erlbaum;1993:189-201. (R01 LM 4583).

Eno K, Sundsten JW, Brinley 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;Wash., DC. New York: McGraw-Hill; [1992]:727-31. (R01 LM 04925).

Fagan L, Wu A, Harris N, Poon A, Shiffman S. Spoken input for medical diagnostic systems. 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]:1059-60. (R01 LM 4864 and 5208).

Fitzpatrick K, Satz D, Hammond W, Grewal R, Hammond E, Collins J, Moylan JA. Automated acuity scoring within a computer based medical record. In: Frisse ME, ed. Supporting collaboration. 16th Annual Symposium on Computer Applications in Medical Care; a conference of the American Medical Informatics Association; 1992 Nov 8-11; Baltimore. New York: McGraw-Hill;[1993]:673-7. (G08 LM 4613).

Fleming JC, Pahl HL, Gonzalez DA, Smith TF, Tenen DG. Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution. J Immunol. 1993 Jan 15;150(2):480-90. (P41 LM 05205).

Forsythe DE. Using ethnography to build a working system: rethinking basic design assumptions. In: Frisse ME, ed. Supporting collaboration. Sixteenth Annual Symposium on Computer Applications in Medical Care; a conference of the American Medical Informatics Association; 1992 Nov 8-11; Baltimore. New York: McGraw-Hill;[1993]:505-9. (R01 LM 5299).

Freeman DT. Computer recognition of brain stem auditory evoked potential Wave V by a neural network. Ann Otol Rhinol Laryngol. 1992;101:782-90. (T15 LM 07059).

Galper AR, Brutlag DL, Millis DH. Knowledge-based simulation of DNA metabolism: prediction of action and envisionment of pathways. In: Hunter L, ed. Artificial Intelligence and Molecular Biology. Menlo Park, CA:AAAI Press;1993:365-95. (R01 LM 4957).

Giuse DA. The future of human-computer interaction in medical applications. J Med Educ Technologies. 1993 Dec:20-5. (R01 LM 04622).

Giuse NB, Giuse DA. Using the medical literature in multicenter knowledge acquisition for medical knowledge bases. J Med Educ Technologies. 1992 Jun:7-11. (R01 LM 4622).

Giuse NB, Giuse DA, Miller RA, Bankowitz RA, Janosky JE, Davidoff F, Hillner BE, Hripcsak G, Lincoln, MJ, Middleton B, Peden JG. Evaluating consensus among physicians in medical knowledge base construction. Methods Inf Med. 1993;32:137-45. (R01 LM 4622).

Guo D, Lincoln MJ, Haug PJ, Turner CW, Warner HR. Comparison of different information content models by using two strategies: development of the best information algorithm for Iliad. In: Frisse ME, ed. Supporting collaboration. 16th Annual Symposium on Computer Applications in Medical Care; 1992 Nov 8-11; Baltimore, MD. New York: McGraw-Hill;[c1993]:465-9. (R29 LM 05260).

Hardison R, Miller W. Use of long sequence alignments to study the evolution and regulation of mammalian globin gene clusters. Mol Biol Evol. 1993;10(1):73-102. (R01 LM 5110).

Hardison R, Xu J, Jackson J, Mansberger J, Selifonova O, Grotch B, Biesecker J, Petrykowska H,

Miller W. Comparative analysis of the locus control region of the rabbit beta-like globin gene cluster:HS3 increases transient expression of an embryonic eta-globin gene. Nucleic Acids Res. 1993;21(5):1265-72. (R01 LM 5110).

Hunt LT, Barker WC. After the sequence search, what next? CODATA Bulletin. 1991;23:19-28. (P41 LM 5206).

Hunter L, Searls D, Shavlik J., eds. ISMB-93; Proceedings First International Conference on Intelligent Systems for Molecular Biology; 1993 Jul 6-9;Bethesda, MD. Menlo Park, CA: AAAI Press; [1993]:459p. (R13 LM 5518).

Isaacs E, Wulfman CE, Rohn JA, Lane CD, Fagan. Graphical access to medical expert systems: IV. experiments to determine the role of spoken input. Methods Inf Med. 1993;32:18-32. (R01 LM 4864 and 5208).

Jacknowitz L. Rural health professionals plug into biomedical information resources. MUMPS Computing. 1992 Nov:14-8. (G08 LM 5047).

Jacknowitz L. West Virginia CONSULT: enhanced information access for health care practitioners in a rural environment. Parsons D, Fleischer CM, Greenes RA, eds. Ann N Y Acad Sci. 1992 Dec 17;670:163-70. (G08 LM 5047).

Jamieson PW. A computational model of levodopa pharmacodynamics in Parkinson's disease. Clin Neuropharmacol. 1991;14(6):498-513. (R29 LM 05074).

Jamieson PW. Neurology research:a journal survey, 1980-1990. Ann Neurol. 1992 Jul;32(1):87-93. (R29 LM 05074).

Jelliffe RW. Effect of assumptions concerning assay error patterns upon pharmacokinetic parameter values and model behavior. Los Angeles, CA 90033:Univ. of Southern California School of Medicine;1993;93-2. 9 pp. (R01 LM 5401).

Jelliffe RW, Sattler F, Gomis P, Tahani B. Adaptive control of drug dosage regimens: basic foundations, relevant issues, and clinical examples. Los Angeles, CA 90033: Lab. of Applied Pharmacokinetics, Univ. of Southern California; 1992;92-2. 20p. (R01 LM 5401).

Jelliffe RW, Schumitzky A, Van Guilder M, Liu M, Hu L, Maire P, Gomis P, Barbaut X, Tahani B. Individualizing drug dosage regimens:roles of population pharmacokinetic and dynamic models, Bayesian fitting, and adaptive control. Ther Drug Monit. 1993;15:380-93. (R01 LM 5401).

Jelliffe RW, Tahani B. A library of serum drug assay error patterns, and some suggestions for improved modeling and simulation of pharmacokinetic behavior. Los Angeles, CA 90033: Lab. of Applied Pharmacokinetics, Univ. of So. Calif.; 1992;92-5. (R01 LM 5401).

Johnson K, Poon A, Shiffman S, Lin R, Fagan LM. A history-taking system that uses continuous speech recognition. In:Frisse ME, ed. Supporting collaboration. 16th Annual Symposium on Computer Applications in Medical Care; a conference of the American Medical Informatics Association; 1992 Nov 8-11; Baltimore, MD. New York: McGraw-Hill;[1993]:757-61. (R01 LM 4864, 5208 and T15 LM 7033).

Jubran A, Gross N, Ramsdell J, Simonian R, Schuttenhelm K, Sax M, Kaniecki DJ, Arnold RJ, Sonnenberg FA. Comparative cost-effectiveness analysis of theophylline and ipratropium bromide in chronic obstructive pulmonary disease; a three-center study. Chest. 1993 Mar;103(3):678-84. (R01 LM 5266).

Kleinmuntz B. Computers as clinicians:an update. Comput Biol Med. 1992 Jul;22(4):227-37. (R01 LM 4583).

Knight JR. Discrete pattern matching over sequences and interval sets. Tucson, AZ: Univ. of Arizona;1993 Aug 31;TR 93-28. 90p. (R01 LM 4960).

Lin R, Lenert L, Middleton B, Shiffman S. A free-text processing system to capture physical findings: Canonical Phrase Identification System (CAPIS). 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]:168-72. (R01 LM 4864).

Lincoln MJ, Turner CW, Haig PJ, Williamson JW, Jessen S, Cundick RM, Cundick K, Warner HR. Iliad's role in the generalization of learning across a medical domain. In: Frisse ME, ed. Supporting collaboration. 16th Annual Symposium on Computer Applications in Medical Care; a conference of the American Medical Informatics Association; 1992 Nov 8-11; Baltimore, MD. New York: McGraw-Hill; [c1993]:174-8. (R29 LM 05260).

Lincoln MJ, Turner C, Bouhaddou O. Validating the clustered structures in knowledge bases. In: Frisse ME, ed. Supporting collaboration. 16th Annual Symposium on Computer Applications in Medical Care; a conference of the American Medical Informatics Association; 1992 Nov 8-11; Baltimore, MD. New York: McGraw-Hill;[c1993]:25. (R29 LM 05260).

Linster M, Musen MA. Use of KADS to create a conceptual model of the ONCOCIN task. Knowledge Acquisition. 1992;4:55-87. (R29 LM 05157).

Long KB, Fowler J, Barber S. VNS Retriever: querying MEDLINE over the Internet. In: Summer '92 USENIX;1992 Jun 8-12;San Antonio, TX. :81-91. (G08 LM 4905).

Lyon HC, Healy JC, Bell JR, O'Donnell JF, Shultz EK, Moore-West M, Wigton RS, Hirai F, Beck JR. PlanAlyzer, an interactive computer-assisted program to teach clinical problem solving in diagnosing anemia and coronary artery disease. Acad Med. 1992 Dec;67(12):821-8. (T15 LM 07044, R01 LM 04487, and K10 LM 00086).

McCarthy BD, Wong JB, Munoz A, Sonnenberg FA. Who should be screened for HIV infection? A cost-effectiveness analysis. Arch Intern Med. 1993 May

10;153:1107-16. (R01 LM 5266 and 4493).

McKibbon KA, Haynes RB, Johnston ME, Walker CJ. A study to enhance clinical end-user MEDLINE search skills: design and baseline findings. In: Frisse ME, ed. Supporting collaboration. Sixteenth Annual Symposium on Computer Applications in Medical Care; a conference of the American Medical Informatics Association; 1992 Nov 8-11; Baltimore. New York:McGraw-Hill;[1993]:73-7. (R01 LM 4696).

Mehldau G, Myers G. A system for pattern matching applications on biosequences. Comput Appl Biosci. 1993;9(3):299-314. (R01 LM 4960).

Miller RA. Editorial commentary: taking inventory of medical decision support software development. Methods Inf Med. 1993;32(1):9-11. (R01 LM 4622).

Miller W. Building multiple alignments from pairwise alignments. Comput Appl Biosci. 1993;9(2):169-76. (R01 LM 5110).

Muhlbaier LH, Pryor DB, Rankin JS, Smith LR, Mark DB, Jones RH, Glower DD, Harrell FE, Lee KL, Califf RM, Sabiston DC. Observational comparison of event-free survival with medical and surgical therapy in patients with coronary artery disease; 20 years of follow-up. Circulation. 1992 Nov;86(5 (Suppl II)):II/198-204. (T15 LM 7003, R01 LM 3373 and G08 LM 4613).

Musen MA. Dimensions of knowledge sharing and reuse. Comput Biomed Res. 1992;25:435-67. (R29 LM 05157 and P41 LM 05208).

Osheroff JA, Bankowitz RA. Physicians' use of computer software in answering clinical questions. Bull Med Libr Assoc. 1993 Jan;81(1):11-9. (T15 LM 07059).

Panko WB, Erhardt-Domino K, Pletcher T, Wilson W. Networking: an overview for leaders of academic medical centers. Acad Med. 1993 Jul;68(7):528-32. (G08 LM 5329).

Pao ML. Perusing the literature via citation links. Comput Biomed Res. 1993;26:143-56. (R01 LM 4177).

Pao ML. Term and citation retrieval: a field study. Information Processing & Management. 1993;29(1):95-112. (R01 LM 4177).

Pauker SG, Kopelman RI. Clinical problem-solving; a rewarding pursuit of certainty. N Engl J Med. 1993 Oct 7;329(15):1103-7. (R01 LM 4493).

Pauker SG, Kopelman RI. Hunting for the cause: how far to go. N Engl J Med. 1993 Jun 3;328(22):1621-4. (R01 LM 4493).

Perlin M. Call-graph caching for incremental modification and reuse of efficient artificial intelligence algorithms. In: Kambhampati S, ed. Working notes of AAAI Spring Symposium on Computational Considerations in Supporting Incremental Modification and Reuse; 1992 Mar; Palo Alto, CA.: Stanford Univ.; 1992:73-8. (R29 LM 04704).

Perlin M. Factored arc consistency for RETE match. Pittsburgh, PA 15213:Sch. of Computer Science, Carnegie Mellon Univ.; 1992 Apr; CMU-CS-92-127. 21p. (R29 LM 4707).

Perlin M. Inner product mapping: direct localization and ordering of DNA probes using chromosome-characterized radiation hybrids. Pittsburgh, PA 15213: Sch. of Computer Science, Carnegie Mellon Univ.; 1992 Jun;CMU-CS-92-162. 24p. (R29 LM 4707).

Perlin MW. Automating the construction of efficient artificial intelligence algorigthms. Pittsburgh, PA 15213: Sch. of Computer Science, Carnegie Mellon Univ.; 1991 Sep; CMU-CS-91-176. 245p. (R29 LM 04707).

Perlin MW. Transforming recursive processes into efficient algorithms:call-graph caching for artificial intelligence. Int J Artif Intell Tools. 1992;1(2):137-74. (R29 LM 04707).

Polaschek JX, Rutledge GW, Andersen SK, Fagan LM. Using belief networks to interpret qualitative data in the ICU. Repiratory Care. 1993 Jan;38(1):60-71. (R01 LM 04316, R01 LM 04136, T15 LM 07033, P41 LM 05208).

Poon A, Johnson K, Fagan LM. Augmented transition networks as a representation for knowledge-based history-taking systems. In: Frisse ME, ed. Supporting collaboration. 16th Annual Symposium on Computer Applications in Medical Care; a conference of the American Medical Informatics Association; 1992 Nov 8-11; Baltimore, MD. New York: McGraw-Hill; [1993]:762-6. (R01 LM 4864 and T15 LM 7033).

Poses RM, Wigton RS, Cebul RD, Centor RM, Collins M, Gleischli GJ. The importance of variability in patients' clinical characteristics and in physicians' responses to them. Med Decis Making. 1993 Oct;13(4):293-301. (R01 LM 4321).

Price DK, Ayres JA, Pasqualone D, Cabell CH, Miller W, Hardison RC. The 5'ends of LINE1 repeats in rabbit DNA define subfamilies and reveal a short sequence conserved between rabbits and humans. Genomics. 1992;14:320-31. (R01 LM 5110).

Pryor DB. Therapeutic choices in the patient with coronary artery disease and left ventricular dysfunction. Circulation. 1993 Jan;87(1):291-3. (G08 LM 4613).

Pryor DB, Shaw L, McCants CB, Lee KL, Mark DB, Harrell FE, Muhlbaier LH, Califf RM. Value of the history and physical in identifying patients at increased risk for coronary artery disease. Ann Intern Med. 1993 Jan 15;118(2):81-90. (G08 LM 4613).

Puerta AR, Egar JW, Samson WT, Musen MA. A multiple-method knowledge-acquisition shell for the automatic generation of knowledge-acquisition tools. Knowledge Acquisition. 1992;4:171-96. (R29 LM 05157 and P41 LM 05208).

Rabold JS. Achieving an integrated information system: the end-user factor. In: Kun KC et al., eds. MEDINFO 92; proceedings of the Seventh World Congress on Medical Informatics; 1992 Sep 6-10; Geneva. North-Holland: Elsevier;1992:1300-4. (G08 LM 4613).

Rich EC, Hillson SD, Dowd B, Lurie N. Residen-

cy reform and health care research. J Gen Intern Med. 1993 Mar;8:165-7. (T15 LM 7041).

Rich EC, Hillson SD, Dowd B, Morris N. Specialty differences in the 'July Phenomenon' for Twin Cities teaching hospitals. Med Care. 1993 Jan;31(1):73-83. (T15 LM 07041).

Schumitzky A. The nonparametric maximum likelihood approach to pharmacokinetic population analysis. Los Angeles, CA 90089: Mathematics Depart., USC; 1992;92-3. (R01 LM 05401).

Shiffman S, Lane CD, Johnson K, Fagan LM. The integration of a continuous-speech-recognition system with the QMR diagnostic program. In: Frisse ME, ed. Supporting collaboration. 16th Annual Symposium on Computer Applications in Medical Care; a conference of the American Medical Informatics Association; 1992 Nov 8-11; Baltimore, MD. New York: McGraw-Hill; [1993]:767-71. (R01 LM 4864 and 5208).

Shiffman S, Wu AW, Poon AD, Lane CD, Middleton B, Miller RA, Masarie FE, Cooper GF, Shortliffe EH, Fagan LM. Building a speech interface to a medical diagnostic system. IEEE Expert. 1991 Feb:41-50. (R01 LM 4864 and 5208).

Sittig DF. Work-sampling: a statistical approach to evaluation of the effect of computers on work patterns in healthcare. Methods Inf Med. 1993;32:167-74. (R29 LM 5284).

Sittig DF, Cheung KH, Berman L. Fuzzy classification of hemodynamic trends and artifacts: experiments with the heart rate. Int J Clin Monit Comput. 1992;9:251-7. (R29 LM 5284 and T15 LM 7056).

Spieler G, Schumitzky A. Asymptotic properties of extended least squares estimators with approximate models. Los Angeles, CA 90089: Dept. of Mathematis, Univ. of So. Calif.; 1992;92-4. 25p. (R01 LM 5401).

Springer GK, Patrick TB. A service-oriented information sources database for the biological sciences. In: Proceedings of the First International Conference on Intelligent Systems for Molecular Biology; 1993 Jul 7-9. Menlo Park, CA: AAAI Press; 1993:380-6. (R01 LM 5513 and T15 LM 7089).

Sundsten JW, Brinkley JF, Conley DM, Eno K, Kastella KG, Rosse C. Three-dimensional computer graphics reconstructions of anatomy on videodisc. J Audiovisual Media in Med. 1992;15:65-7. (R01 LM 04925).

Tcheng JE, Jackman JD, Nelson CL, Gardner LH, Smith LR, Rankin S, Califf RM, Stack RS. Outcome of patients sustaining acute ischemic mitral regurgitation during myocardial infarction. Ann Intern Med. 1992;117(18-24. (G08 LM 4613)).

Tong DA, Widman LE. Model-based interpretation of the ECG: a methodology for temporal and spatial reasoning. Comput Biomed Res. 1993;26:206-19. (R01 LM 5530).

Tu SW, Shahar Y, Dawes J, Winkles J, Puerta AR, Musen MA. A problem-solving model for episodic skeletal-plan refinement. Knowledge Acquisition. 1992;4:197-216. (R29 LM 5157 and P41 LM 5208).

Wong JB, Salem DN, Pauker SG. You're never too old. N Engl J Med. 1993 Apr;328(13):971-5. (R01 LM 04493).

Woodruff RC, Oster P. Mid-America Drosophila melanogaster Stock Center Stock List Version 12-91. Bowling Green, Ohio: Bowling Green State University; 1991.

Wulfman CE, Rua M, Lane CD, Shortliffe EH, Fagan LM. Graphical access to medical expert systems: V. integration with continuous-speech recognition. Methods Inf Med. 1993;32:33-46. (R01 LM 4864 and 5208).

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:

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

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

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

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

BRINKLEY, James F., M.D., Ph.D.
Research Assistant Professor
Department of Biological Structure
University of Washington

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

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

MITCHELL, Joyce A., Ph.D.
Director, Information Science Group
University of Missouri-Columbia

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

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

CANTOR, Charles R., Ph.D.
Director, Center for Advanced Research in Bioltechnology
Boston University
Boston, MA

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

FITZGERALD, Paula, M.D., Ph.D.
Senior Research Fellow
Department of Biophysical Chemistry
Merck Sharp & Dohme
Rahway, NJ

HUNKAPILLER, Michael W., Ph.D.
Executive Vice President
Applied Biosystems Division
Perkin-Elmer Corporation
Foster City, CA

KIM, Sung-Hou, Ph.D.
Director, Structural Biology Division
Lawrence Berkeley Laboratory
University of California
Berkeley, CA

WILLIAMS, Myra N., Ph.D.
Vice President, Information Technology
Glaxo Research Institute
Triangle Park, NC 27709

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:

FULLER, Sherrilynne, Ph.D. (Chair)
Director, Health Sciences Library and Information Center
University of Washington
Seattle, WA

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

BROERING, Naomi C.
Director, Biomedical Info. Resource Center
Dahlgren Memorial Library
Georgetown University Medical Center
Washington, D.C.

BUNTING, Alison
Associate University Librarian for Science
Louise Darling Biomedical Library
University of California
Los Angeles, CA

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

FIELDS, Christopher A., Ph.D.
Laboratory Director
Comparative Genomic Department
The Institute for Genomic Research
Gaithersburg, MD

HAYNES, R. Brian, M.D.
Chief, Health Information Research Unit
McMaster University
Hamilton, Ontario

IYENGAR, S. Sitharama, Ph.D.
Professor and Chairman of Computer Science
Louisiana State University
Baton Rouge, LA

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

MAVROVOUNIOTIS, Michael L., Ph.D.
Associate Professor
Chemical Engineering Department
Northwestern University
Evanston, IL

MUSEN, Mark A., M.D., Ph.D.
Assistant Professor
Depts. of Medicine and Computer Science
Stanford University
Stanford, CA

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

ROBERTS, Richard J., Ph.D.
Research Director
New England Biolabs
Beverly, MA

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

VANNIER, Michael W., M.D.
Professor of Radiology
Washington University
St. Louis, MO

VRIES, John K. M.D.
Assoc. Vice Chancellor for Medical Informatics
University of Pittsburgh Medical Center
Pittsburgh, PA

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:

WEAVER, William Lynn, M.D., F.A.C.S. (Chair)
Chief, Surgical Service
VA Medical Center
Buffalo, NY

ALTAMORE, Rita Ann, M.D., M.P.H.
Senior Lecturer
Department of Health Services
University of Washington
Seattle, WA

COLON, Angel Rafael Jr., M.D.
Professor of Pediatrics
Georgetown Univ. School of Medicine
Washington D.C.

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

HOAK, John C., M.D.
Director, Div. of Blood Diseases & Resources
National Heart, Lung, and Blood Institute
Bethesda, MD

PITTMAN, Constance Shen, M.D.
Professor of Medicine
University of Alabama
Birmingham, AL

THOMAS, John A., Ph.D.
Vice President for Academic Affairs
University of Texas Health Science Center
San Antonio, TX

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

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.

Dr. Billings was successful in acquiring funds to construct a new building to house the Library, and, in 1887, the collection was moved to this brick building on the Washington Mall. It was torn down in the late 1960's to make room for the Hirshhorn Museum of Art. The Library moved into its present building in Bethesda in 1962 (see back cover); a 10-story structure (the Lister Hill Center) was erected adjacent to it in 1980 to house the Library's computer facilities, audiovisual studios, and research and development laboratories.

The National Library of Medicine today is the world's largest research library in a single scientific field. Its holdings, which date back to the eleventh century, number more than five million books, journals, microfilms, pictures, audiovisuals, and other forms of recorded medical knowledge. An extensive computerized literature retrieval system, known as MEDLARS, provides rapid and convenient access to virtually all the materials in the Library's collection.

National Library of Medicine

8600 Rockville Pike
Bethesda, Maryland 20894