

AMIA 2016

Anticipating the NLM's 3rd Century

PATRICIA FLATLEY BRENNAN, RN, PHD

DIRECTOR

NATIONAL LIBRARY OF MEDICINE

High School Participants



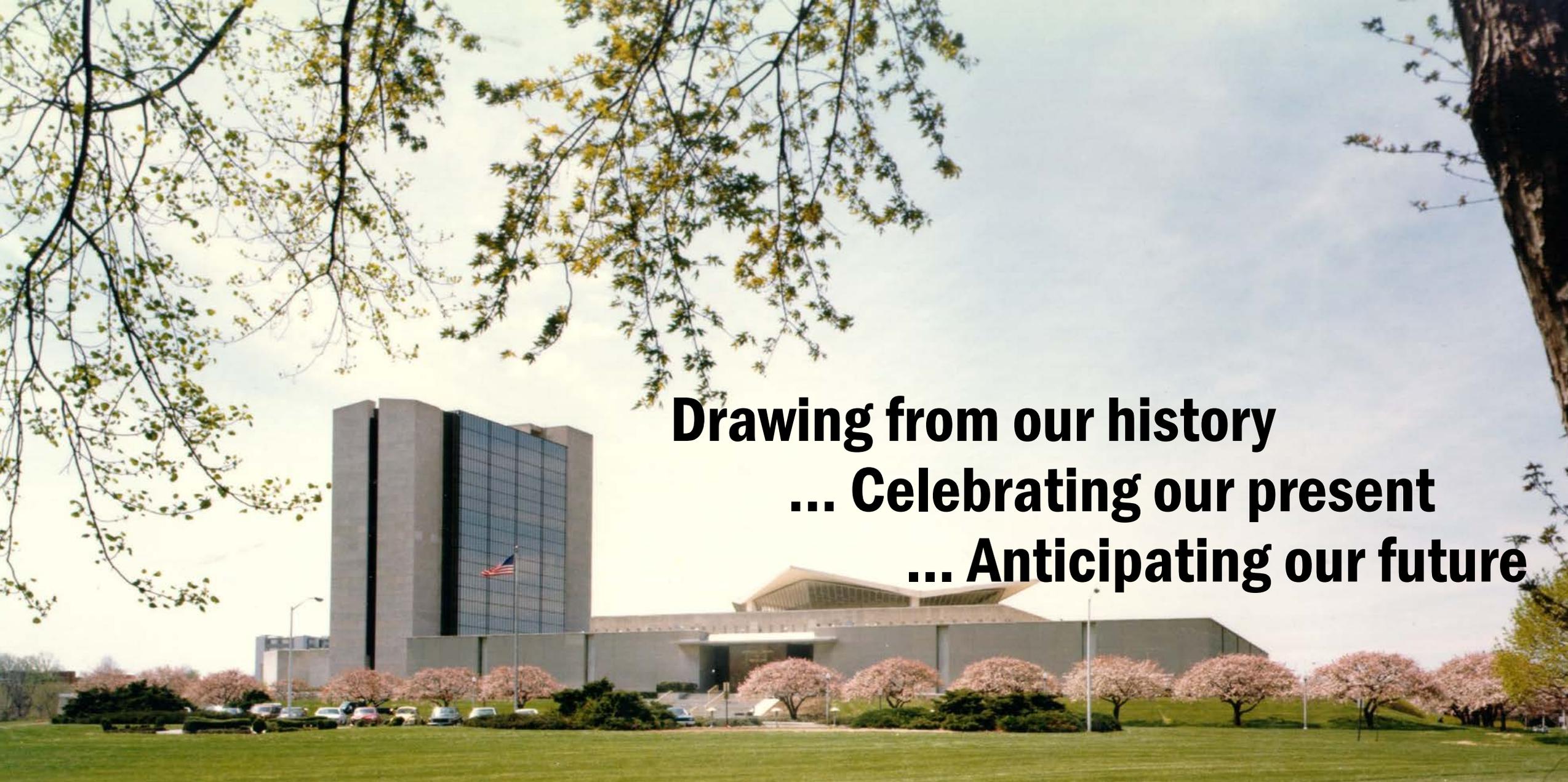
- Fiona Cai
- Rohan Chalasani
- Kevin Chen
- Catherine Chu
- Anay Gupta
- Callie Jones
- Shania Khatri
- Elliot Nam
- Chi Nguyen
- Veer Sangha
- Anoochree Sengupta
- Olivia Zhang

New NLM Staff here today

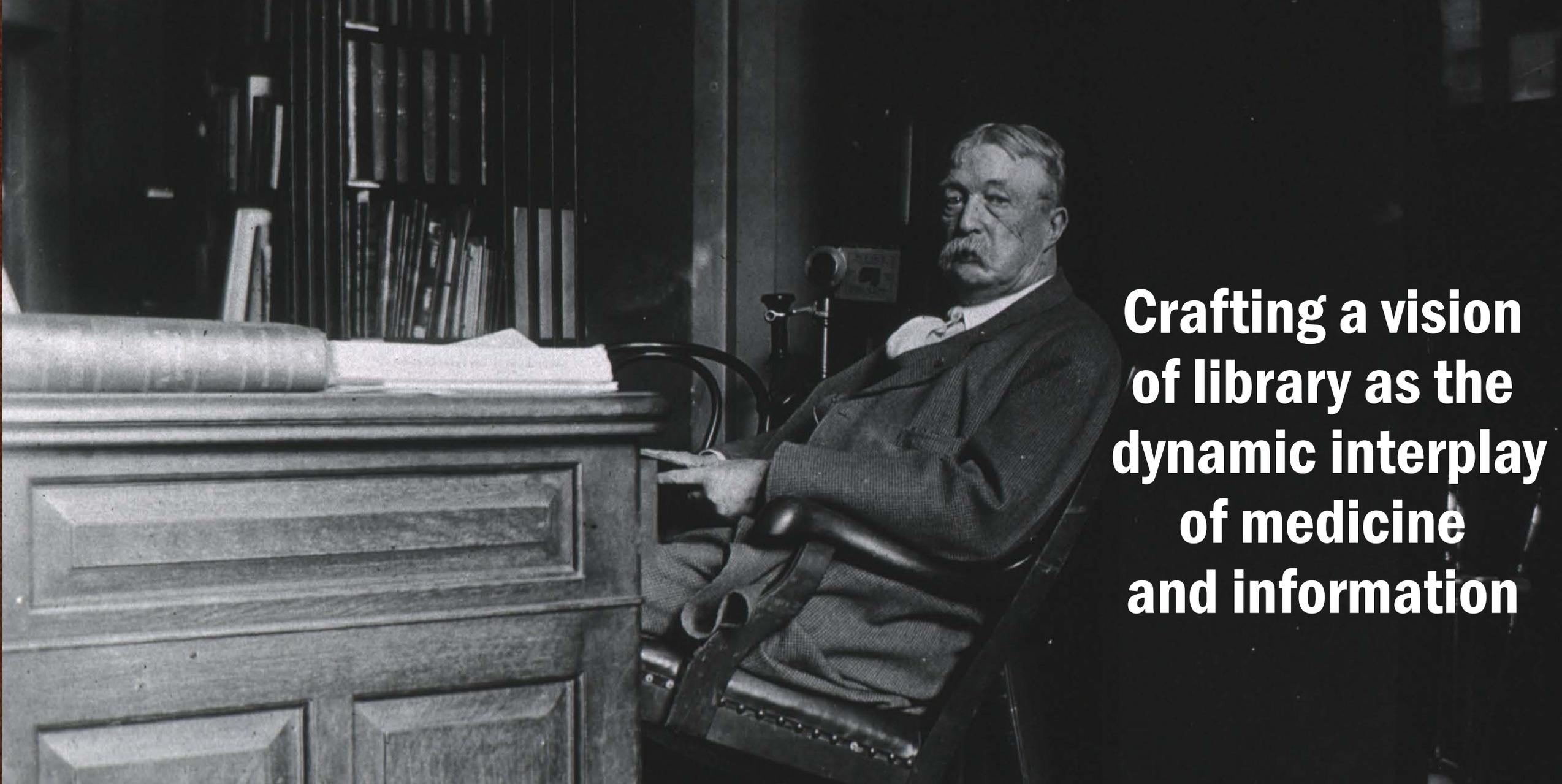
- Library Associates
 - **Kendra Godwin**
 - **Megan Kellner**
 - **Tyler Moses**
 - **Candace Norton**
- Public Health Policy Analyst
 - **Dana Casciotti, PhD**

How do I think NLM will fare
under the new administration?

“NIH has a long history of bi-partisan support and stands ready to work with the new Administration to improve people's health and reduce the burden of disease through biomedical research.”



**Drawing from our history
... Celebrating our present
... Anticipating our future**



**Crafting a vision
of library as the
dynamic interplay
of medicine
and information**



U.S. National Library of Medicine

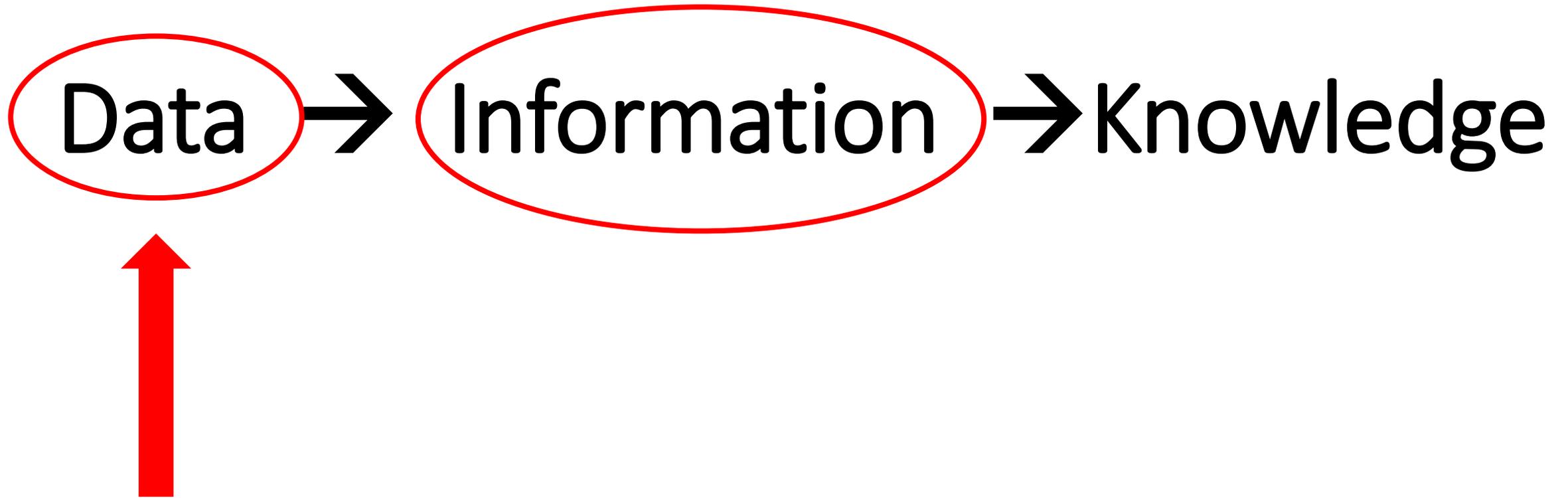


The NLM will shape the future of data-driven discovery

...beyond genomics

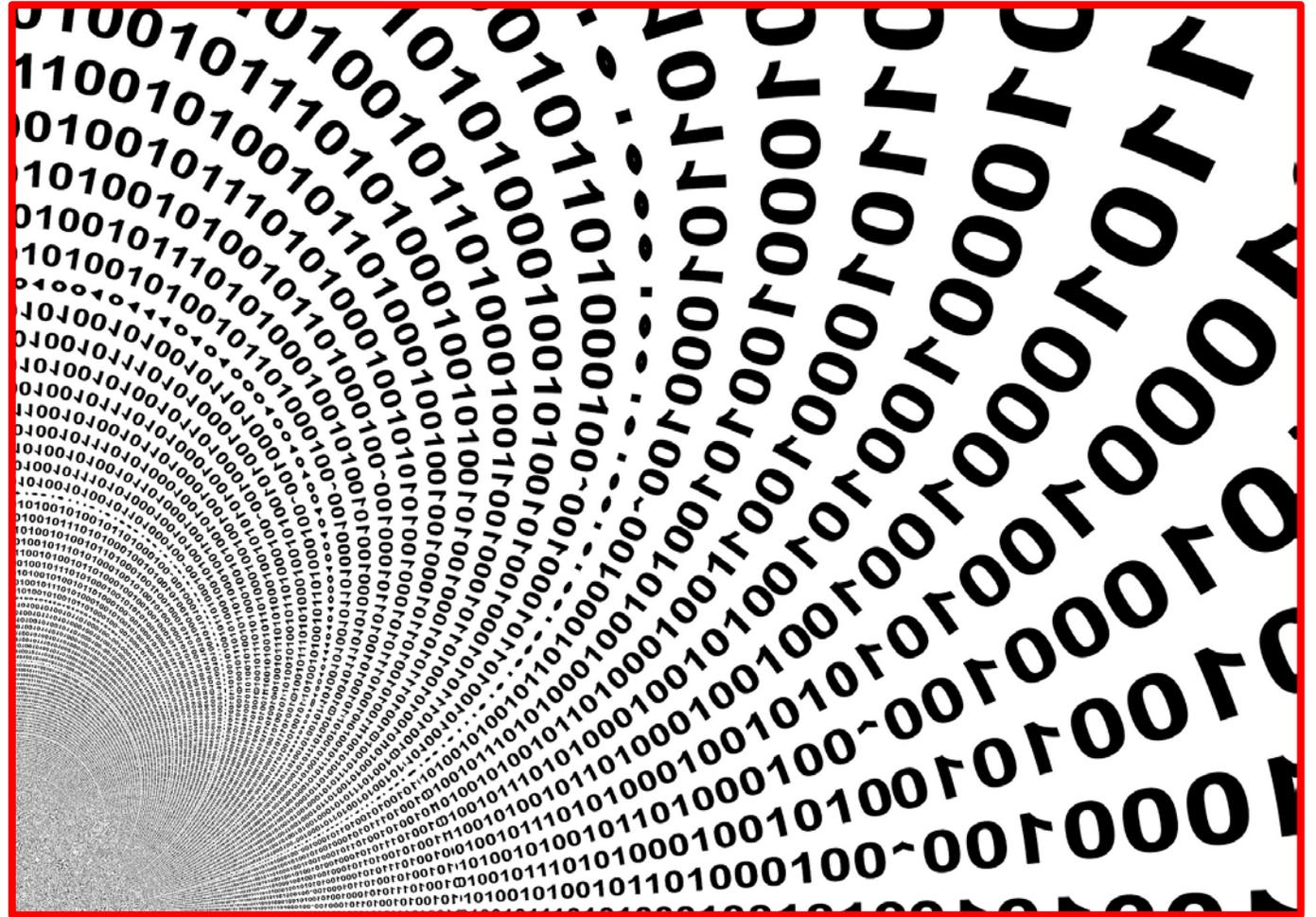
...more than in the cloud!





Our future builds on our past

- 1838-~1960 – Index
- ~1970-1983 – Digitize
- 1984-2015 – Network
- 2016-∞ – Data





In the words of Nina Matheson...

- ...be tool builders and...system developers...and solvers information problems (1982)
- Choices need to be made
- The major barrier to change is often *not a love of the status quo* but the lack of a clear picture of where technology leads
- Some futurists underestimate what can be accomplished in one years and over-estimate what can be accomplished in ten years.

What Matheson & Cooper envisioned, still holds...

- that books will not disappear, but their uses will concentrate in those areas where their portability, relative low cost and ease of production for **stable** information in incontestable
- As more and more information is available only to those with the financial means, fears grow about the potential of **disenfranchisement for those less affluent**
- The high costs of **connectivity**...
- The ... problem is to filter the external knowledge base and **bring in only that which is relevant and useful at the time** and place for it to accelerate the effectiveness of work

Nina's view of our time...

- Lessening dependence on centralized work places
- Electronic mail might save \$4.35 per message
- Growth in majors in technical fields
- Information base doubles every 10 years, faster in health
- Concerns about data security and the lack of confidentiality
- Academic resources include faculty, libraries, information & instructional material

Now remember, when she was writing...

- Searching data bases can range from \$16 to \$100 per connect hour in addition to \$5 to \$8 per hour telecom costs and 5¢ to 20¢ per citation
- “A new crisis is emerging that is equal to that of the 1960’s” – Societal view is information access as an advantage; now it is viewed as essential
- In 1982 ,the business community was pushing computers and “the flow of technology is so rapid as to acquire a momentum of its own and sweep us into lifestyles we many not like”
- Libraries had been in a bad state and the NLM responded with \$\$ to advance information flows through academic medical centers

Recommendation for the future

- ... the quantity of material to be reviewed...still exceeds the time available to devote to this necessary work...
 - Technology *could be* used to filter the databases and there *is a need for* changes to personal information file management
- ...NLM *should* develop state of the art network communication throughout the RMLN
- ...work closely with other national libraries, industry, and library groups to develop appropriate archival storage and retrieval devices and systems, and to intensify research efforts in developing appropriate technologies to improve information transfer and utilization mechanisms



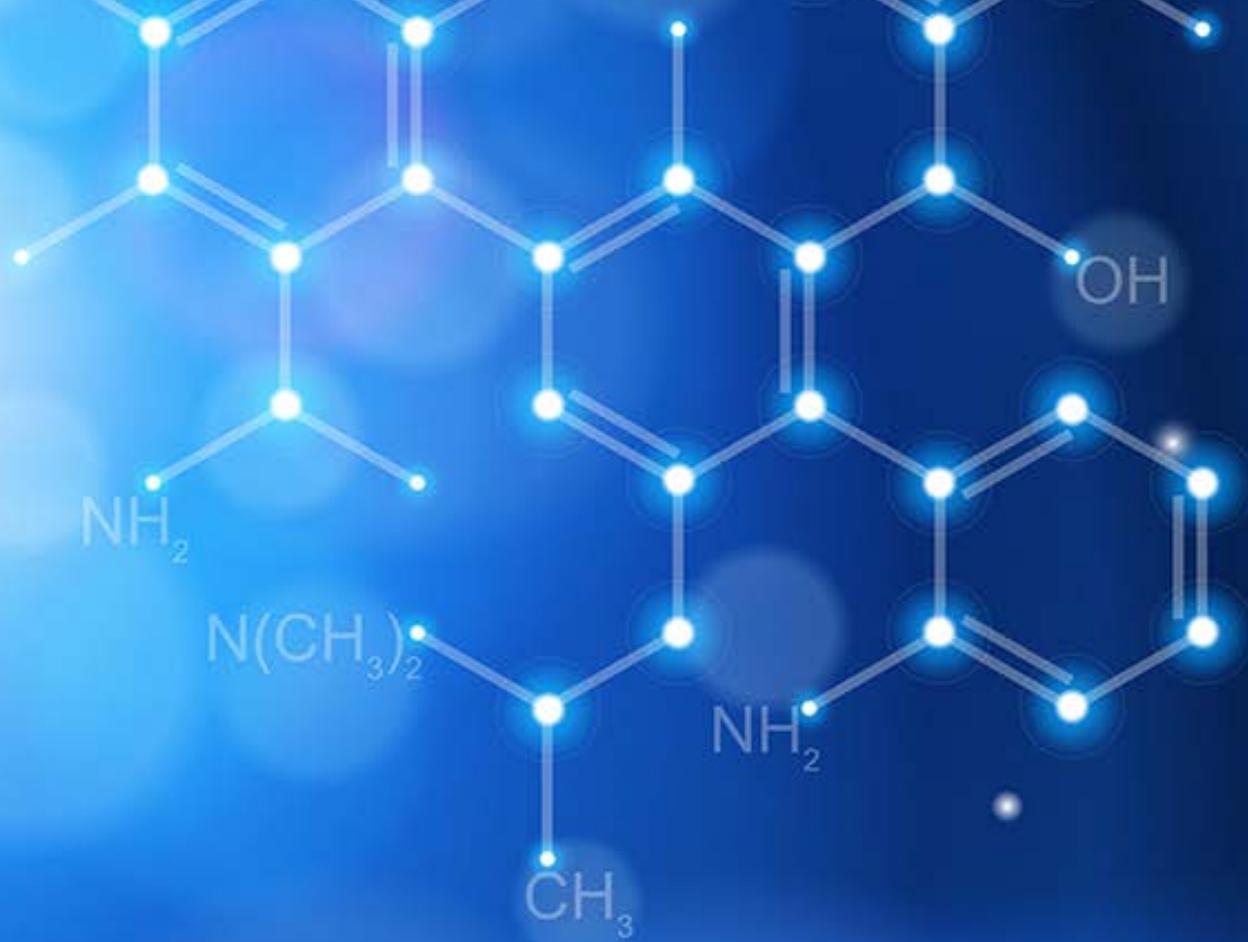
It worked,
and the idea of
what a library *is*
became wonderful!



U.S. National Library of Medicine



Supporting discovery



National Library of Medicine Collection

National Library of Medicine Collection

- ~17,000 serial titles
- All Serials – journals, annuals, statistics, etc.
- Discoverable in NLM Catalog & LocatorPlus
- NLM provides ILL and ensures archiving

PubMed

- 26.5 M records
- Only bibliographic records

MEDLINE

- ~5600 selected journal titles
- Only bibliographic records
- 93% of PubMed

PMC = PubMed Central Archive

- ~2,000 full participation journals
- 4 million full text articles
- ~1M federally funded public access articles
- Bibliographic records display in PubMed

NCBI: Research-driven service

Service-inspired research



Literature

PubMed, PMC, Books

~4,000,000 daily users

Browsing, reading
Small downloads

Molecular and Clinical Data

Gene, Nucleotide, Protein
Structure, BLAST, PubChem, Clinical Trials

~300,000 daily users

Batch retrieval
Data analysis
Extensive NCBI curation
Big Data

Specialized Datasets

SNP, GEO, SRA
dbGaP, ClinVar

~15,000 daily users

Retrieving whole datasets
Expensive per user
Extensive Curation
Big Data



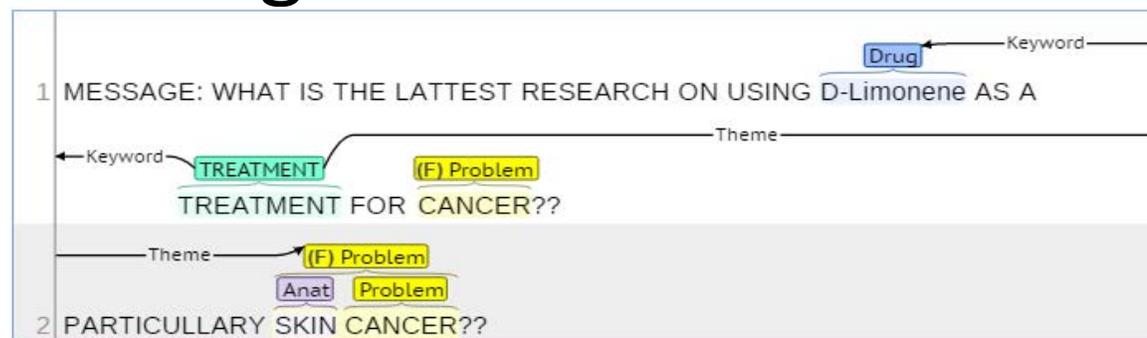
Biomedical Informatics R&D at the Lister Hill Center

- MetaMap Lite
 - Open source: <http://metamap.nlm.nih.gov/#Downloads>
- Consumer-health Question Answering

- Classification for Medical Questions
- Recognizing Question Entailment
- Combining Knowledgebases

• RxNORM & RxNAV

- Accesses continue to grow beyond a Billion
- Soft launch of RxNav 2.0



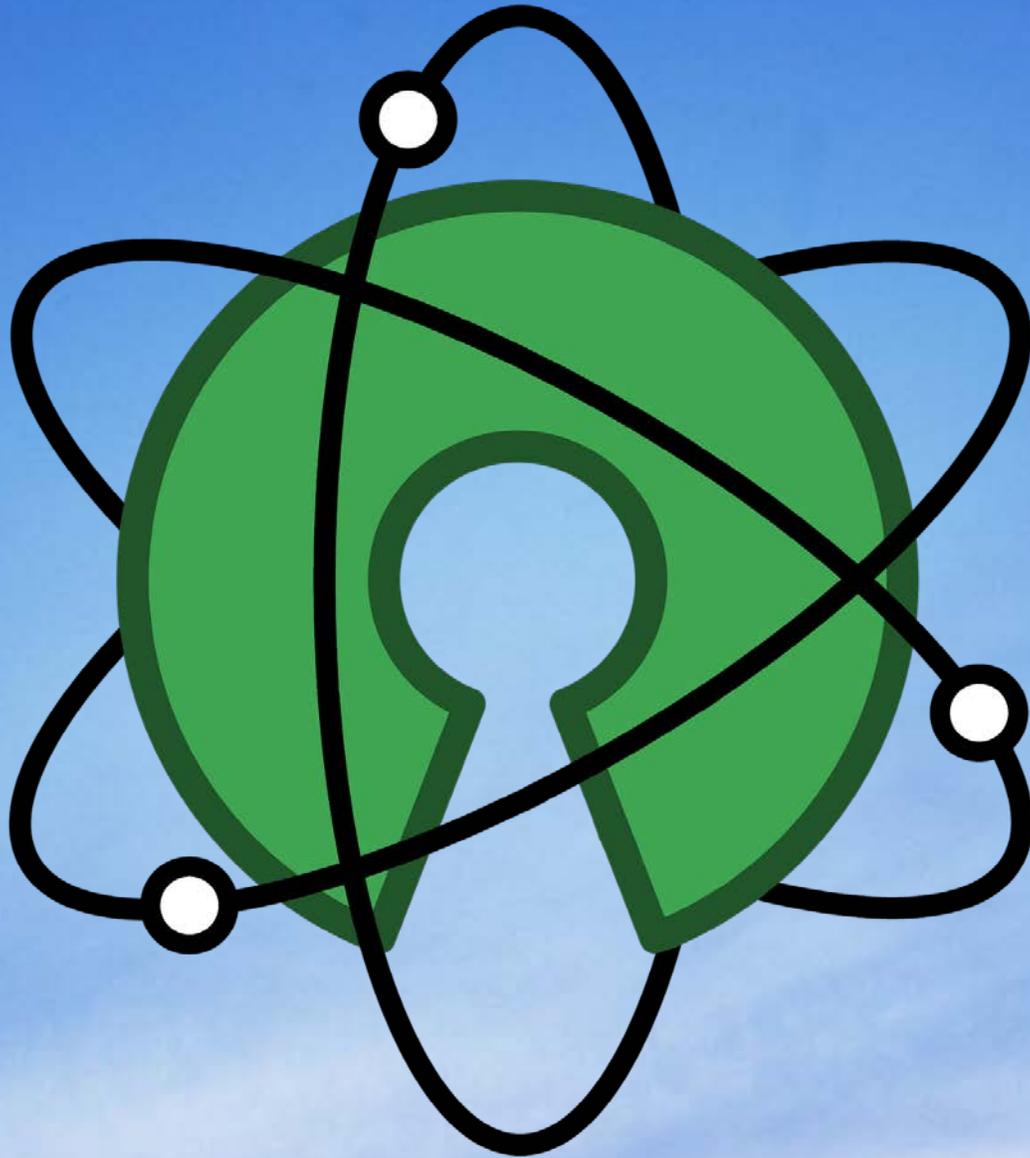
Biomedical Image Informatics R&D at the Lister Hill Center

Image Analysis, Feature Extraction, Machine Learning, Visual Content Understanding, Information Retrieval

- Open-iSM, an image + text search engine
<https://openi.nlm.nih.gov>
- Automated analysis of digitized histology images to detect cervical cancer
- Automated screening of CXR
 - Normal vs. Pulmonary disease (TB)
 - Currently deployed in rural Kenya
- Automated extraction of bibliographic data from scanned and Web articles to populate MEDLINE[®]/PubMed[®]
- Face matching to search for missing disaster victims in People Locator[®]
 - <https://pl.nlm.nih.gov>
- Automated detection of malaria from image analysis of microscope (cell) images

Preserving the past





Opening science



U.S. National Library of Medicine



NLM Pill Image Recognition Challenge

Consumer Safety through Public Data Science Research

Can your cell phone tell you what that pill is?

August 1, 2016 Prize announced:

**1st Mobile Pill Finder – Michigan State Univ.
\$25,000**



HEALTH DATA

surgery 0
clinical test
medications
blood pressure
lab test 52%
vaccination 82%
BMI normal



10-may-14

patient #08001

gender ♂
age 23
HR 95 bpm
120/60
ECHO D
CD PWR<500
Frq 2.0 MHz
1800 mm
AO 100%



Improving health care through standards & specialized information



U.S. National Library of Medicine



Now Available: Final Rule for FDAAA 801 and NIH Policy on Clinical Trial Reporting

[Find Studies](#) ▾ [About Clinical Studies](#) ▾ [Submit Studies](#) ▾ [Resources](#) ▾ [About This Site](#) ▾

ClinicalTrials.gov currently lists **228,054 studies** with locations in all 50 States and in **191 countries**.

Text Size ▾

Search for Studies

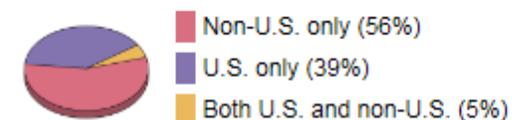
Example: "Heart attack" AND "Los Angeles"

[Advanced Search](#) | [See Studies by Topic](#)
[See Studies on Map](#)

Search Help

- [How to search](#)
- [How to find results of studies](#)
- [How to read a study record](#)

Locations of Recruiting Studies



Total N = 40,060 studies
(Data as of October 17, 2016)

- [See more trends, charts, and maps](#)

For Patients and Families

- [How to find studies](#)
- [See studies by topic](#)
- [Learn about clinical studies](#)
- [Learn more](#)

For Researchers

- [How to submit studies](#)
- [Download content for analysis](#)
- [About the results database](#)
- [Learn more](#)

For Study Record Managers

- [Why register?](#)
- [How to register your study](#)
- [FDAAA 801 requirements](#)
- [Learn more](#)

Learn More

- [Final Rule Webinar Series](#)
- [Tutorials for using ClinicalTrials.gov](#)
- [Glossary of common site terms](#)
- [For the press](#)
- [Using our RSS feeds](#)

[HOME](#)

[RSS FEEDS](#)

[SITE MAP](#)

[TERMS AND CONDITIONS](#)

[DISCLAIMER](#)

[CONTACT NLM HELP DESK](#)



PubMed Central®



PMC

PubMed Central® (PMC) is a free full-text archive of biomedical and life sciences journal literature at the U.S. National Institutes of Health's National Library of Medicine (NIH/NLM).

4.1 MILLION Articles

are archived in PMC.

Content provided in part by:

1949

Full Participation
Journals

323

NIH Portfolio
Journals

4015

Selective Deposit
Journals



PubMed

Direct Deposit of Data
in support of any
PMC article
as a supplement
by October, 2017

and life sciences
National Library of

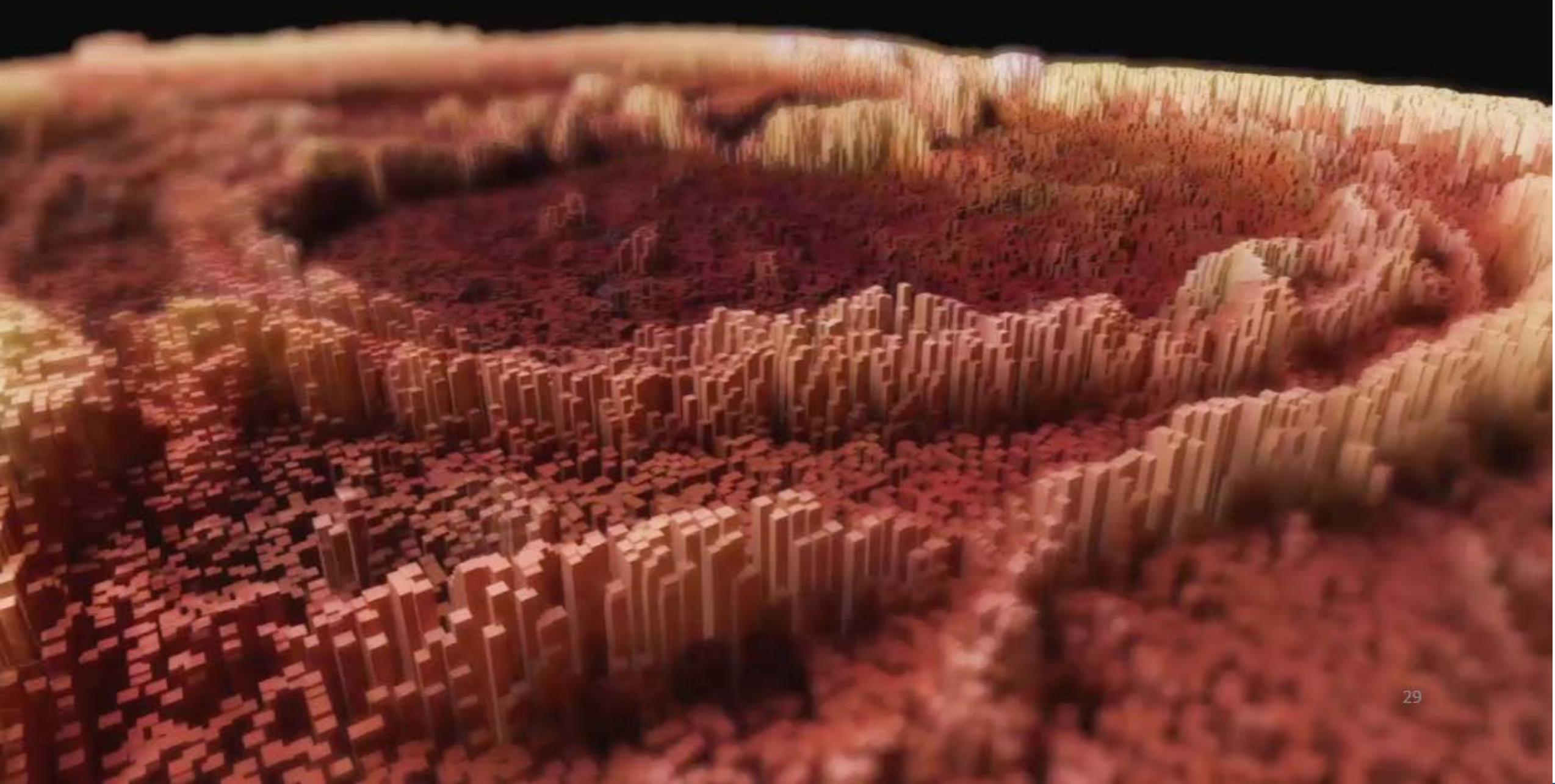
Articles

Content provided in part by:

1949
Full Participation
Journals

323
NIH Portfolio
Journals

4015
Selective Deposit
Journals

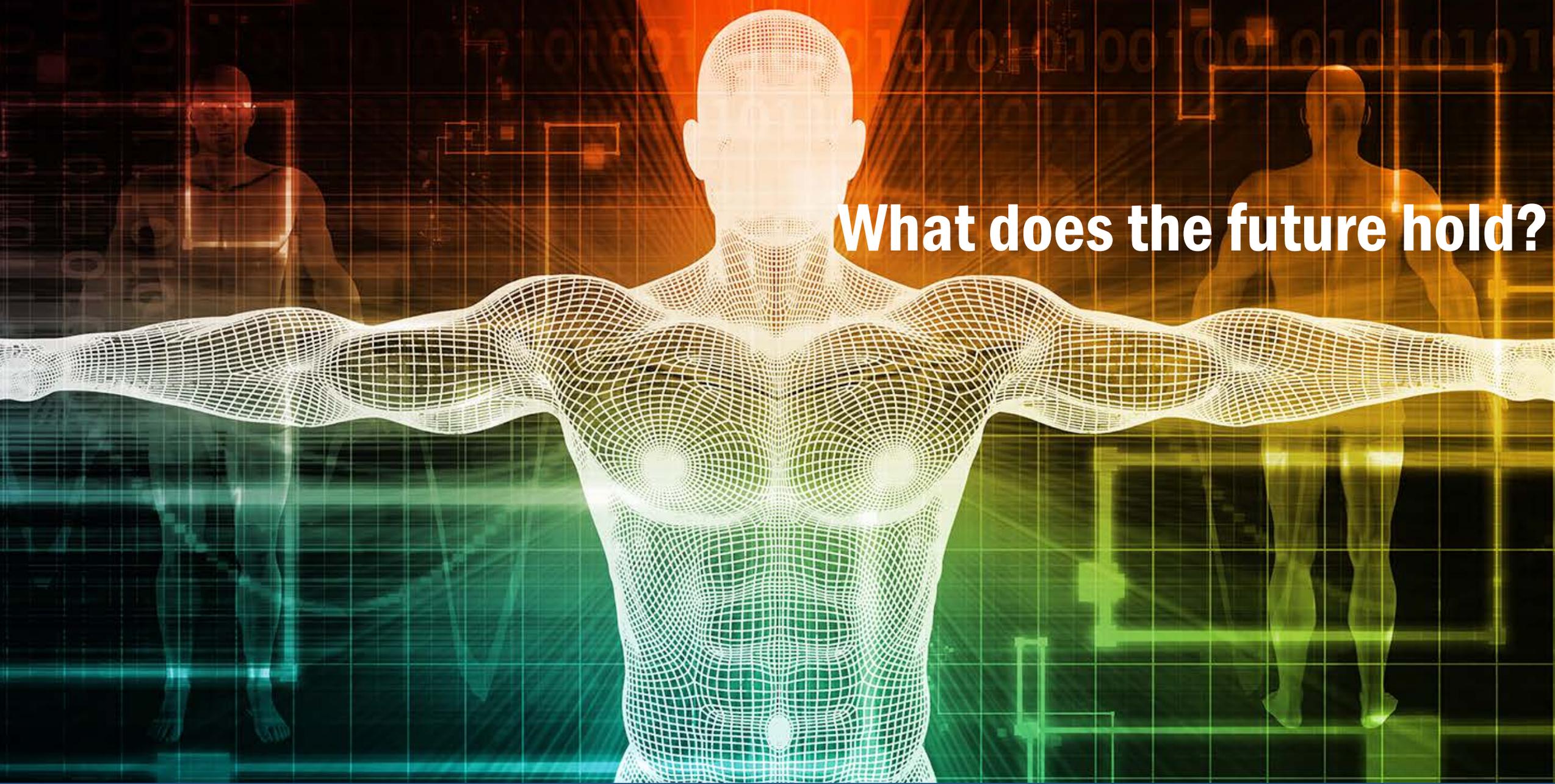


However...

The challenges to innovation identified by Matheson and Cooper persist – not *in spite of the* advances afforded by her vision of integrated library systems and academic information resources management networks



But because the information substrate is expanding from bibliographic elements to DATA!

The image features a central, glowing white wireframe human figure with arms outstretched. The background is a dark space filled with a grid of orange and green lines, suggesting a digital or data environment. In the background, there are faint silhouettes of human figures, one on the left and one on the right, appearing to be part of a virtual or augmented reality space. The overall aesthetic is high-tech and futuristic.

What does the future hold?

Look back to Matheson & Cooper!

- Content
- Context
- Computation
- Communication





Content: Data



Computation



Communication



**Partnering with experts,
professional and lay**

Accelerating discovery

Charting the future



Recommendation #2: 'Open Science'

NLM should
catalyze
research
promote
informatics

Recommendation #3: 'Data Science'

NLM should be the intellectual and programmatic epicenter for data science at NIH and stimulate its advancement throughout biomedical research and application

Strategic vision

- Data (data...data...data...data...data...data...)
- Grow our research programs
- Expand outreach to citizens

What could our future look like?

- Evolving data storage, communications, and computer security technologies
- Methods for generation, formalization, management, and sharing of knowledge resources
- Balance of basic and applied research in NLM's research support portfolio.
- Training for data scientists, data-informed investigators, data librarians
- Standards and policy development in the promotion of open science
- Partnership with other NIH components and agencies promoting best practices for data storage, access, discovery and analysis.

Developing the NLM Strategic Plan

September 2016-December 2017

- Board of Regents charge
- Stakeholder engagement
- NLM Staff visioning
- NIH IC Directors and stakeholder input
- Assessment of the physical assets
- Tying budget to goals
- Functional audit
- Division and program visitation

What is role of the National Library of Medicine in...

1. advancing data science, open science, and biomedical informatics
2. advancing biomedical discovery and translational science
3. supporting the public's health: clinical systems, public health systems and services, and personal health
4. building collections to support discovery and health in the 21st century

Cross-cutting themes

1. Standards
2. Partnerships
3. User Communities
4. User Engagement and Educational Outreach
5. International Engagement
6. Health Disparities
7. Infrastructure: computing
8. Infrastructure: physical plant
9. Research Needs and Funding
10. Workforce Development

NLM Strategic Planning Committee

- Daniel Masys, MD, co-chair, University of Washington
- Jill Taylor, PhD, co-chair, NY State Department of Health
- Robert Greenes, MD, PhD, Arizona State University
- Eric Horvitz, MD, PhD, Microsoft Research
- Sandra Martin, MSLS, Wayne State University
- James Olds, PhD, National Science Foundation
- Mike Heurta, PhD and Barbara Rapp, PhD, NLM staff

Come along – help us chart the future

Request for Information (RFI): Strategic Plan for the National Library of Medicine, National Institutes of Health

Notice Number: NOT-LM-17-002

Key Dates

Release Date: November 8, 2016

Response Date: January 9, 2017

Related Announcements

None

Issued by

National Library of Medicine ([NLM](#))

Purpose

The National Library of Medicine is undertaking a Strategic Planning Process and is soliciting input from its broad stakeholder community.

We recognize that many of our stakeholders generously replied to the 2015 RFI regarding future directions of NLM. Input provided in 2015 is already under consideration and need not be re-submitted. The 2015 RFI was issued by NIH on behalf of the NLM Working Group of the Advisory Committee to the NIH Director (ACD) to obtain input for their June 2015 report (<http://acd.od.nih.gov/reports/Report-NLM-06112015-ACD.pdf>) on a vision for the future of NLM in the context of NLM's leadership transition and emerging NIH data science priorities. The current RFI is issued to obtain public input on goals and priorities for NLM's next strategic plan.

Background

As defined in statute, the purpose of the NLM is 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." As the world's largest biomedical library, NLM presents a highly visible face of NIH across the United States and around the globe. Through its information systems, biomedical informatics and data science research portfolio, extensive training programs, and many partnerships, NLM plays an essential role in furthering fundamental research; catalyzing and supporting the translation of basic science into new treatments, products, and improved practice; and providing useful decision support for health professionals, the public health and emergency response workforce, and patients.

In the ten years since development of the NLM's last long range plan, there have been significant advances in biomedical informatics; major new initiatives at the NIH in data science, precision medicine, and open access to biomedical information; and changes in the environment and infrastructure of our country's health systems. NLM is committed to building a data infrastructure that will support the future of biomedical research.

Planning Themes

Read the RFI

Data Privacy Software Development

NIH is here to get a sense of what the community feels is on the horizon for data privacy software development.

- Interested in applied methods
- All data types & digital objects of interest:
 - Clinical, EHR, Image
 - Omics
 - Behavioral
 - Environmental Exposure
 - Health-relevant information broadly

Please meet with: Ed Ramos

ed.ramos@nih.gov

NLM Booth #404

Today (Tuesday) 11:00 am - 1:00 pm



Reaching me



<https://nlmdirector.nlm.nih.gov>



patti.brennan@nih.gov



@NLMdirector

