NIH U.S. National Library of Medicine

Synopsis of A Platform for Biomedical Discovery and Data-Powered Health

Strategic Plan 2017–2027

NLM is poised to address the challenges laid out since its inception — not by building a single service to address each one, but by knitting together the best of several services to efficiently and effectively advance health and biomedical discovery through information.

—Patricia Flatley Brennan, RN, PhD, NLM Director

The 21st Century has seen a tremendous shift in the practice and products of science — from cultural shifts within the scientific community towards a more open science enterprise to the steady stream of big and complex data enabled by next-generation technologies.

As the largest biomedical library in the world, the National Library of Medicine (NLM) is committed to meeting the evolving needs of the research and clinical communities and to serving the public at large. As an institute of the National Institutes of Health (NIH), NLM will bring its research focus in information science, informatics, and data science to meet the challenges of this rapidly changing biomedical space.

NLM has recently completed its strategic plan spanning 2017 to 2027, the details of which are summarized in Table 1. 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 healthcare. NLM resources — from PubMed and GenBank to ClinicalTrials.gov and MedlinePlus — have opened the channels for communication necessary to advance research from bench to bedside, while promoting rigor and reproducibility. NLM's products must keep pace with the increasing volume and variety of data generated by the scientific community to meet the demand for immediate access to needed information. By increasing the speed at which information is organized, disseminated, and accessed, NLM will accelerate the speed and precision of discovery.

The NLM Strategic Plan envisions NLM as a platform for biomedical discovery and data-powered health, integrating streams of complex and interconnected research outputs that can be readily translated into scientific insights, clinical care, public health practices, and personal wellness.

NLM envisions three pillars as the foundation for such a platform:

1. Innovate, create, and maintain a sustainable digital ecosystem to keep pace with the data demands of the research enterprise

2. Engage a wide range of audiences to ensure the right information gets delivered to them at the right time.

3. Inspire and empower the data-driven workforce of the future.

The NIH has affirmed its dedication to the future of data-powered health and data-driven discovery through large-scale data-intensive programs like the BRAIN Initiative, the *All of Us* Research Program, and the Cancer Moonshot. These pioneering initiatives have promoted revolutionary techniques in data generation, processing, and analysis, which in turn demand innovations in data infrastructure, management, and sharing. NLM's strategic plan echoes these commitments to data science and open science, which remain central to its mission and vision for advancing human health and discovery.

NLM will achieve its vision in the coming decade by engaging stakeholders in the public and private sectors, including researchers, librarians, health professionals, entrepreneurs and innovators, and the public. NLM takes enormous pride in the partnerships and engagement it has attracted since its inception and remains committed to preserving and protecting the public's trust, while harnessing the power and potential of data.

1 Accelerate data-driven discovery and advance human health

In the not-so-distant past, a given scientific investigation could be broken down into three interconnected research products: (1) the data collected during the course of experiments; (2) the tools and software used to draw conclusions from the data; (3) the resulting publication that reports conclusions drawn from the data and analysis.

Today these research products are largely digital, as are other products and processes, including models and simulations, data-standards, workflows, and more. Nevertheless, while databases and repositories exist for a range of digital research objects, they are rarely linked. Furthermore, greater emphasis is placed on end publications — rather than the underlying data and analysis — providing only a flattened view of the scientific process. An emerging digital environment — including new methods, more openness, and greater computational and storage capacity — provides the opportunity to link the various processes and products of science into an interconnected ecosystem of digital research objects, bringing a multi-dimensional view of the research endeavor into full relief.

NLM aims to support innovation that explores alternate paths of re-analysis and discovery through dynamic visualizations and executable articles.

NLM intends to create computable libraries of data, models, literature, and more to devise an interlocked environment that enables and enhances discovery and stimulates insight. NLM will continue to develop new collections and resources, while also forming strategic partnerships to integrate and connect existing valuable resources supported by its sister institutes and centers of NIH and beyond. In the process, NLM will leverage its leadership to advocate for research practices and policies that foster open science and scholarship. NLM will link its expanding collections into an ecosystem that ensures digital research objects are FAIR (findable, accessible, interoperable, and reusable). Facilitating data reuse will work to enhance reproducibility and rigor, a central priority for NIH-supported research. And, ensuring discoverability and data access can underscore the interconnections between research outputs throughout the course of scientific inquiry.

Transforming the information ecosystem into a dynamic and interlocking network of research products hinges on innovation powered by cutting edge research conducted and supported by NLM. The increasing pace of data generation, deposition, and demand have rendered manual curation unsustainable. Rather, the future lies in curation at scale to build automatic and autonomous databases equipped to match research outputs. To query those expanding databases and systems will require next-generation min-ing approaches driven forward by advances in machine learning and natural language processing. NLM aims to support innovation that explores alternate paths of re-analysis and discovery through dynamic visualizations and executable articles, challenging the paradigm of publication as an endpoint of inquiry. As these technologies and capabilities develop, NLM is committed to furthering research and fostering partnerships that advance the appropriate security protocols that promote the chain of trust in health information.

2 Maximize NLM's impact through enhanced engagement

A library is only as effective as the users it reaches. Smart community engagement is therefore critical to assuring NLM's resources most effectively reach our wide range of audiences — from librarians to researchers and clinicians, from teenagers to their parents, from policymakers to healthcare professionals.

As data inputs and outputs have shifted to encompass voice, music, image, and virtual reality applications, the power of data has become more accessible to a wider population. The revolution in information technologies has created an opportunity to move beyond a one-size-fits-all approach to user engagement to meet users' individualized information needs.

Person-centric and community-conscious design strategies lie at the heart of NLM's vision for enhanced engagement. Anticipating the needs of different populations under different contexts is essential to deploying tailored approaches for information dissemination. NLM will therefore employ rapid cycle design and delivery approaches to gain insights into what information users seek and why. A fine-grained understanding of user needs will allow NLM to better serve communities across the United States. Delivering personalized and actionable health information is particularly powerful as it allows individuals to make the best health decisions for their unique needs.

NLM remains committed to building bridges of trust between the biomedical community and the public at large. Even in an increasingly digital age, human networks remain crucial to maximizing NLM's impact. NLM will continue to leverage the 6,500 member libraries of the National Network of Libraries of Medicine (NNLM), which act as trusted ambassadors between NLM and the communities they serve. NNLM has become a key actor in NIH's *All of Us* Research Program, working with public libraries to engage local communities, raise awareness of the program, and improve health literacy. By remaining engaged with the broader public, NLM will continue to level the health information playing field to ensure the power of data-driven health.

NNLM has become a key actor in NIH's *All of Us* Research Program, working with public libraries to engage local communities, raise awareness of the program for populations unrepresented in biomedical research, and improve health literacy.

The pace of change in biomedical data science is moving at lightning speed… It is critical that NIH pave the path forward for data science, and this move will enable researchers, medical practitioners and many others to use the wealth of vast knowledge and data available to them through the NLM.

— Francis S. Collins, MD, PhD, NIH Director

3 Build a data-ready workforce for the future

Workforce development is a cornerstone of NLM's mission. To continue powering and advancing a vibrant biomedical and health ecosystem rooted in data, NLM will expand its training programs in biomedical informatics and data science.

As a leader in information science, NLM will continue to provide direction for building a broader biomedical workforce capable of data-driven discovery in concert with NIH institutes and centers. Biomedical research and human health increasingly require data literacy to transform data into actionable knowledge and improve health. Therefore, NLM must provide training both for scholars seeking expertise in biomedical informatics or data science *and* for biomedical or clinical domain experts seeking a basic level of competency to facilitate data management and analysis.

Data science has become a growing priority across NIH institutes and centers, playing a prominent role in the NIH-Wide Strategic Plan for 2016-2020 through visionary data-intensive initiatives. To support this priority, NLM's intramural and extramural training programs will emphasize data science and management of large and complex biomedical big data. NLM will expand its intramural programs through full post-doctoral fellowships and educational rotations in bioinformatics and computational biology, as well as post-graduate fellowships in library and information science. Beyond the NIH campus, NLM will support the current generation of biomedical informatics and data science trainees, as well as middle- and late-stage scientists who will serve as mentors for the next generation. Within the data revolution is an additional opportunity to build the culture that will power research into the future — one that is rooted in open science principles and scholarship. As a result, open science practices, such as data stewardship and reuse, will be embedded within NLM's training initiatives.

The data revolution also calls for a new generation of data-ready librarians and informationists to transform libraries into epicenters of data literacy.

The data revolution also calls for a new generation of data-ready librarians and informationists to transform libraries into epicenters of data literacy. Librarians and informationists should therefore play a critical role in advancing the data science skills and data literacy of the communities they serve — whether scholars, clinicians, patients, or the general public. While librarians can help users tap into and make sense of vast stores of knowledge, they can also empower users to channel the power of data for themselves through educational programming. Digital training innovations, such as hackathons, provide fertile ground for enhancing the data skills of their patrons.

As biomedical data science continues to emerge as an attractive field, NLM recognizes the importance of developing a talented workforce, capable of taking advantage of our increasingly technological world. NLM therefore acknowledges its responsibility to craft incentives

to recruit and retain talent within its programs. NLM is committed to fostering a variety of thought and plurality of methods, which are essential for innovation, while ensuring that the needs and concerns of various communities are represented within the future fabric of the field.

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 Engage the next generation and promote data literacy

Table 1. NLM Strategic Plan Overview for 2017 to 2027.

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

Cover image by Donny Bliss, Lister Hill National Center for Biomedical Communications,
National Library of Medicine, National Institutes of Health 2017