



**For more information about the National Library of Medicine, see:**

<http://www.nlm.nih.gov>



U.S. National Library of Medicine



# A Platform for Biomedical Discovery and Data-Powered Health

## National Library of Medicine Strategic Plan 2017–2027

---

Report of the NLM Board of Regents



U.S. National Library of Medicine

---

December 2017

U.S. Department of Health and Human Services  
National Institutes of Health

National Library of Medicine Cataloging-in-Publication Data

Names: National Library of Medicine (U.S.). Board of Regents, author.

Title: A platform for biomedical discovery and data-powered health : National Library of Medicine strategic plan 2017-2027 / report of the NLM Board of Regents.

Other titles: National Library of Medicine strategic plan 2017-2027.

Description: [Bethesda, Md.] : U.S. Dept. of Health and Human Services, Public Health Service, National Institutes of Health, National Library of Medicine, 2018. | Series: NIH publication | "December 2017." | Includes bibliographical references and index.

Subjects: MESH: Libraries, Medical. | Data Curation. | Medical Informatics Applications. | Biomedical Research. | Research Support as Topic. | Training Support.

Classification: NLM Z 675.M4





# From the Director



I am pleased to present to you the NLM Strategic Plan for the period 2017-2027. Over 750 NLM staff members, hundreds of national and international librarians, informatics professionals, biomedical scientists, data scientists, clinicians, public health specialists, and other stakeholders advised us as we charted a pathway towards our third century, and for this I am grateful. We consider the environment around us:

- **First and foremost is the promise of data-driven discovery.** Data-driven discovery requires sophisticated library and information science to open the door to thrilling new prospects for improving the public health, as well as informatics and data science to deliver insights and solutions.
- **Second, clinical care is rapidly migrating from hospital to home.** This migration challenges NLM to reach into these places where health occurs, not just where care is delivered, which are less formal and less well controlled.
- **Third, the spirit of open science suffuses the discovery landscape.** Governmental and scientific forces are aligning under a philosophy that innovation is accelerated if data flow freely, that the results of government-sponsored research should be open to the public as quickly as possible, and that linking scientists, citizens, and industry yields social benefits.
- **Fourth, the very nature of libraries is changing.** Libraries continue to be essential places for knowledge repositories and community gathering, yet the advent of self-directed search, e-publishing and consolidation of hospital library services challenges librarians and libraries to devise new services and solutions.

This strategic plan positions NLM to address existing and emerging challenges in biomedical research and public health. We will achieve this by creating a vibrant workforce; building on our core functions of acquiring, collecting, and disseminating the world's biomedical literature; and extending these skills and developing new ones to make data findable, accessible, interoperable, and reusable (the FAIR principles). We will continue to lead, conduct, and support research in biomedical information science, informatics, and data science to ensure that robust terminologies provide systematic characterization of complex health phenomena from cells to society and to devise new methodologies that uncover the knowledge held in data. We will expand our training programs to incorporate data science

and maintain our commitment to outreach excellence and support of a diverse workforce. And we will do this through our valued partnerships across biomedical research, the NIH, libraries, and the public.

The National Library of Medicine takes seriously the enormous public trust awarded us worldwide for the quality and integrity of our data and information. Our commitment is to preserve and protect this public trust while accelerating the yield from data-driven discovery. We invite you to be our partner in the adventure.

A handwritten signature in black ink that reads "Patricia Flatley Brennan". The signature is written in a cursive, flowing style.

Patricia Flatley Brennan, RN, PhD, Director

# NLM Today

Every day more than four million people use NLM resources; every hour, a petabyte of data moves in or out of our computing systems.

Our MedlinePlus Connect links electronic health records to trusted consumer health information. We foster rigor and reproducibility by linking results to the more than 250,000 clinical trials registered in our [ClinicalTrials.gov](https://clinicaltrials.gov) database, and we advance biomedical communication and capacity in sub-Saharan Africa.

Our Database of Genotypes and Phenotypes (dbGaP) and the ClinVar database of genetic variation, serves both scientists and the public health by helping identify the genetic components of disease and rapidly identify foodborne pathogens.

Our role as the coordinator of terminology standards within the Department of Health and Human Services enables interoperability among electronic health records.

NLM preserves over 10 centuries of precious medical manuscripts and other artifacts contributing thousands of items from its premier historical collections as a partner in the Medical Heritage Library to digitize treasures from the world's leading medical libraries.

We educate the next generation of innovators in biomedical informatics and data science, and our researchers develop new gene editing techniques, devise creative ways to locate and analyze medical images, and apply advanced analytics to decode clinical narratives and to map the spread of disease across communities.

We are the world's largest biomedical library and so much more.



# Executive Summary

The National Library of Medicine (NLM) envisions a future in which data and information transform and accelerate biomedical discovery and improve health and health care.

Achieving this vision will require NLM to refocus and enhance its research, development, training, and information services to make more biomedical data findable, accessible, interoperable, and reusable, to invent the tools and services to turn data and information into knowledge and insight, and to develop the workforce capable of doing so. It will require new forms of partnership and engagement with stakeholders in the public and private sectors, including researchers, librarians, health professionals, entrepreneurs and innovators, underserved communities, and the public.

This strategic plan outlines NLM's priorities for enabling a future of biomedical discovery and data-powered health. It heralds a new research paradigm in which data-driven science complements experimental and observational approaches, care is characterized by high levels of personalization, and personally collected data become a foundation of self-knowledge and personal health management.

To become a platform for biomedical discovery and data-powered health, NLM must collect and integrate an expanding set of information resources and enable them to be analyzed by tools emerging from the informatics and data science research front. It must establish pathways for dissemination and engagement across a broad range of users to drive information to the right place at the right time. It must expand training initiatives for informatics and data science research, and for technical support that will drive and shape the future of biomedical discovery. It must accelerate its investments in research in informatics and data science and align its research, development, and resource priorities toward achieving three goals, with an organizational structure that optimizes implementation .

## GOAL 1

### **Accelerate discovery and advance health through data-driven research**

- 1.1 Connect the resources of a digital research enterprise
- 1.2 Advance research and development in biomedical informatics and data science
- 1.3 Foster open science policies and practices
- 1.4 Create a sustainable institutional, physical, and computational infrastructure

## GOAL 2

### **Reach more people in more ways through enhanced dissemination and engagement**

- 2.1 Know NLM users and engage with persistence
- 2.2 Foster distinctiveness of NLM as a reliable, trustable-source of health information and biomedical data
- 2.3 Support research in biomedical and health information access methods and information dissemination strategies
- 2.4 Enhance information delivery

## GOAL 3

### **Build a workforce for data-driven research and health**

- 3.1 Expand and enhance research training for biomedical informatics and data science
- 3.2 Assure data science and open science proficiency
- 3.3 Increase workforce diversity
- 3.4 Engage the next generation and promote data literacy



# Introduction

The National Library of Medicine (NLM) is a global leader and trusted agent in the collection, organization and dissemination of biomedical information. Tracing its roots to the library of the U.S. Army Surgeon General in 1836, its statutory mission is “...to assist with the advancement of medical and related sciences and to aid in the dissemination and exchange of scientific and other information important to the progress of medicine and to the public health.”

NLM fulfills this mission by collecting, organizing, and providing access to the biomedical literature, as well as the growing volumes of molecular biology and clinical research data; engaging with users to discern and meet their information needs; advancing research and development in biomedical informatics and data science; and serving as the primary supporter of pre- and post-doctoral research training in biomedical informatics and data science in the United States.

As one of the 27 Institutes and Centers of the National Institutes of Health (NIH), NLM supports NIH priorities and NIH’s 2016-2020 Strategic Plan objectives addressing biomedical research opportunities, priority-setting for innovation, enhancing scientific stewardship, and managing for results. NLM will support the NIH mission of transforming discovery into health by providing a dynamic information platform that accelerates biomedical discovery and translation into clinical care, public health practices, and personal wellness. It will also work to achieve the recommendations of the NIH Advisory Committee to the Director that “NLM should lead efforts to support and catalyze open science” and “be the intellectual and programmatic epicenter for data science at NIH.”

NLM’s vision for the coming decade is to unleash the potential of data and information to accelerate and transform discovery and improve health and

health care. Central to this vision is the idea of data-powered health, in which holistic knowledge of the person is enriched with greater awareness of biological processes, environmental exposures, and human connection. With data-powered health, extant data complements experimental and observational approaches, care processes are characterized by high levels of personalization, and personal data collection becomes a foundation of self-knowledge and personal health management.

## Achieving this vision will require NLM to:

- 1** Create and link an expanding collection of integrated and interoperable information resources, powered by analytic tools from the informatics and data science research front. This will require increased investments in informatics and data science and an organizational structure that optimizes implementation.

---

- 2** Deepen and establish new forms of engagement across a broad range of users to drive information to the right place at the right time.

---

- 3** Expand training initiatives for informatics and data science research to build a workforce to drive the future of biomedicine.

# Goals and Objectives

## GOAL ONE:

### Accelerate discovery and advance health by providing the tools for data-driven research

---

Scientific investigations now yield not only amazing advances but also spawn large and potentially valuable data sets that can be exploited to achieve significant additional advances and even greater efficiencies to the progress of research. NLM will invest in novel approaches to bring its collections together—literature, images, biological data, environmental exposures, to name a few—and create a rich integrative platform for data-driven discovery and data-powered health.

NLM will create computable libraries of data, models, and literature and devise an interlocked environment that enables discovery; this effort will be guided by principled decisions in collections management and curation, and propelled by practices that foster open science and scholarship. It will also build upon research advances in biomedical informatics and data science. Partnerships with domain scientists, health professionals, information professionals, and the public will guide the design of discovery and analysis tools to improve the insight borne of data, information, and knowledge. Organizational infrastructure will also be modified to support new science and services.



## Objective 1.1: Connect the resources of a digital research enterprise

NLM will enhance its efforts to collect, organize, and disseminate well-understood (literature) and non-traditional research objects, notably, but not only, data. These collections of information, whether physical or digital, are fuel for ideas, knowledge, insight, and progress. The value of individual collections can be multiplied when collections of different, but related, objects can be searched, interrogated, and analyzed, and their insights integrated, aggregated, and shared.

NLM will lead this process by creating, collecting, and curating digital research objects such as articles, data sets, analytic strategies, and visualization tools, as well as executable objects such as workflows, analytic pipelines, simulations, and predictive models. Systems will be designed around the FAIR principles of making information Findable, Accessible, Interoperable, and Re-usable. Modern approaches to attribution will be devised.

NLM will modernize NLM collections and services, and enhance the integration and interoperability of existing collections. In so doing, it will ensure the responsiveness of collections to a broad range of information needs. NLM will ensure the operational and curatorial efficiency of collection management. This may include developing new methods of automated indexing and designing innovative approaches to presenting information that promotes interoperability among resources and, importantly, facilitates linking to resources not housed in NLM.

NLM will develop new collections to reflect scientific progress and meet new scientific needs. Collection development and retention decisions will be based on program priorities and informed via partnerships with stakeholders. While some new resources will be incorporated through acquisition by NLM, others will be integrated through federation, coordination, or other connections with the integrated platform. Recognizing the importance of global knowledge resources, NLM will form partnerships for

connecting with these resources across and outside NIH, and around the world.

NLM recognizes the potential for addressing the continuing challenge of health disparities through integrated and aggregated collections of diverse data and analytic tools. In creating a platform for discovery, NLM is committed to offering a resource that is useful for advancing basic and translational research on health disparities. Factors important for such research include a range of biological, genomic, social, behavioral, and environmental determinants of health, as well as population characteristics such as sex, gender, age, race, and ethnicity. In addition to incorporating relevant content in the information collections, NLM will promote the development

---

NLM will enhance its efforts to collect, organize, and disseminate well-understood (literature) and non-traditional research objects, notably, but not only, data.

---

and use of standards for data collection and reporting that will enable aggregation of data and identification of subpopulations for analysis. In unique ways like these, and more, NLM contributes to the Trans-NIH Minority Health and Health Disparities Strategic Plan.<sup>1</sup>

NLM will augment its literature and genomics collections to create a platform for science and health from new sources, such as:

**Data generated from scientific research:**

New investigations generate not only hypotheses resolution but also data as a significant scientific product. The number of data repositories for research datasets is increasing, along with the volume of data and the variety of data. NLM will lead in developing mechanisms to make data repositories discoverable



and set the policies to guide the selection of repositories and their investment. NLM will also develop approaches and infrastructure for accommodating new classes of research data, in addition to enabling access to externally-hosted resources. In addition, NLM will engage with the entire NIH to determine how to accommodate data generated in large NIH research initiatives such as the BRAIN Initiative, All of Us, and the Cancer Moonshot.

**Curation and standards:** NLM develops, houses, supports, and stewards many health data standards that promote interoperability. These sets of standards include: terminologies and vocabularies, data formats, value sets, common data elements, and coding systems related to clinical phenomena. There remain many other types of data for which standard representations must be developed, including biological, environmental, behavioral, and interactional. NLM will devise new methods of generating standards, maintaining version control, and applying them at the appropriate point in the data life cycle. An important research direction will develop strategies for curation at scale, and those that accelerate the standards development and application process. In partnership with community stakeholders, NLM will consider how to use its collections and their interconnectivity to facilitate application of standards, as well as how to help user communities increase the uniformity and the interoperability of standards.

**Data science tools and other executables:** NLM will play a role in creating new methodologies and new ways of organizing collections of data science tools. Examples include: libraries of mathematical models; indexed visualization tools; reusable data analytic models for health applications; open source software and algorithms; research workflows; statistical, diagnostic, or predictive models; or collections of machine-executable knowledge. These will not replace, but will be integrated with, resources that reside within NIH Institutes and Centers, their grantees' institutions, or other public and private

organizations. NLM will work with stakeholders to identify tools for dissemination and to develop policies to guide decisions for incorporating them within the NLM platform.

**Clinical data:** NLM must create the controls for effective stewardship of data generated in the course of clinical care as well as from clinical investigations. NLM will work with EHR vendors and clinical care institutions to foster linkage and reuse, as well as to house special collections. Aggregated collections of such data, properly curated, will enable analyses of subpopulations based on aspects of their medical care and on demographic characteristics such as gender, age, race, and ethnicity. NLM engagement arises from strong commitment to facilitate research in minority health and health disparities.

NLM will enrich its flagship clinical trial registry and repository, ClinicalTrials.gov, through improved front-end and back-end design. Enhancements to expand interoperability with data and information in other resources, such as individual-level data collected in the course of the trials supported by NIH, will be considered, along with improvements in operational efficiency in producing, curating, and ensuring long-term preservation of this resource.

**Status indicators of the health of people and communities:** Growing recognition of the importance of social, behavioral, and environmental aspects of health will lead NLM to convene stakeholders to develop data standards and processes for accessing information and accelerating standards for these indicators. Collections relevant to the community span microbial colonies to behaviors to disease outbreaks to broad based environmental indicators. NLM must anticipate future needs for data through effective partnerships and implement mechanisms for the data life-cycle and the value proposition at each state.

**Scientific communication:** NLM will stimulate new forms of scientific communication and become the library of the future, one of connections between and among literature,

data, models, and analytical tools. Creating efficient ways to link the literature with associated datasets enables knowledge generation and discovery. Initial approaches to connect datasets associated with papers in PubMed Central will be enriched with a complex environment interlocking data, literature reports, science profiles, and research proposals.

NLM will anticipate developments such as preprints and novel alternatives to the journal issue as a mode of organizing articles. This includes the concept of an “executable article,” in which human-readable text can be combined with computer-interpretable representations of the evidence and guidance contained in a publication. Decomposing articles into components including interactive visualizations, decision analyses, executable models and simulations (e.g., of physiologic or molecular processes), and predictive and

decision-analytic inference systems will enhance the value and extension of science.

**Health information for the public:** NLM provides portals to authoritative health information from NIH and other government agencies, as well as private sources. In the future, the goal is to provide that information in a way that makes it more actionable. In addition, there is a need to complement existing information by providing linkages to relevant data from multiple sources, such as environmental contaminant or exposure data and disease outbreak data, to enable an integrated presentation of information in understandable formats such as customizable dashboards. NLM will engage with commercial and non-commercial web search and health information providers to preserve its responsibility as an authoritative provider of health information.

## Objective 1.2: Advance research and development in biomedical informatics and data science

As a platform for discovery, NLM plays a significant role throughout the entire research lifecycle, from initial inspiration through scholarly communication. NLM must anticipate research directions of the biomedical and clinical health science, prepare the tools necessary to accelerate discovery, and enable efficiencies in communication and dissemination. NLM has the responsibility to foster problem-inspired research across the various health science disciplines while accelerating the development of domain-independent, reusable approaches to discovery, curation, analytics, and collections.

Biomedical scientists of today tell of a future where food itself becomes medicine, where understanding the context of health extends far beyond the intracellular environment to the built environment of homes and cities, where the data flows needed to characterize health phenomena

move at speed of light and are far too dense to be captured by standard databases. NLM, in close partnerships with the biomedical and clin-

---

NLM must anticipate research directions of the biomedical and clinical health science, prepare the tools necessary to accelerate discovery, and enable efficiencies in communication and dissemination.

---

ical research disciplines, will build the discovery and knowledge resources and tools needed while tailoring its investments to leverage the best of the disciplines.

First among these resources and tools will be the devising of a system of digital research objects – articles, datasets, computational pathways, analytical models, visualization tools, reference standards, and the like, each of which will have a unique identifier and could be coupled in novel modes to accelerate discovery in a broad array of disciplinary explorations. NLM will provide the accelerators and the repositories of future discoveries, and enable newer and more efficient dissemination mechanisms.

Connecting collections of digital objects so that the whole is greater than the sum of its parts presents big computational and scientific challenges. Surmounting these challenges will require data science and informatics research and development, including innovation in curation, analytics, visualization, modeling, knowledge generation and next-generation mining approaches. These innovations, in turn, depend on advances in the science and mathematics underpinning such approaches. This research agenda will require more investment in intramural and extramural research support and pursuing partners whose participation will augment these efforts.

Below are examples that illustrate how such research is integrally related to the mission and future of NLM and NIH.

- With the science and health sectors producing more digital research objects—whether research-generated datasets, electronic health records, or computational models—value can be added through curation. A key

area for research is how curation at scale might be accomplished for hundreds of millions of such objects in a way that would be workable across biomedical domains, health care sectors, public health, and consumer health. Automatic, autonomous curation strategies will allow for operational efficiency as well as accelerate the speed of discovery.

- Analytics beyond statistics, data presentation beyond visualization, and “data mining using meaning” across heterogeneous data sources are important areas for advancing data science. Research and development of artificial intelligence in its many forms, including natural language processing and deep learning, require stimulation, demonstration, testing, and curation.
- Computable biomedical knowledge—such as diagnostic, predictive, and decision-analytic models, or practice guidelines represented as coded digital objects—is an increasingly important complement to human-readable knowledge represented in books and journals.
- Define the structure and functions of “executable articles” that form an interactive library where the resources “talk to each other” and enable movement from data to knowledge and knowledge to action.
- A rich set of data resources presents the opportunity to pursue basic research on data-driven questions in biomedical domains. NLM will partner with scientists across NIH to generate research questions and identify value solutions.

## Objective 1.3: Foster open science policies and practices

NLM will become an advocate for open science, and thereby democratize access to the products and processes of scientific research—that is, making them widely accessible. NLM will continue to develop tools that support open science, including information and data systems

that preserve and make publications, data, and other research outputs discoverable and accessible to the broadest possible audiences. It will continue to develop technologies and systems that provide necessary protections and access limitations to such materials, where necessary,

---

NLM will become an advocate for open science, and thereby democratize access to the products and processes of scientific research—that is, making them widely accessible.

---

e.g., to protect privacy, confidentiality, security, and certain proprietary rights.

Beyond its systems and services, NLM has an

essential role to play in helping develop and implement policies and practices that facilitate open science and data-driven research, while addressing the complex challenges of investigator freedom and patient privacy. These include policies related to data sharing, privacy protection, and informed consent that can motivate open science practices consistent with legal and ethical considerations. With productive relationships across NIH and the Department of Health and Human Services, NLM is positioned to participate in the reformulations necessary to allow open science to flourish, and can bring to these policy debates the practical experience in system building that can ensure effective policy and implementation.

## Objective 1.4: Create a sustainable institutional, physical, and computational infrastructure

For NLM to become the hub for data science and open science activities at NIH, it must also make an organizational commitment to creating an institutional home, which will provide for organizational stability, fiscal accountability for data science efforts, and coordination of resources, investments, and communication. In devising the best structure to accomplish its data science mission, NLM must consider the key activities to be accomplished; how to best leverage existing data science efforts; and the possibility of a hybrid structure in which a permanent, clearly identified institutional locus for data science and open science exists in concert with the distributed nature of data science talent and investment.

NLM will make creative use of a constrained physical space while advocating for a larger and more appropriate physical environment. Sufficient space is needed for the current staff, which is distributed across several sites.

Additionally, fit-for-purpose space is needed to house and support its physical and digital collections to guarantee long-term access. As a national biomedical library, NLM takes seriously the responsibility to ensure the preservation of its unique collections of books and medical heritage embodied in information artifacts.

NLM's mission is dependent on the viability of its data centers and the computation platforms therein. It needs to ensure adequate capacity and reliability in the electrical, mechanical, and structural infrastructure that supports NLM data centers to meet future computational demands. It must experiment with innovative computer and network solution strategies. Technical and policy solutions must accompany curation and collection management activities. Key technical challenges include authentication and authorization, assuring compliance with permissions required for data access and re-use, and security of data sets and session permission.

# Goals and Objectives

## GOAL 2:

# Reach more people in more ways through enhanced dissemination and engagement pathways

---

NLM must understand and meet the needs of the stakeholders it serves: scholarly, clinical and community. A future of data-powered health will be built upon new models for interacting with information and data in innovative ways that can accelerate discovery and lead to knowledge that informs decisions and action. Advancing science and improving health is tied to potential users being aware of NLM resources, having access to them, and understanding their use. NLM continually strives to understand user needs and how they vary over time and across different categories of users, and to create new and better ways to deliver the right information, at the right level of sophistication, and at the right time.

In designing its dissemination initiatives, NLM is committed to developing approaches to innovation and outreach in which information access can reduce health disparities and support underserved communities that bear a disproportionate burden of disease.

## Objective 2.1: Know NLM users and engage with persistence

As new resources and services bring new users with new information needs, NLM will need to create pathways to reach them with relevant information at times the information is needed. Users will range from librarians to researchers, policy-makers to teens, clinicians to parents, pharma to public health laboratories. Each of these user groups will require unique patterns of engagement. User engagement encompasses promoting awareness of information resources, understanding information needs, facilitating access, and ensuring ability to use information resources.

NLM will develop strategies to increase its visibility among intended audiences globally and nationally. Modern human factors and human computer design strategies will guide public-facing utilities redesign that feature the National Library of Medicine and National Institutes of Health brand. Attention to ethnographic design principles that draw from user experience are needed. Partnerships with NIH Institutes will also be systemically cultivated and highlighted.

Recognizing that people live in communities, NLM will employ person-centric and community-conscious design strategies. This requires an investment in research to better understand health information needs and how to address them. It will prioritize increasing awareness of NLM services among low resource or low literacy communities and among underrepresented groups, in keeping with NLM's commitment to employ information interventions to reduce health disparities.

NLM will employ rapid cycle design and delivery approaches that reflect knowledge both of what the user seeks from its resources as well as why they are seeking it. It is also important

to understand user needs in the moment. Approaches to dissemination must reflect different motivations for seeking information as well as disparities in health and data literacy among NLM constituencies.

Human networks remain important and will be leveraged to expand outreach across the U.S. NLM will continue to utilize the 6,500 member libraries of the National Network of Libraries of Medicine (NNLM) to serve in a trusted, visible manner for engaging and reaching out to the

---

User engagement encompasses promoting awareness of information resources, understanding information needs, facilitating access, and ensuring ability to use information resources.

---

wide range of NLM users. With the NNLM, NLM will emphasize expanding partnerships with public libraries and community groups to improve awareness of NLM's resources and to obtain feedback for improvement. The NNLM is partnering with the NIH All of Us Research Program to support community engagement efforts by public libraries and raise awareness about the program for populations unrepresented in biomedical research. The partnership also aims to increase the capacity of public library staff to improve health literacy.



## Objective 2.2:

### Foster distinctiveness of NLM as a reliable, trustable source of health information and biomedical data

The information and data resources provided by NLM are becoming ever-more essential to biomedical discovery and contemporary health care while also becoming increasingly invisible as this information and data become embedded in the processes and products of our users. NLM is essential to public health authorities, who query its genomic data banks daily to discern the nature of a foodborne pathogen; essential to hospitals for demonstrating quality of care through using a MedlinePlus Connect embedded link to

---

NLM will aim to become recognized widely as the pre-eminent source of trusted information for the nation's health.

---

clinical systems to provide patients with tailored, trusted health instructions; essential to the over four million users who daily access its literature and biological databases. New devices found along the Internet-of-Things open an entirely new dissemination outlet for NLM resources. As NLM resources become embedded in search engines and informational instructions around the world, NLM must find ways to demonstrate its

oversight and promote the chain of trust in health information.

Those who recognize the NLM brand—libraries, research scientists, health professionals and the lay public—know and trust its information resources. Even though millions of people use NLM's information services every day, there remains a persistent lack of name recognition and underappreciation for the vast number and variety of services provided by the NLM. NLM must make itself accessible and known to those who do not yet recognize its brand nor know of its services. Public awareness campaigns, better branding of the NLM-provided resources, and educating scientists and the public in best practices for health information and health data management will extend the chain of trust and improve discovery and care through greater reliance on trusted health information.

NLM will aim to become recognized widely as the pre-eminent source of trusted information for the nation's health. It will increase public awareness and the public's trust in NLM-provided materials not only by devising ways to convey the NLM imprimatur on materials, but also through creative engagement that raises the general awareness of NLM as a place to turn to with health questions or biomedical data needs. Envision a campaign "NLM: your most trusted partner in health".

## Objective 2.3:

### Support research in biomedical and health information access methods and information dissemination strategies

NLM will support programs of research in areas such as rapid cycle information processing, and will undertake new research to bring new

information products to new users. It will focus on understanding how searches are initiated, how information is used, and how questions are

posed and answered. The Library will provide important advances in delivering guidance for action or decision making and provide insight through advanced visualizations of complex data. NLM will continue supporting and undertaking informatics and data science research to design advanced interfaces and query systems that facilitate access to information systems in new ways and reflect the information needs and behaviors of individuals. This will include investigating the best methods to deliver effective information to support team science,

family caregiving, and care coordination. Some examples of research areas that contribute to innovation in information access and delivery include human-computer interaction, natural language processing, virtual and augmented reality for system interaction, probabilistic inference mechanisms for sense-making from data visualization, and predictive analytics. To meet the needs of a diverse audience for information services, NLM must also conduct research to address matters of literacy, numeracy, and making information actionable in the moment.

## Objective 2.4: Enhance information delivery

NLM will provide access to information on multiple delivery platforms and through multiple user-centric interface formats, including ways that are very different from today's formats. Strategies will be developed to provide better interlinking of resources to improve discovery.

In addition to designing fresh approaches for querying databases on demand, NLM will consider implementing push models that will

addressing information delivery challenges may lead to innovations in information delivery as well as new platforms.

To promote ease of access to NLM resources, interfaces that are more uniform across NLM platforms will be developed. It will explore voice, gesture, and behavioral interfaces, while attempting to ensure a predictable experience. Use of a common interface will not only improve accessibility of resources, it will complement the enhancements in interoperability and findability of related resources. Through test bed experimentation with new ways to improve search quality and usability, NLM will offer a better user experience for finding information faster, with a clean modern interface, and with better support for mobile devices.

---

To promote ease of access to NLM resources, interfaces that are more uniform across NLM platforms will be developed. It will explore voice, gesture, and behavioral interfaces, while attempting to ensure a predictable experience.

---

Given the range of uses of NLM's resources, NLM will continue to offer multiple delivery platforms, including application programming interfaces (APIs), mobile, social media, web-based, and stand-alone applications that can operate without internet access. New presentation modes that do not require the user to read, but still allow user control, will be explored. Fundamental to these efforts is recognizing that the fastest growing group of NLM users are actually digital devices that engage with the systems automatically.

anticipate user needs, operate continuously to search for new information, and engage them with results in a just in time manner. New approaches to input and output formats will be explored from the research front, including voice, music, image, and virtual reality applications. Partnerships with other NIH Institutes that are

# Goals and Objectives

## GOAL 3:

### Build a workforce for data-driven research and health

---

To assure a future of data-driven discovery and health, it is essential to have a biomedical informatics and data science workforce prepared to make conceptual and methodological advances in analytics, visualization, mining, and other methods needed to use data for discoveries and to make it interoperable with existing knowledge. A new generation of librarians who are as capable of collecting and disseminating data resources as they are with more familiar forms of scholarly communications is also needed. The roles of the informationist and the librarian must expand to support a world where massive industrial search engines provide immediate access to vast stores of knowledge, while threatening to overwhelm the scholar, patient, and clinician with unfettered lists of potentially relevant resources.

NLM must also play a role in creating a broader biomedical sciences workforce capable of data-driven discovery. New approaches to investigation will exist side-by-side with familiar experimental and exploratory methodologies. Training must foster the development of scholars with expertise in both biomedical informatics or data science *and* expertise in a biomedical or clinical domain. Both are needed in order to develop and deploy usable and useful applications harnessing the power of data to advance knowledge and improve health.

Workforce development is a cornerstone of NLM’s mission. Training initiatives that prepare researchers for the evolving challenges of data-driven research will complement training that provides rapid, just-in-time and just-as-needed preparation for scientists and others. Current programs will be expanded and revised to seize scientific opportunities, form new partnerships, build capacity, and ensure a robust workforce for scientific discovery in the coming decades. Novel models of training, from three-minute videos to immersive, group-based problem-solving challenges, will engender team science skills and interdisciplinary learning. Initiatives to increase the size and expertise of the workforce will include a focus on strategies to achieve greater demographic diversity in biomedical informatics and data science. NLM will also engage in efforts to inspire the next generation in the power of data and information for progress in science and better health.

## Objective 3.1: Expand and enhance research training for biomedical informatics and data science

NLM has primary responsibility for ensuring an adequate supply of advanced research scientists to conduct research in biomedical informatics and data science and to serve as mentors and trainers of the next generation. Important skills for all researchers include being prepared to “compute in context;” extract meaning and insight from aggregations of data; and create new ways to analyze, visualize, mine, and integrate data and information. Specialists must be developed to meet the challenge of enabling dynamic, real-time curation at the scale and complexity that the future of biomedical data and information holds.

NLM will train a new generation of data scientists to complement its current doctoral training for translational research, bioinformatics, consumer health informatics, and clinical and public health informatics. The emphasis on data science and management of large and complex biomedical big data will be expanded in NLM’s existing university-based training programs. This will include infusing data science with security and privacy

management concepts in addition to next-generation curation, dissemination, and analytic approaches within data science. This increased focus will produce cadres of scientists who can develop a new array of tools and approaches for data, moving analytics beyond statistics, pre-

---

Specialists must be developed to meet the challenge of enabling dynamic, real-time curation at the scale and complexity that the future of biomedical data and information holds.

---

sentation beyond visualization, and more. NLM is strongly committed to ensuring an enduring professoriate for biomedical informatics, library and information science, and data science.

Intramural training at NLM will be expanded through full post-doctoral fellowships in bioinformatics and computational biology, educational rotations in biomedical informatics, and post-graduate fellowships in library and information science. Leveraging an intramural research program on advanced analytics, NLM will offer

short-term and extended training engagements involving interdisciplinary teams of scholars to craft the next generation artificial intelligence and machine learning methodologies. Particular emphasis will be placed on generating innovative methods to manage and disseminate the vast electronic resources of the NLM.

## Objective 3.2:

### Assure data science and open science proficiency

Data science—accelerating insights from data through purposeful methodologies—and open science—a commitment to making all aspects of the research process from design to data accessible to scientists, professionals, and society at large—combine to create a new paradigm for training. Preparing a workforce to address biomedical data science problems will involve enhancing the computational and statistical skills of researchers with biomedical knowledge, and training computer scientists to apply their work to biomedical problems. In developing training approaches, NLM will combine these two cross-training goals to support the emerging field of biomedical data science. To expand its training initiatives, NLM will partner with other agencies, universities, and online education providers.

**Prepare clinicians and biomedical researchers to utilize data resources and interpret data science-based research findings.** NLM will collaborate across NIH to develop a core curriculum for data science training embedded in institute specific training programs. It will also help ensure that domain specific university training and mid-career training awards meet proficiencies for data science. Working with partners, NLM will identify strategies to facilitate data science instruction in innovative ways such as modular approaches or self-directed learning, encompassing such topics as advanced analytics, data standards, data management, reproducibility, and data sharing.

**Training for stewardship.** Data stewardship is fundamental to open science and must become a fundamental component of every research

project and program. The commitment to open science is expanding globally, with data sharing and reuse becoming the norm across many biomedical research disciplines. Investigators need assistance to ensure that data generated in research projects will be findable, accessible, interoperable, and reusable (FAIR) and will be effectively managed and curated across the research life cycle. As NIH develops new policies that guide data management and data sharing,

---

Preparing a workforce to address biomedical data science problems will involve enhancing the computational and statistical skills of researchers with biomedical knowledge, and training computer scientists to apply their work to biomedical problems.

---

NLM will develop tools and resources that help research teams understand and comply with policies.

### Promote competencies in open science and data science for information professionals.

As librarians and other information professionals meet the evolving needs of scientific research communities, they play a particularly important role in fostering open science, data management, data sharing, and development of repositories. Programs that enable such roles in promoting data and open science support—provided by NLM, through NNLM, or in partnerships with professional associations—ensure the availability of a trained workforce to support data science, data reuse, and responsible data stewardship.

With the growing participation in citizen science, public librarians require special support and information on data management and interpretation of data-driven discovery.

NLM will share responsibility with educational institutions, industry, and other federal agencies to ensure the development of a technical workforce capable of supporting data-driven discovery and advanced data management. Post-baccalaureate training programs in data management, medical library science, and clinical research data support are three such opportunities.

## Objective 3.3: Increase workforce diversity

NLM is committed to developing a diverse workforce, inclusive of a broad range of racial and ethnic groups, individuals with disabilities, and individuals from economically or educationally disadvantaged backgrounds. Beyond workforce participation, NLM's commitment to diversity extends to advocating for diversity of thought and plurality of methods. A commitment to diversity is grounded in the belief that participation of a diverse workforce improves team performance and engenders more robust knowledge representations and more culturally-competent means of supporting investigations and delivering health information.

The responsibility of NLM for meeting its commitments to enhancing diversity will be met in several ways. First and foremost is a commitment to ensuring curricular integrity in its training programs, which requires explicit consideration of the completeness and robustness of knowledge representation and literature selection and organization. Second, NLM must increase the representativeness and participation of individuals from underrepresented groups in its university-based training programs. NLM

will devise new incentives to recruit and retain underrepresented minorities within the programs it sponsors, and NLM will work with stakeholders to consider multi-focused strategies for increasing visibility and interest in these fields among underrepresented populations.

NLM will partner with NIH agencies and professional societies to support programs that help high school and college students consider careers in health information technology, medical

---

Beyond workforce participation, NLM's commitment to diversity extends to advocating for diversity of thought and plurality of methods.

---

librarianship, and data science. NLM will leverage existing partnerships, such as the consortium of minority serving institutions that are members of NLM's Environmental Health Information Partnership (EnHIP).



## Objective 3.4:

### Engage the next generation and promote data literacy

Generations who grew up in a technological-rich world bring a certain skill set to health sciences and biomedical services. Developing a workforce that is well prepared to meet the challenges of 21st century science still requires initiatives to develop an enthusiasm for science and data at an early age. NLM will cultivate and inspire the next generation of biomedical informatics researchers, data scientists, and information professionals. In collaboration with public and private partners, NLM will develop initiatives to spark students' interest in science, technology, engineering, and mathematics (STEM). These initiatives will include ways to identify skills that will be needed and design targeted educational resources for learners from elementary school to college.

Engagement with data is not limited to those who work in formal research or clinical roles. The public needs a basic level of scientific and data literacy. NLM will work with organizations, educational institutions, and public libraries to contribute to increasing public understanding and appreciation of science.

---

---

Developing a workforce that is well prepared to meet the challenges of 21st century science still requires initiatives to develop an enthusiasm for science and data at an early age.

---

---

# Appendices

## Appendix A

### References and Notes

National Institutes of Health (2015)

*NIH-Wide Strategic Plan, Fiscal Years 2016-2020: Turning Discovery Into Health.* [online]

Bethesda, MD: National Institutes of Health.

Available at: <https://www.nih.gov/sites/default/files/about-nih/strategic-plan-fy2016-2020-508.pdf>

[Accessed 23 Jan 2018].

National Institutes of Health

Advisory Committee to the Director, National Library of Medicine Working Group (2015).

*Final Report—June 11, 2015.* [online]

Bethesda, MD: National Institutes of Health.

Available at: <https://acd.od.nih.gov/documents/reports/Report-NLM-06112015-ACD.pdf>

[Accessed 23 Jan 2018].

National Institutes of Health, Data and Informatics Working Group (2012).

*Draft Report to The Advisory Committee to the Director.* [online]

Available at: <https://acd.od.nih.gov/documents/reports/DataandInformaticsWorkingGroupReport.pdf>

[Accessed 23 Jan 2018].

---

<sup>1</sup> NLM contributes to the Trans-NIH Minority Health and Health Disparities Strategic Plan and complies with Section 2031(c)(B) of the 21st Century Cures Act.

## Appendix B

### NLM Authorizing Legislation and Executive Branch Assignments

**1956** The National Library of Medicine Act of 1956 (Public law 84-941) established the National Library of Medicine (NLM) by renaming the Armed Forces Medical Library and transferring the Library to the Public Health Service in the Department of Health, Education, and Welfare (DHEW). The Act defines NLM's purpose "... to assist the advancement of medical and related sciences and to aid the dissemination and exchange of scientific and other information important to the progress of medicine and to the public health." Its basic functions are to acquire, preserve, and make available materials pertinent to medicine, broadly defined; to prepare and make available indexes, catalogs, and bibliographies of the materials; and to provide reference and research assistance. The Act also established the NLM Board of Regents to advise, consult with, and make recommendations to the Secretary on matters of Library policy, including the acquisition of materials, the scope, content, and organization of Library services, and the rules under which its materials, publications, facilities, and services are made available to users.

- 1965** The Medical Library Assistance Act of 1965 (Public Law 89-241) authorized the NLM to make extramural grants for research, training, and development of information resources and to establish Regional Medical Libraries, the origin of what is now called the National Network of Libraries of Medicine. Subsequent reauthorizations broadened the scope of the NLM's grant authority.
- 1967** In response to the expressed interest by the Special Subcommittee on Investigation of DHEW of the House Committee on Interstate and Foreign Commerce, the Surgeon General transferred the Public Health Service Audiovisual Facility to NLM from the Center for Disease Control and renamed it the National Medical Audiovisual Center (NMAC).
- In response to a Presidential request based on recommendations made in the 1966 report of the President's Science Advisory Committee, entitled "Handling of Toxicological Information," DHEW assigned NLM the responsibility to build a Toxicology Information Program (TIP).
- 1968** The NLM was transferred administratively to the National Institutes of Health as part of a broader reorganization of the DHEW.
- The Lister Hill National Center for Biomedical Communications (LHNCBC) was established at NLM via Senate Joint Resolution 193.
- 1988** The Health Omnibus Programs Extension Act established the National Center for Biotechnology Information (NCBI) at NLM to focus and expand the digital collection, storage, retrieval, and dissemination of the results of biotechnology research, and to support and enhance the development of new information technologies to aid in the understanding of the molecular processes that control health and disease.
- 1989** The amendment to the Public Health Service Act that established the Agency for Health Care Policy and Research (AHCPR) required the new agency to work with--and to transfer funds to--NLM to develop and enhance information services in the field of health services research, encompassing health technology assessment and the development of practice guidelines.
- 1993** The National Institutes of Health Revitalization Act of 1993 established the National Information Center for Health Services Research and Health Care Technology (NICHSR) at NLM to collect, store, analyze, retrieve, and disseminate information on health services research, clinical practice guidelines, and health care technology. It also expanded NLM's basic functions to include publicizing NLM services and promoting use of health information technology.
- 1997** The FDA Modernization Act directed NIH to develop a registry of selected clinical trials of FDA-regulated drugs for serious and life-threatening conditions. The Director, NIH assigned responsibility for creating the registry to NLM, resulting in the release of ClinicalTrials.gov in 2000.
- 2004** The Secretary of Health and Human Services (HHS) designated NLM as the HHS central coordinating body for clinical terminology standards.
- 2007** The FDA Amendments Act required expansion of ClinicalTrials.gov to serve as a registry of clinical trials of a broad set of FDA-regulated drugs and devices and to collect summary results data from registered trials.

- 2008** The Consolidated Appropriations Act required NIH-funded investigators to submit final peer-reviewed journal manuscripts to NLM's PubMed Central, making NIH's previously voluntary public access policy mandatory. This requirement was made permanent in the Omnibus Appropriations Act, 2009.
- 2013** The President's Office of Science and Technology Policy issued a directive requiring all Federal agencies with more than \$100 million per year in R&D expenditures to establish and implement public access policies similar to NIH's and to develop policies for increasing access to digital data. As of 2018, all HHS agencies and five other Federal departments and agencies had elected to use PubMed Central to implement their policies.

## Appendix C

# NLM Strategic Plan Planning Process

Overall development of the 2017-2027 strategic plan was guided by the Board of Regents Strategic Planning Subcommittee, which was appointed at the September 2016 meeting of the Board of Regents. Members are Daniel R. Masys (co-chair), Jill Taylor (co-chair), Robert A. Greenes, Eric J. Horvitz, James L. Olds, and Sandra I. Martin. Discussions with the Board were held during its meetings in 2017, with final approval at the February 2018 meeting.

Working in conjunction with NLM Planning staff, the Strategic Planning Subcommittee of the Board of Regents identified four themes as a framework for considering NLM's priorities and future directions, and around which public input was solicited.

These themes were:

- Advancing data science, open science, and biomedical informatics
- Advancing biomedical discovery and translational science
- Supporting the public's health: clinical systems, public health systems and services, and personal health
- Building collections to support discovery and health in the 21st century

In addition, the following topics were considered across the four themes: partnerships, user communities, user engagement and educational outreach, international engagement, health disparities, standards, infrastructure, workforce development, research needs and funding.

NLM engaged the broader community and its own staff in the strategic planning process through several approaches: Request for Information (RFI), Expert Working Groups, Institutional Training Site Visits, Staff Survey, and NLM Town Hall meetings.

### **Request for Information**

On November 8, 2016, NLM issued an RFI (<https://grants.nih.gov/grants/guide/notice-files/NOT-LM-17-002.html>) based on the four themes to solicit input from its broad stakeholder community on goals and priorities for NLM's next strategic plan. The RFI

was active through January 23, 2017. NLM received 111 responses from a wide range of stakeholders including librarians, researchers, historians, first responders, nurses, informaticians, associations, and the public.

In addition, more than 600 responses were reviewed in response to the 2015 RFI (<https://grants.nih.gov/grants/guide/notice-files/NOT-OD-15-067.html>) issued by NIH on behalf of the NLM Working Group of the Advisory Committee to the NIH Director (ACD); that report was designed to obtain input on a vision for the future of NLM in the context of NLM's leadership transition and emerging NIH data science priorities. The subsequent report of the ACD Working Group report issued in June 2015 also informed our planning process (<http://acd.od.nih.gov/reports/Report-NLM-06112015-ACD.pdf>), as did the ACD Data and Informatics Working Group Report of 2012 (<https://acd.od.nih.gov/documents/reports/DataandInformaticsWorkingGroupReport.pdf>), and the NIH-Wide Strategic Plan, Fiscal Years 2016-2020: Turning Discovery Into Health (<https://www.nih.gov/sites/default/files/about-nih/strategic-plan-fy2016-2020-508.pdf>).

### **Expert working groups**

Working groups of outside experts were convened around each of the four themes, with meetings held from March to May 2017. In these working groups, more than 100 people met to consider ways to propel the NLM forward, strategically and meaningfully. A fifth internal working group of NLM staff met to consider the reports from the four external groups, generate additional ideas, and synthesize and prioritize recommendations.

### **Training Site Visits**

NLM leadership made a series of visits to NLM's university-based training programs for research careers in biomedical informatics and data science. The purpose of the visits was to discuss the scope and characteristics of NLM's training programs, particularly local perceptions about the benefits and challenges being faced and their strategies for incorporating data science into their programs. Separate meetings were held with three groups at each location: faculty, program administrators, and trainees. Ten programs were visited between October 2016 and June 2017, including five visits to private institutions and 5 to public institutions.

### **NLM Staff Input**

A staff survey soliciting input about NLM operations was available online or in paper format, which enabled anonymous responses to be placed in secure suggestion boxes in each of the office buildings where NLM staff work. The NLM planning team reviewed the 112 survey responses.

Town Hall meetings open to all Federal employees and contractors at NLM were held in February 2017 and August 2017. The NLM leadership team shared information and responded to questions about the planning process and the needs of NLM moving into the next decade. The first meeting brought together 500 staff members, and 351 staff attended the second meeting. Staff questions covered a broad range of topics, including data science, collections, media, global health, partnerships at NIH and in the community, reorganization, communication, workforce development, facilities, and NLM culture.

## Working Group Participants

### Role of NLM in Advancing Biomedical Discovery and Translational Science

**Chair:**

*Arthur S. Levine, MD*

Senior Vice Chancellor for the Health Sciences  
John and Gertrude Petersen Dean, School of Medicine  
Professor of Medicine and Molecular Genetics  
University of Pittsburgh

**Members:**

*Todd Anderson, PhD*

Professor and Chair  
Department of Environmental Toxicology  
Institute of Environmental and Human Health  
Texas Tech University

*Gregory L. Armstrong, MD*

Director, Office of Advanced Molecular Detection  
U.S. Centers for Disease Control and Prevention  
National Center for Emerging and Zoonotic  
Infectious Diseases

*Rob Califf, MD, MACC*

Professor of Medicine  
Duke University School of Medicine

*Joshua C. Denny, MD, MS, FACMI*

Professor, Biomedical Informatics & Medicine  
Director, Vanderbilt Center for Precision Medicine  
Vice President for Personalized Medicine  
Vanderbilt University Medical Center

*Filipa Godoy-Vitorino, PhD*

Associate Professor  
Department of Natural Sciences  
Inter American University of Puerto Rico,  
Metropolitan Campus

*Kristi Holmes, PhD*

Director, Galter Health Sciences Library  
Director of Evaluation, Northwestern University  
Clinical and Translational Sciences Institute  
(NUCATS)  
Associate Professor, Preventive Medicine  
(Health and Biomedical Informatics) and Medical  
Education  
Northwestern University Feinberg School of  
Medicine

*George Hripcsak, MD, MS*

Professor and Chair of Biomedical Informatics  
Columbia University

*Maricel G. Kann, PhD*

Associate Professor  
Department of Biological Sciences  
University of Maryland, Baltimore County

*Sean Mooney, PhD*

Professor, Biomedical Informatics and Medical  
Education  
Chief Research Information Officer (CRIO)  
UW Medicine  
University of Washington

*Sheila Prindiville, MD, MPH*

Director, Coordinating Center for Clinical Trials  
National Cancer Institute, NIH

*Richard J. Roberts, PhD*

Chief Scientific Officer  
New England Biolabs



*Julie Segre, PhD*  
Senior Investigator  
National Human Genome Research Institute,  
NIH

*Christine E. Seidman, PhD*  
Thomas W. Smith Professor of Medicine and  
Genetics, and Investigator, HHMI  
Harvard Medical School, Brigham & Women's  
Hospital, Howard Hughes Medical Institute

*Jacquelyn Taylor, PhD, PNP-BC, RN,  
FAHA, FAAN*  
Associate Professor and Associate Dean  
of Diversity and Inclusion  
Yale School of Nursing

*Jill Taylor, PhD*  
Director, Wadsworth Center  
Faculty, Wadsworth School of Laboratory  
Sciences  
New York State Department of Health

*Consuelo H. Wilkins, MD, MSCI*  
Executive Director, Meharry-Vanderbilt Alliance  
Associate Professor of Medicine,  
Vanderbilt University Medical Center and  
Meharry Medical College  
Vanderbilt University School of Medicine

## Role of NLM in Supporting the Public's Health: Clinical Systems, Public Health Systems and Services, and Personal Health

### **Chair:**

*Suzanne Bakken, PhD, RN, FAAN, FACMI*  
The Alumni Professor of Nursing and Professor of Biomedical Informatics  
Columbia University

### **Members:**

*Julia Adler-Milstein, PhD*  
Associate Professor  
Schools of Information and Public Health  
University of Michigan

*Catherine Arnott-Smith, PhD, AMLS, MSIS*  
Associate Professor, School of Library and  
Information Studies  
University of Wisconsin-Madison  
Discovery Fellow, Living Environments  
Laboratory  
Wisconsin Institute for Discovery

*Renae E. Barger, MLIS*  
Associate Director, Research, Instruction and  
Clinical Information Services  
Program Lead, National Network of Libraries of  
Medicine, Middle Atlantic Region  
University of Pittsburgh, Health Sciences Library  
System

*Theresa Cullen, MD, MS*  
Associate Director, Global Health Informatics  
Program  
Regenstrief Institute, Inc.  
Visiting Associate Professor  
Department of Family and Community Medicine  
Indiana University

*Marcela Gaitán, MPH, MA*  
Senior Director for External Relations  
National Alliance for Hispanic Health

*Graciela Gonzalez-Hernandez, MS, PhD*  
Associate Professor of Informatics, Department  
of Epidemiology, Biostatistics and Informatics  
Perelman School of Medicine  
University of Pennsylvania

*Robert Greenes, MD, PhD*  
Professor and Ira A. Fulton Chair in Biomedical  
Informatics  
Arizona State University

*Shannon D. Jones, MLS, M.Ed., AHIP*  
Director of Libraries & Associate Professor  
Medical University of South Carolina Libraries

*Marty LaVenture, PhD, MPH, FACMI*  
Director  
Office of Health IT, e-Health and Informatics  
Minnesota Department of Health

*Cynthia A Lindquist, PhD, MPA*  
President  
Cankdeska Cikana (Little Hoop) Community  
College (CCCC)

*Thomas S. Nesbitt, MD, MPH*  
Associate Vice Chancellor, Strategic  
Technologies and Alliances  
Director, Center for Health & Technology  
Professor, Department of Family and Community  
Medicine  
UC Davis Health

*Casey Overby, PhD*  
Assistant Professor, Divisions of General Internal  
Medicine and Health Sciences Informatics  
Department of Medicine  
Johns Hopkins University

*Wanda Pratt, PhD*  
Professor, Information School  
Adjunct, Division of Biomedical & Health  
Informatics, Medical School  
University of Washington

*Eric C. Schneider, MD, MSc, FACP*  
Senior Vice President for Policy and Research  
The Commonwealth Fund

*William W. Stead, MD*  
Chief Strategy Officer  
McKesson Foundation Professor of Biomedical  
Informatics  
Professor of Medicine  
Vanderbilt University Medical Center

*Paul Tang, MD, MS*  
Vice President  
IBM Watson Health

*Jill Taylor, PhD*  
Director, Wadsworth Center  
Faculty, Wadsworth School of Laboratory  
Sciences  
New York State Department of Health

*Andrew (Andy) M. Wiesenthal, MD, SM*  
Managing Director  
Health Care Practice  
Deloitte Consulting, LLP

## Role of NLM in Building Collections to Support Discovery and Health in the 21st Century

### **Chair:**

*Patricia (Pat) L. Thibodeau, MLS, MBA*  
Associate Dean for Library Services and Archives (Retired April 2017)  
Duke Medical Center Library & Archives

### **Members:**

*Amy Bernard, PhD*  
Product Architect  
Allen Institute for Brain Science

*Laura Brown, MA, MFA*  
Executive Vice President  
Ithaka  
JSTOR Managing Director

*Sayed Choudhury, MS*  
Associate Dean, Research Data Management  
Director, Digital Research and Curation Center  
Sheridan Libraries  
Johns Hopkins University

*Yolanda Cooper, BGS, MLS*  
University Librarian  
Robert W. Woodruff Library  
Emory Libraries and Information Technology

*Wayne Crocker, MLS*  
Director of Library Services  
Petersburg Public Library  
(Petersburg, VA)

*Noémie Elhadad, PhD*  
Associate Professor  
Biomedical Informatics, Computer Science  
Columbia University

*E. Thomas Ewing, PhD*  
Professor of History  
Virginia Tech

*Lorraine J. Haricombe, PhD*  
Vice Provost and Director of Libraries  
University of Texas at Austin

*Michele Klein-Fedyshin, MSLS, BSN, RN, AHIP*  
Research & Clinical Instruction Librarian  
Health Sciences Library System  
University of Pittsburgh

*John A. Kunze*  
Identifier Systems Architect  
University of California Curation Center  
California Digital Library

*Kenneth Mandl, MD, MPH*  
Director  
Computational Health Informatics Program  
Boston Children's Hospital  
Donald A.B. Lindberg Professor of Pediatrics and  
Professor of Biomedical Informatics  
Harvard Medical School

*Sandra Martin, MSLS*  
Director, Shiffman Medical Library  
Wayne State University

*Neil Rambo, MLibr*  
Director  
Health Sciences Library  
New York University

*Nigam H. Shah, MBBS, PhD*  
Associate Professor of Medicine  
Stanford University

*Jonathan Teich, MD, PhD*  
Assistant Professor of Medicine and Emergency  
Medicine  
Harvard / Brigham and Women's Hospital

## Role of NLM in Advancing Data Science, Open Science, and Biomedical Informatics

### **Chair:**

*Russ Altman, MD, PhD*  
Director, Stanford Biomedical Informatics Training Program  
Professor, Bioengineering, Genetics, Medicine, Biomedical Data Science  
(and Computer Science, by courtesy)  
Stanford University

### **Members:**

*Suzanne Allard, PhD*  
Associate Dean for Research, College of  
Communication & Information  
Professor, School of Information Sciences  
University of Tennessee, Knoxville

*Francine Berman, PhD*  
Edward P. Hamilton Distinguished Professor in  
Computer Science  
Rensselaer Polytechnic Institute

*Christine L. Borgman, PhD, MLS*  
Distinguished Professor and Presidential Chair in  
Information Studies  
University of California, Los Angeles (UCLA)

*Wendy Chapman, PhD*  
Chair, Department of Biomedical Informatics  
University of Utah

*Greg Farber, PhD*  
Director, Office of Technology Development and  
Coordination  
National Institute of Mental Health, NIH

*Lisa Federer, MLIS, MA*  
Research Data Informationist  
NIH Library

*Eric Horvitz, MD, PhD*  
Director and Technical Fellow  
Microsoft Research

*Debra R. Lappin, JD*  
Principal  
Faegre Baker Daniels Consulting

*Erez Lieberman Aiden, PhD*  
Assistant Professor  
Baylor College of Medicine

*Clifford Lynch, PhD*  
Executive Director  
Coalition for Networked Information (CNI)  
Washington, DC

*Elizabeth Marincola, MBA*  
Senior Advisor for Science Communications and  
Advocacy  
The African Academy of Sciences

*Patricia Matthews-Juarez, PhD*  
Professor, Department of Family and Community  
Medicine/Vice President for Faculty Affairs and  
Development  
Meharry Medical College, and  
Chair of the NLM Environmental Health  
Information Partnership Advisory Board

*Lucila Ohno-Machado, MD, PhD, MBA*  
Professor of Medicine  
Chair, Department of Biomedical Informatics  
Associate Dean for Informatics and Technology  
University of California San Diego

*Brian Nosek, PhD*  
Executive Director and Co-Founder  
Center for Open Science (COS)

*James L. Olds, PhD*  
Assistant Director  
Directorate for Biological Sciences  
National Science Foundation

*Kevin B. Read, MLIS, MAS*  
Assistant Curator  
NYU Health Sciences Library  
NYU School of Medicine

*Chris Shaffer, MS, AHIP*  
University Librarian and Associate Professor  
Oregon Health & Science University

*R. Joe Stanley, PhD, MS*  
Associate Professor, Electrical and Computer  
Engineering  
Associate Chairman, Computer Engineering  
Missouri University of Science and Technology  
(Rollo, MO)

*Xuemao Wang, MBA, MLIS, MLS*  
Dean and University Librarian  
Special Advisor to Provost on China Initiative  
University of Cincinnati Libraries

*Susan J. Winter, PhD*  
Associate Dean for Research  
Co-Director, Center for Advanced Study of  
Communities and Information (CASCI)  
College of Information Studies  
University of Maryland

## Working Group of NLM Internal Experts

### Chair:

*Michael F. Huerta, PhD*

NLM Coordinator of Data & Open Science Initiatives  
Associate Director for Program Development

### Members:

*Stacey Arnesen, MS*

Head, Office of Disaster Information Management  
Research Center (DIMRC)  
Division of Specialized Information Services

*Joyce Backus, MS*

Associate Director for Library Operations

*Ivor D'Souza, MS*

Director, Information Systems  
Office of Computer and Communications  
Systems

*Todd Danielson, MS*

Executive Officer and Associate Director for  
Administrative Management

*Dina Demner-Fushman, MD, PhD*

Staff Scientist, Computer Engineering Branch  
Lister Hill National Center for Biomedical  
Communications

*Valerie Florance, PhD*

Associate Director for Extramural Programs

*Kathryn Funk, MLIS*

Program Specialist, PubMed Central  
National Center for Biotechnology Information

*Pertti (Bert) Hakkinen, PhD*

Acting Head, Office of Clinical Toxicology  
Division of Specialized Information Services

*Stefan Jaeger, PhD*

Visiting Scientist, Computer Engineering Branch  
Lister Hill National Center for Biomedical  
Communications

*David Landsman, PhD*

Chief, Computational Biology Branch  
Bioinformatics of Chromatin Structure Group  
National Center for Biotechnology Information

*Wei Ma, MS*

Chief, Applications Branch  
Office of Computer and Communications  
Systems

*Jennifer Marill, MS, MA*

Chief, Technical Services Division  
Division of Library Operations

*James G. Mork, MS*

Computer Scientist, Computer Engineering  
Branch & Lead Developer and Technical Project  
Lead, Medical Text Indexer (MTI) project  
Lister Hill National Center for Biomedical  
Communications

*Duc Nguyen, MS*

Chief, Medical Language Branch  
Office of Computer and Communications  
Systems

*Ilene Mizrahi, PhD*

Staff Scientist, GenBank  
National Center for Biotechnology Information

*Jeffrey S. Reznick, PhD*

Chief  
History of Medicine Division  
Division of Library Operations

*Valerie Schneider, PhD*

Staff Scientist, RefSeq  
National Center for Biotechnology Information

*Jerry Sheehan, MS*

Assistant Director for Policy Development  
Office of the Director

*Hua-Chuan Sim, MD*

Chief Program Officer, Clinical and Public Health  
Informatics Research, Applied informatics and  
Career Awards  
Extramural Programs Division

## NLM Board of Regents Subcommittee

### **Co-Chairs:**

*Daniel R. Masys, MD*  
Affiliate Professor  
Biomedical and Health Informatics  
University of Washington School of Medicine

*Jill Taylor, PhD*  
Director, Wadsworth Center  
Faculty, Wadsworth School of Laboratory Sciences  
New York State Department of Health

### **Members:**

*Robert Greenes, MD, PhD*  
Professor and Ira A. Fulton Chair in Biomedical  
Informatics  
Arizona State University

*Eric Horvitz, MD, PhD*  
Director and Technical Fellow  
Microsoft Research

*Sandra Martin, MSLS*  
Director, Shiffman Medical Library  
Wayne State University

*James L. Olds, PhD*  
Assistant Director  
Directorate for Biological Sciences  
National Science Foundation

## NLM Planning Team

*Patricia Flatley Brennan, RN, PhD*  
Director  
National Library of Medicine, NIH

*Betsy L. Humphreys, MLS*  
Deputy Director (retired June 2017)  
National Library of Medicine, NIH

*Jerry Sheehan, MS*  
Deputy Director (effective July 2017)  
National Library of Medicine, NIH

*Michael F. Huerta, PhD*  
NLM Coordinator of Data & Open Science  
Initiatives  
Associate Director for Program Development

*Barbara A. Rapp, PhD*  
Chief, Office of Planning and Analysis  
Office of Health Information Programs  
Development

*Rebecca M. Goodwin, JD*  
Data Science Specialist  
Office of Health Information Programs  
Development

*Elizabeth Robboy Kittrie, MMHS*  
Strategic Advisor for Data Science and Open  
Science  
Office of Health Information Programs  
Development





U.S. National Library of Medicine

U.S. Department of Health and Human Services  
Public Health Service  
National Institutes of Health  
National Library of Medicine