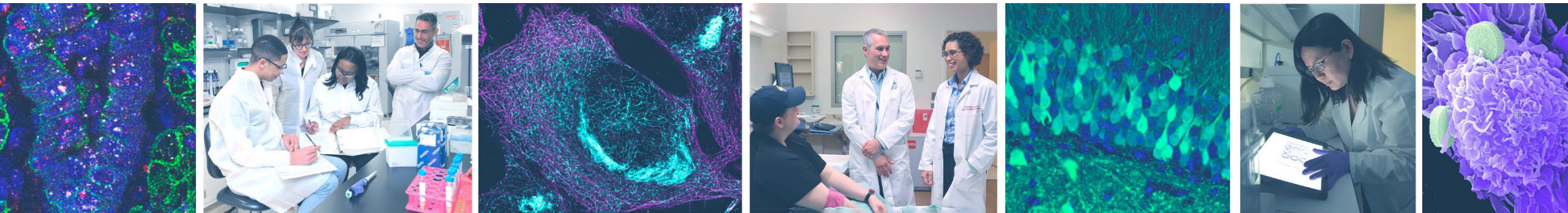


Building and Maintaining Data Communities: A Vision of the Future for the National Library of Medicine

128th Advisory Committee to the Director Meeting

June 14, 2024



Monica M. Bertagnoli, MD
Director, National Institutes of Health



Accelerating Discovery and Data-Powered Health



Accelerate discovery
and advance health
through data-driven
research



Reach more people
in more ways
through enhanced
dissemination and
engagement



Build a workforce
for data-driven
research and health



Powering Resources for the Scientific Community

- MedlinePlus®
- PubMed Central®
- GenBank®
- ClinicalTrials.gov
- dbGaP
- Sequence Read Archive (SRA)
- Digital collections
- Terminology standards
- ...and more



Three of HHS's Top Five Websites

- ncbi.nlm.nih.gov
- [PubMed.gov](https://pubmed.gov)
- [MedlinePlus.gov](https://medlineplus.gov)

Serving Science and Society Since 1836

Research Enterprise for Biomedical Informatics & the World's Largest Biomedical Library



1836-1968:

A Collection of Books and Journals

With its roots in the office of the U.S. Army Surgeon General, Congressional authorization moved NLM to the Public Health Service.

In 1962, NLM opened its doors on the NIH Campus in Bethesda, MD.



1968-2000:

Foundation of a Modern Library

Expansion and impact in the Information Age:

- Lister Hill National Center for Biomedical Communications
- National Center for Biotechnology Information



2000-2036:

The 21st Century Library

Leading innovative research to accelerate NIH's mission and reach scientists and society with trusted health information.



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

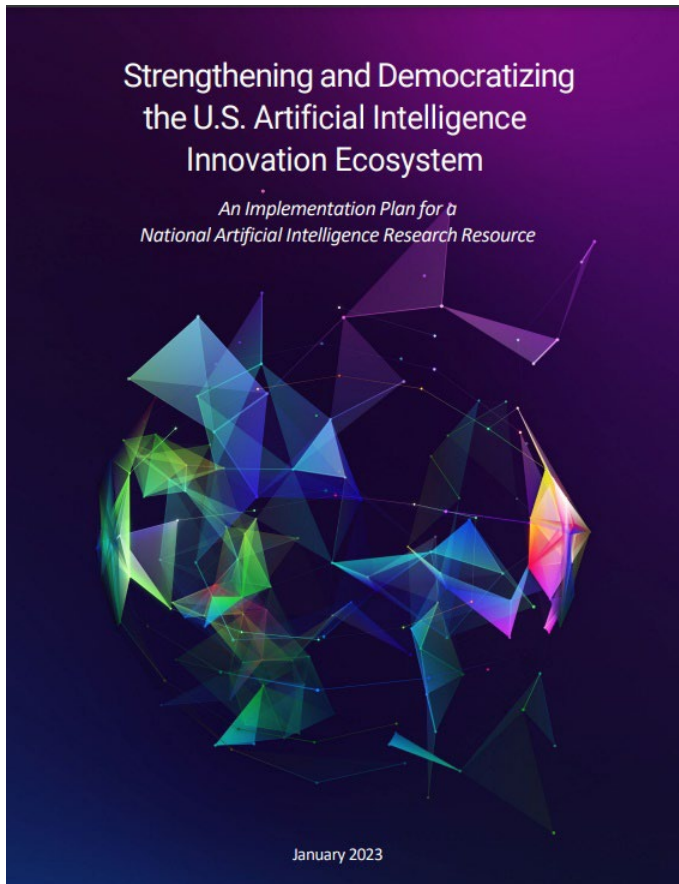
NLM envisions three pillars as the foundation for such a platform:

- 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.

Delivering the data required to use Artificial Intelligence to improve health



National AI Research Resource: a shared research infrastructure facilitating access to compute, software, datasets, models, training and user support for researchers and students



Objective: To strengthen and democratize the U.S. AI Innovation ecosystem in a way that protects privacy, civil rights, and civil liberties

Goals:



Spur
innovation



Increase the
diversity of
talent in AI

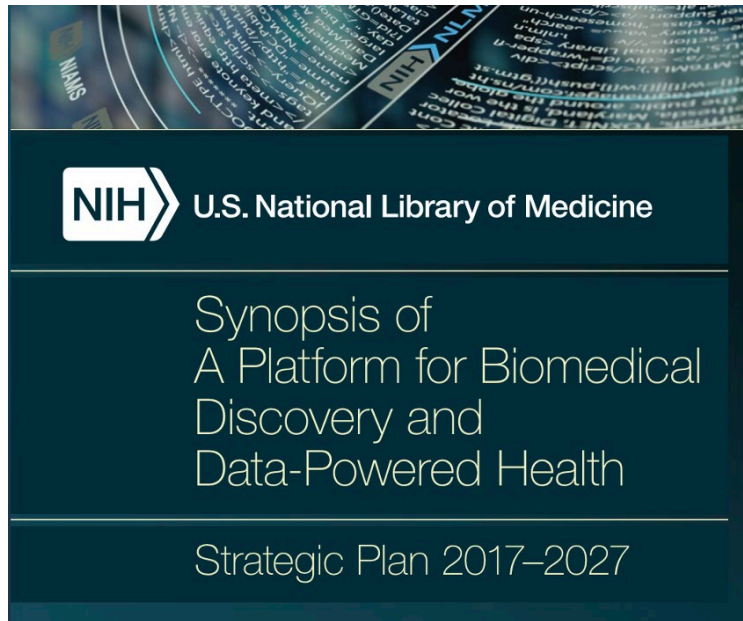


Improve U.S.
capacity for AI
R&D



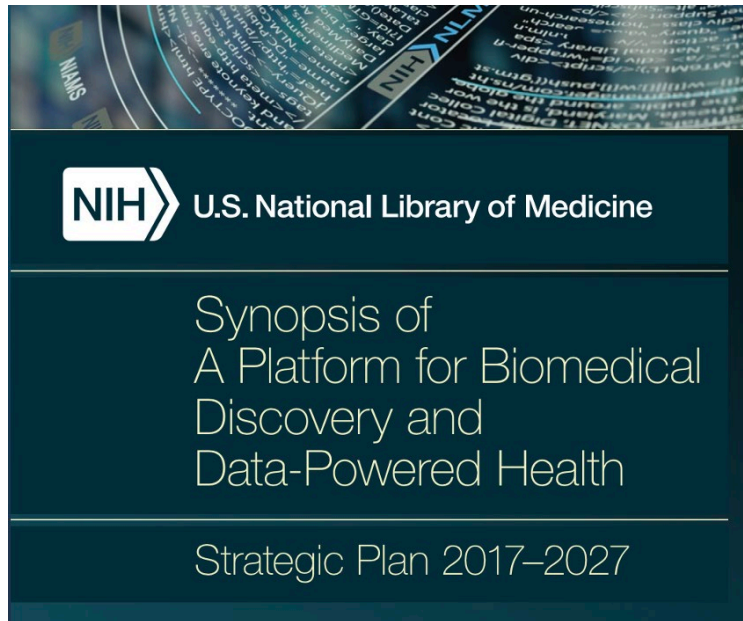
Advance
trustworthy AI

NSF | NIH | DOE | NASA | NOAA



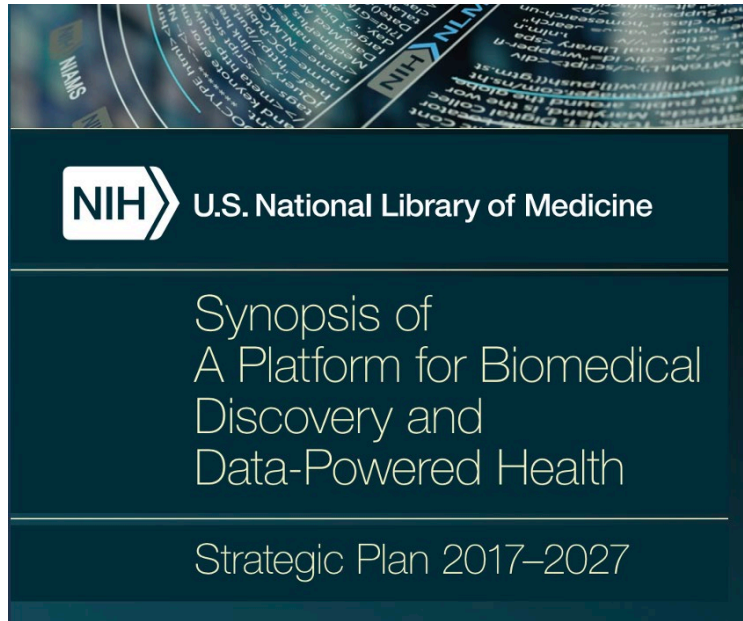
Federated biomedical research data *sharing infrastructure*:

- Publication data sets for **all** NIH-funded research
 - Centralized catalog of data holdings and use models
 - Dramatic expansion of data storage capabilities
- HHS-integrated infrastructure to obtain data from the clinical care environment



Collaborative data sharing and ***use environments*** and data access models

- Examples: Genomic Data Science Analysis, Visualization, and Informatics Lab-space (AnVIL), Health Equity Action Network (HEAN), Medical Imaging and Data Resource Center (MIDRC), National COVID Cohort Collaborative (N3C), NCI Clinical Trials Data Archive, NCI Genome Data Commons



Education programs to promote equitable resource availability and use

- Data science learning center

NIH Helping to End Addiction Long-term® Initiative (NIH HEAL Initiative®)

Goal: speed scientific solutions to stem the national opioid and pain public health crises

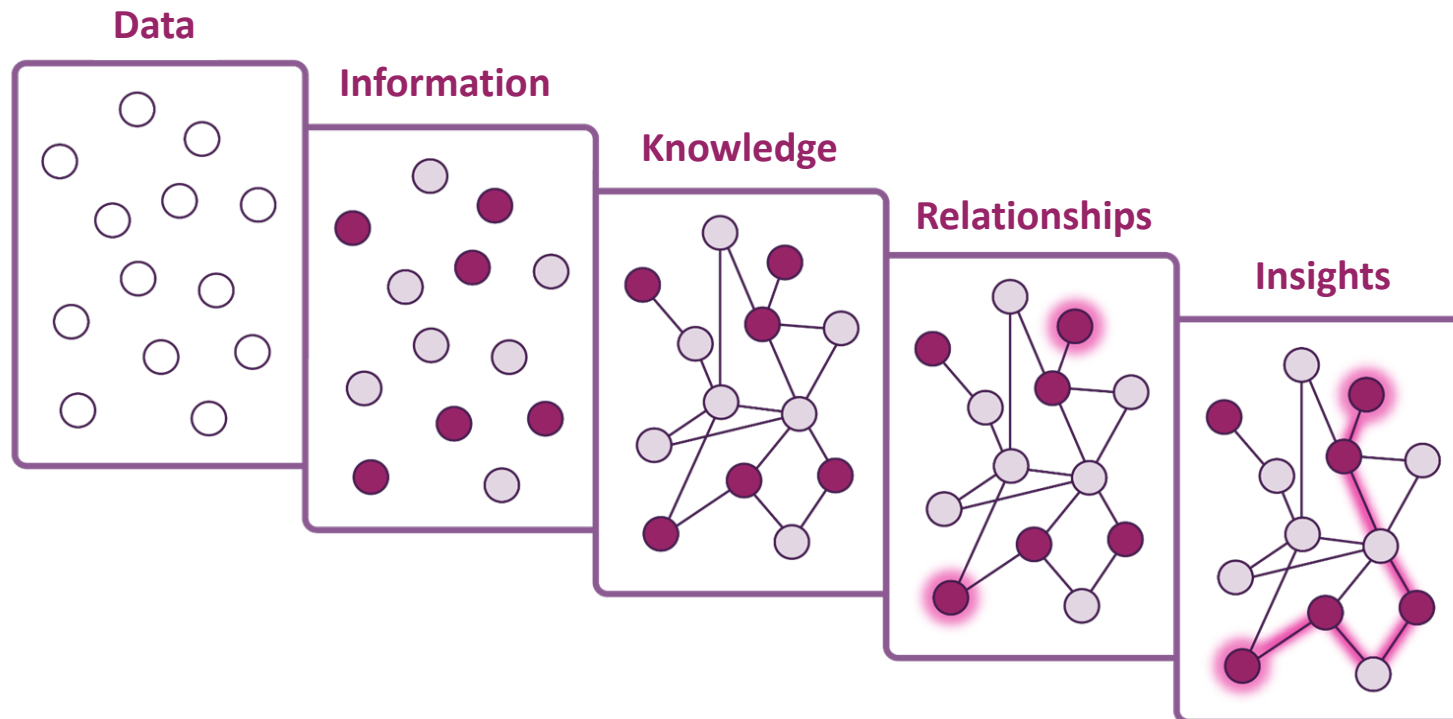
(established 2018)



Image credit: Zoran Zeremski/Shutterstock

The HEAL Data Ecosystem: Setting an Example for Open Science

New solutions to opioid use disorder and pain management demand transparency. Open science and maximal data sharing are critical for driving scientific discoveries.



User-friendly infrastructure and centralized data management resources to translate independent, disconnected data into connected, well-annotated insights

A Platform for Biomedical Discovery and Data-Powered Health



Serve as an epicenter for NIH-funded research to **advance information science, analytics and data science**

Provide educational programs to equitably **expand the biomedical informatics and data science capabilities** of the research workforce

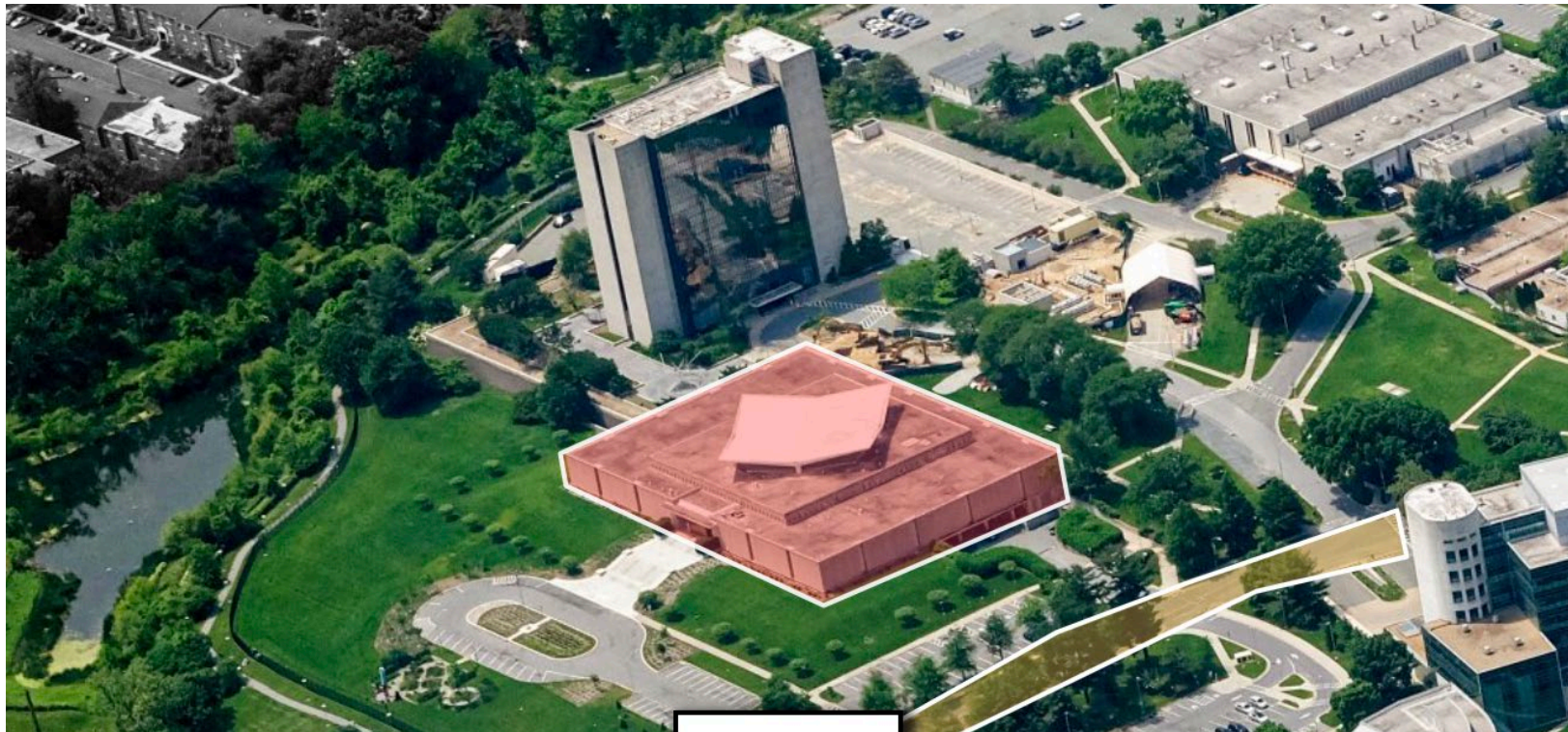
Serve as an interactive and productive **leader** within the **international community of biomedical data science organizations**, working to provide stability and functionality of the broader biomedical data ecosystem

Maintain an open, scalable, **federated digital ecosystem** for access to biomedical data and analytics

Provide a focus for **collection** and **from the clinical care** environment (i.e. "real world" data)

Provide a "front door" where biomedical or clinical research domain experts, trainees, patients, advocates, or others seeking answers to novel research questions can **obtain access to required data and analytic tools**, as well as **consultation in data science**

Serve as an interactive and productive **leader** within the **international community of biomedical data science organizations**, working to provide stability and functionality of the broader biomedical data ecosystem



Provide a "front door" where biomedical or clinical research domain experts, trainees, patients, advocates, or others seeking answers to novel research questions can **obtain access to required data and analytic tools**, as well as **consultation in data science**



Maintain an open, scalable, **federated digital ecosystem** for access to biomedical data and analytics



Provide a focus for **collection** and **from the clinical care** environment (i.e. "real world" data)



Data from the clinical care environment

A cross-agency initiative addressing an urgent need for high-quality, comprehensive data from the clinical care environment to generate evidence required for decision-making to improve health

Why:

- Facilitate learning health system initiatives by achieving better data to assess health outcomes that matter to individual people and to society overall
- Eliminate costly data formatting and collection redundancies that create silos
- Reduce clinical care site burden for data submission
- Increase data quality and speed time to data access for use in agency decision-making
- Apply artificial intelligence methods to health data that are comprehensive and represent the diversity of the US population

This new data initiative will create a learning lab to drive development and implementation of data standards and data collection systems that enable data sharing and fit-for-purpose use by any HHS agency and office

Serve as an epicenter for NIH-funded research to **advance information science, analytics and data science**

Extramural and Intramural Research Programs:

Conducting innovative research and training in computational biology and health sciences



Evolutionary Genomics and Biomolecular Structure



Health Information Standards and Clinical Discovery



Image Processing



Natural Language Processing



Networks, Gene Regulation, and Chromatin



Statistical Methods

Provide educational programs to equitably **expand the biomedical informatics and data science capabilities** of the research workforce

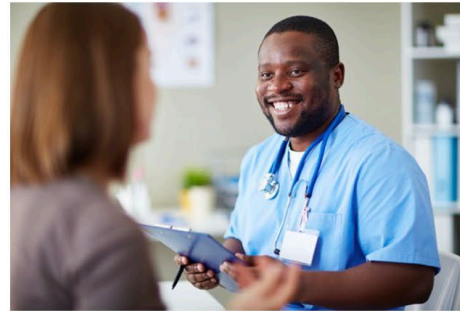
Biomedical and Clinical Informatics at NLM



Health IT and Health Data Standards

Efficient health care information exchange in the US and worldwide is made possible by NLM's work with IT Data Standards.

[LEARN ABOUT NLM'S CONTRIBUTIONS TO HEALTH IT](#)



Unified Medical Language System (UMLS) Terminology Services

This set of tooling services brings together many health and biomedical vocabularies and standards to enable interoperability between computer systems.

[EXPLORE UMLS](#)



Biomedical Informatics Training Program

This training program provides biomedical and clinical informatics training and research opportunities for individuals at various stages in their career.

[INVESTIGATE TRAINING OPPORTUNITIES](#)

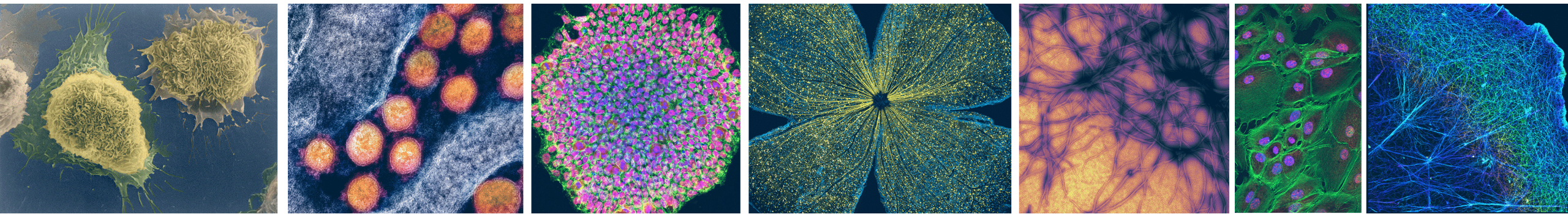


[Home](#) > Job Openings

Job Openings at NLM

Open Positions

- [Staff Scientist 1 - Algorithms for Computational Biology](#) - *Applications will be accepted until the position is filled.*
- [Software Development Section Lead](#) - *Applications will be accepted from May 6, 2024 through June 6, 2024.*
- [Staff Scientist 1 - Sequence Read Archive Development Team](#) - *Applications will be accepted until the position is filled.*
- [Director, National Library of Medicine](#) - *Applications must be received by 11:59 P.M. ET, Monday, July 1, 2024*
- [Staff Scientist 1 - Human Regulatory Genomics](#) - *Applications will be accepted until the position is filled.*



NIH

Turning Discovery Into Health

