ACD Workgroup Proposal

Dr. Sean Mooney, CIT Director

NIH Center for Information Technology

6/13 ACD MEETING

Workgroup Proposal

- Introduction to CIT and IT at NIH
- Need for Guidance on Information Technology, Cyberinfrastructure, and Cybersecurity
- Proposed Workgroup Logistics
- Discussion

NIH Center for Information Technology

Today, all parts of the biomedical research life cycle contain data, computers, software, analytics, and their connectivity.

How can we cultivate the most advanced technology and the most capable workforce to tackle such a broad mission?

NIH Center for Information Technology

Digital NIH

- Strategic plan and visionary roadmap for our shared future in technology
- Developed through a committee and intended for 2023-2028
- A great starting point for the future of CIT

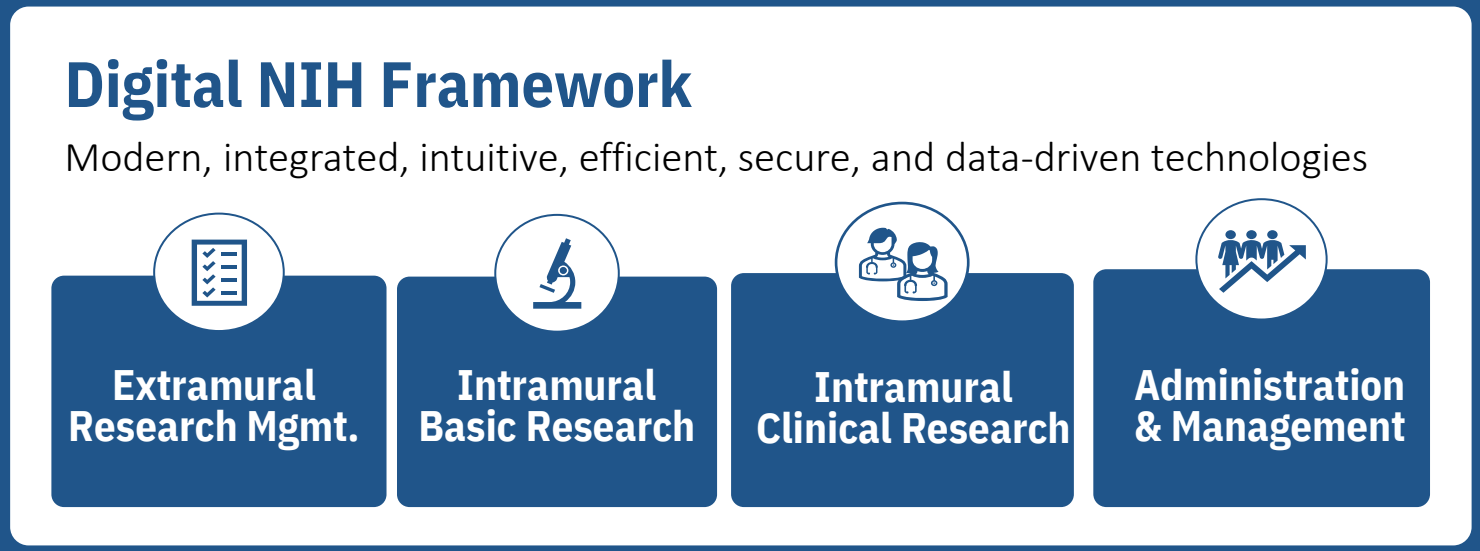
Co-Chairs	• Andrea Norris (OD/CIT)	• Patricia Flatley Brennan (NLM)	
Committee Members	• Jill Barnholtz Sloan (NCI) • Andy Baxevanis (NHGRI) • Raymond Dillon (OD) • Miles Fabian (NIGMS) • Inna Faenson (OER) • Gregory Farber (NIMH) • Greg Germino (NIDDK)	• Darla Hayes (OD) • Dyung Le (OD) • Janice Lee (NIDCR) • Colleen McGowan (ORS) • Elaine Ostrander (NHGRI) • Kate O'Sullivan (NHLBI) • Taunton Paine (OD)	• Kim Pruitt (NLM) • Rebecca Rosen (NICHD) • Jeff Shilling (NCI) • Xavier Soosai (CIT) • Michael Tartakovsky (NIAID) • Alastair Thompson (NHLBI)

NIH Center for Information Technology

6/13 ACD MEETING

Digital NIH

Innovation, technology, and computation for the future of NIH



NIH Center for Information Technology

6/13 ACD MEETING

Digital NIH

Innovation, technology, and computation for the future of NIH

Digital NIH Framework

Modern, integrated, intuitive, efficient, secure, and data-driven technologies

Extramural Research Mgmt.

Intramural Basic Research

Intramural Clinical Research

Administration & Management

Common **architecture** with well-defined standards to enable integration

NIH Center for Information Technology

6/13 ACD MEETING

Digital NIH

Innovation, technology, and computation for the future of NIH

Digital NIH Framework

Modern, integrated, intuitive, efficient, secure, and data-driven technologies

Extramural Research Mgmt.

Intramural Basic Research

Intramural Clinical Research

Administration & Management

Common **architecture** with well-defined standards to enable integration

Innovative, cutting-edge storage, analytics, and computational **infrastructure**

NIH Center for Information Technology

6/13 ACD MEETING

Digital NIH

Innovation, technology, and computation for the future of NIH

Digital NIH Framework

Modern, integrated, intuitive, efficient, secure, and data-driven technologies

Extramural Research Mgmt.

Intramural Basic Research

Intramural Clinical Research

Administration & Management

Common **architecture** with well-defined standards to enable integration

Innovative, cutting-edge storage, analytics, and computational **infrastructure**

Increased technically competent **workforce**

NIH Center for Information Technology

6/13 ACD MEETING

Digital NIH

Innovation, technology, and computation for the future of NIH

Digital NIH Framework

Modern, integrated, intuitive, efficient, secure, and data-driven technologies

- Common **architecture** with well-defined standards to enable integration
- Innovative, cutting-edge storage, analytics, and computational **infrastructure**
- Increased technically competent **workforce**
- **Technology** to support anywhere, anytime workplace of the future

NIH Center for Information Technology

6/13 ACD MEETING

Digital NIH

Innovation, technology, and computation for the future of NIH

Digital NIH Framework

Modern, integrated, intuitive, efficient, secure, and data-driven technologies

Extramural Research Mgmt.

Intramural Basic Research

Intramural Clinical Research

Administration & Management

Common **architecture** with well-defined standards to enable integration

Innovative, cutting-edge storage, analytics, and computational **infrastructure**

Increased technically competent **workforce**

Technology to support anywhere, anytime workplace of the future

Risk-based, embedded **cybersecurity protections**

NIH Center for Information Technology

The CIT View of IT @ NIH

Cloud Technologies

High-Performance Computing

Applications

Research Computing

Identity & Access Management

Network

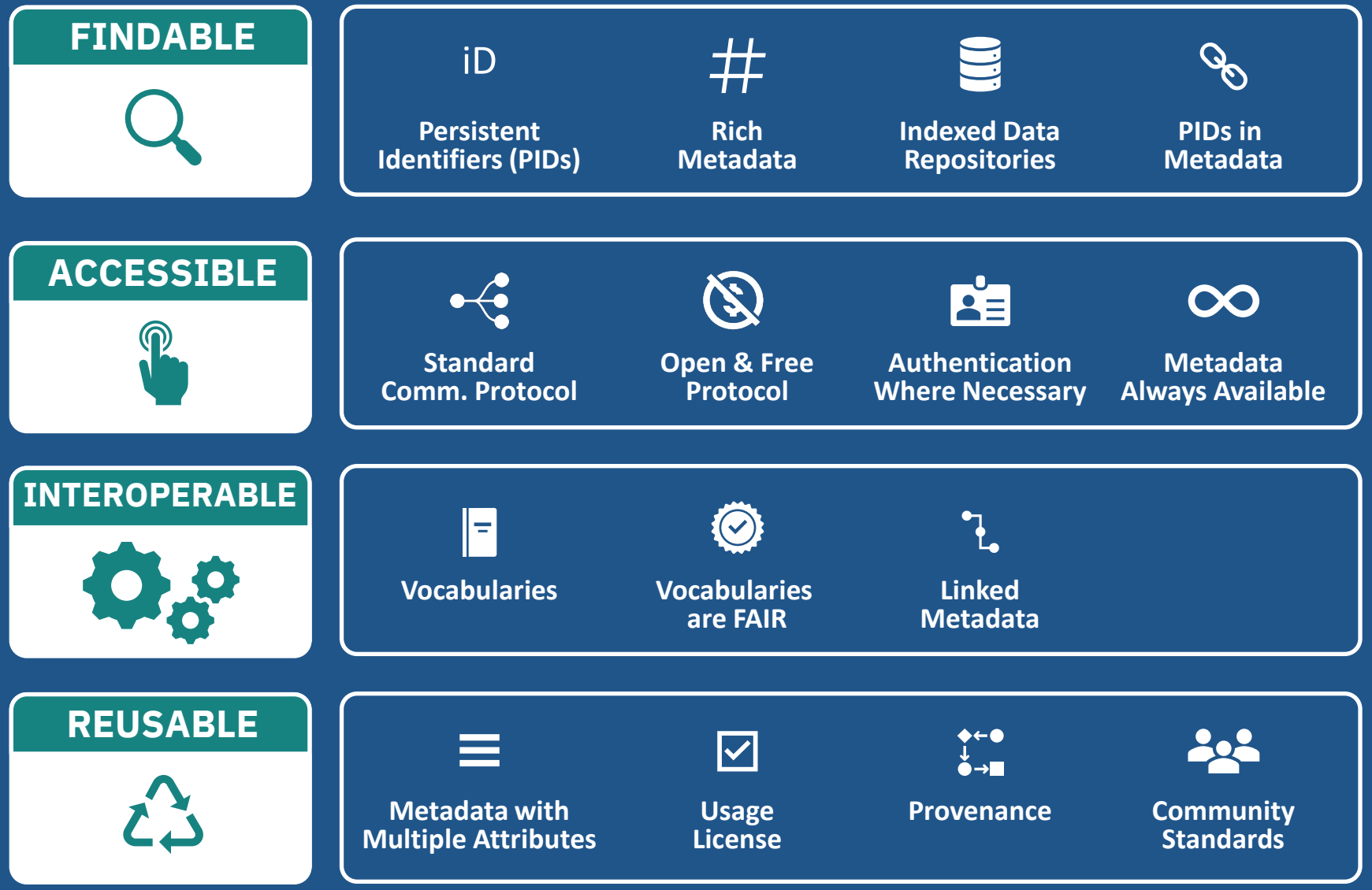
Cyberinfrastructure

Collaboration

Cybersecurity

NIH Center for Information Technology

We need **enterprise-level thinking** around FAIR data... and I need your help!



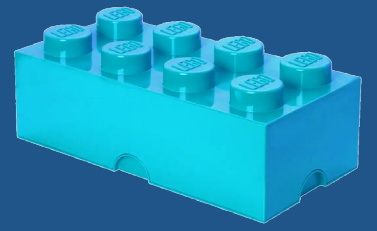
NIH Center for Information Technology

“New directions in science are launched by **new tools** much more often than by new concepts.”

Freeman Dyson, Theoretical Physicist

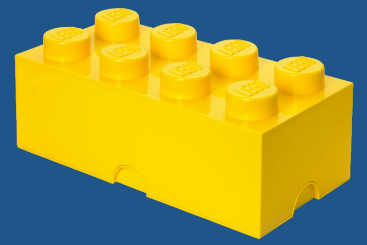
We need to support both the new data generating tools and the platforms to support their data

Bioinformatics Workflows



Components of Cloud Based Data Platforms

Specialized Computing: Quantum, GPUs, etc.



LLMs and Generative AI

NIH Center for Information Technology

We Need Intelligent Investments

- Standards for platform interoperability
- AI/data science platforms on controlled access data
- A balance of innovation, harmonization, and cybersecurity
- Efficiencies
- Improved IT governance

NIH Center for Information Technology

Mission-Based Approach to Cybersecurity

Security is a balance. It needs to be strong enough to keep us safe and light enough not to interfere with our work, like a hard hat.

More effective

MISSION

Open environment for research, collaboration, and clinical care

Stronger

SECURITY

Confidentiality, integrity, availability, and trust

WHAT WE PROTECT

Life & Safety
Personally Identifiable and/or Personal Health Information
Research Data Integrity
Financial Stewardship
Administration and Operations

NIH Center for Information Technology

Clinical Research Information Systems and Cyberinfrastructure for the NIH Campus

There are lots of **opportunities in clinical research IT and cyberinfrastructure**.

Supporting Clinical Trials

Data Resources and Data Governance

AI and Data Science Capabilities

Translation of Research Back to the EHR

6/13 ACD MEETING

Supporting Clinical Trials Within the NIH

This is an immediate opportunity to improve system-ness of the intramural NIH.

NIH Center for Information Technology

We have many axes to contemplate improving:

- Many **REDCap instances** across campus
- (To my knowledge) no **Enterprise Clinical Trials Management System (CTMS)** or EHR/CRO integration
- Few enterprise tools for **patient eligibility identification**
- Opportunities to improve **enterprise data systems** and **data governance** of intramural clinical research data
- Opportunities to **integrate research into the EHRs** including the Clinical Center and affiliates

NIH Center for Information Technology

Success Stories

A look at some (but not all!) of the things we do and how they benefit NIH

6/13 ACD MEETING

The NIH Network

- **4,300+** miles of network capacity
- **65,000+** devices connected
- **100 Gbps** network infrastructure
- **30,000** connections a day
- **150 TB** of Internet traffic a day

50
Provider-Grade Routers

NIH Center for Information Technology

3,000
Switches

9,000
Access Points

200
Firewalls

1,100
Uninterruptible Power Supplies

50
Servers

NIH Center for Information Technology

Enabling NIH-ers to Work from Anywhere

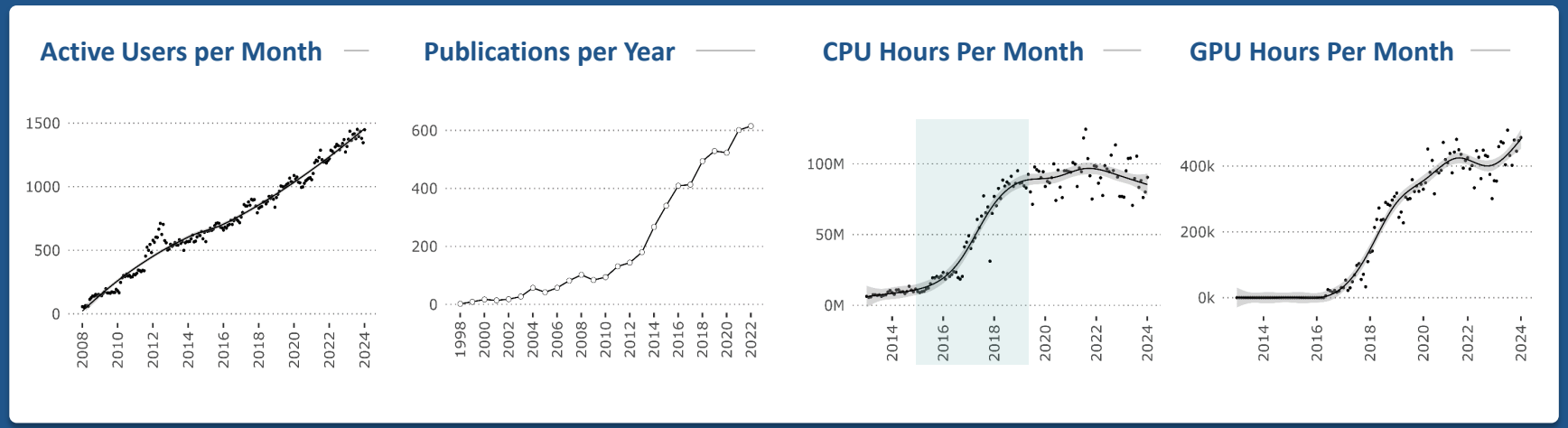
In FY 2023:

- More than **2.5M** virtual meetings
- More than **1B** emails sent and received
- More than **1,000,000** GB in OneDrive and SharePoint

Webex	Zoom	eFax	SharePoint Storage
Meetings Organized 176,856	**Meetings Organized** 466,096	**Total Pages Exchanged** 3,782,184	265 TB
Meetings Participants 1,129,868	**Meetings Participants** 3,226,990	**Total Transactions** 357,929	
Zoom Webinar	**Teams**	**Email Usage**	**OneDrive Storage**
Meetings Organized 1,268	**Meetings Organized** 1,808,768	**Total Email Sent** 107,699,158	777 TB
Meetings Participants 172,196	**Meetings Participants** 5,494,315	**Total Email Read** 633,206,257	
		Total Email Received 999,153,779	

NIH Center for Information Technology

Supporting Growing Demand for High-Performance Computing

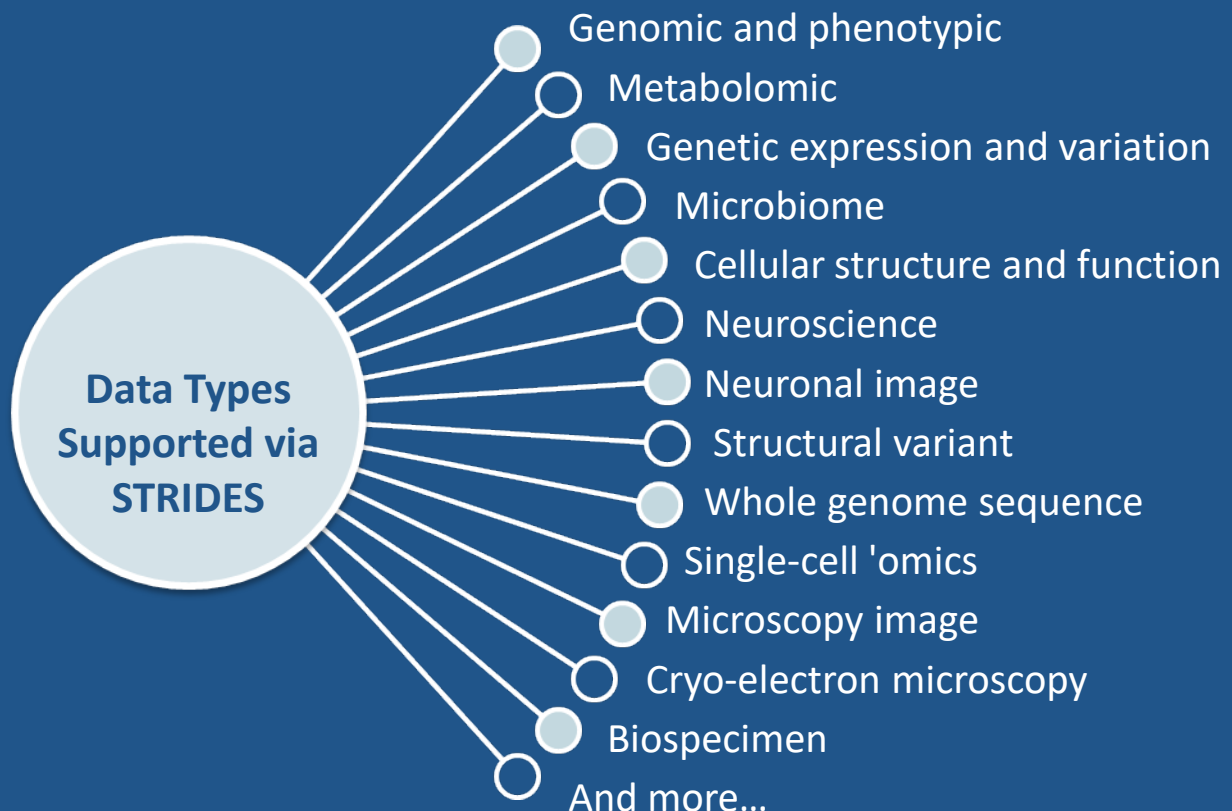


Biowulf, the NIH supercomputer, is:

- **105,000** compute cores
- **828** graphics processing units (GPUs)
- **50** PB storage
- **~ 1,000** biomedical applications supported
- **100** Gbps connection to the NIH network
- **5,050+** scientific papers supported
- Responsible for sequencing the last **8%** of the human genetic code

NIH Center for Information Technology

STRIDES | A partnership with cloud vendors to provide low-cost access



260+ PETABYTES OF DATA

606M+ COMPUTE HOURS

2,225+ RESEARCH PROGRAMS

$92M+ COST SAVINGS

5500+ PEOPLE TRAINED

as of February 29, 2024

NIH Center for Information Technology

Building for the Future

- In biomedical science, we spend 100s of millions each year unnecessarily **due to duplicative platforms we choose to support.**
- This is **expensive**. Not interoperable. Not sustainable.
- Having one platform for all data is obviously (I think) unrealistic.
- However, we can provide a common toolkit to make these platforms more efficient, interoperable, sustainable, and ***impactful.***

6/13 ACD MEETING

Digital Ecosystem and Cyberinfrastructure

Cyberinfrastructure should be like Legos:

reusable, interoperable, and open.

NIH Center for Information Technology

NIH Center for Information Technology

An Example: NIH Researcher Authentication Service (RAS)

A common task for computer resources is to log in to them. **Authentication.**

Once we log into a resource, we need to be authorized to see certain digital assets (data, web pages, pictures of cats). **Authorization.**

RAS solves this problem by enabling researchers to not reinvent the wheel and save both time and money using a common toolkit for logging in and accessing resources.

RAS is another impactful collaboration between ODSS and CIT.

RAS is a Lego that contributes to global biomedical cyberinfrastructure.

6/13 ACD MEETING

The People Building the Cyberinfrastructure Ecosystem of the NIH

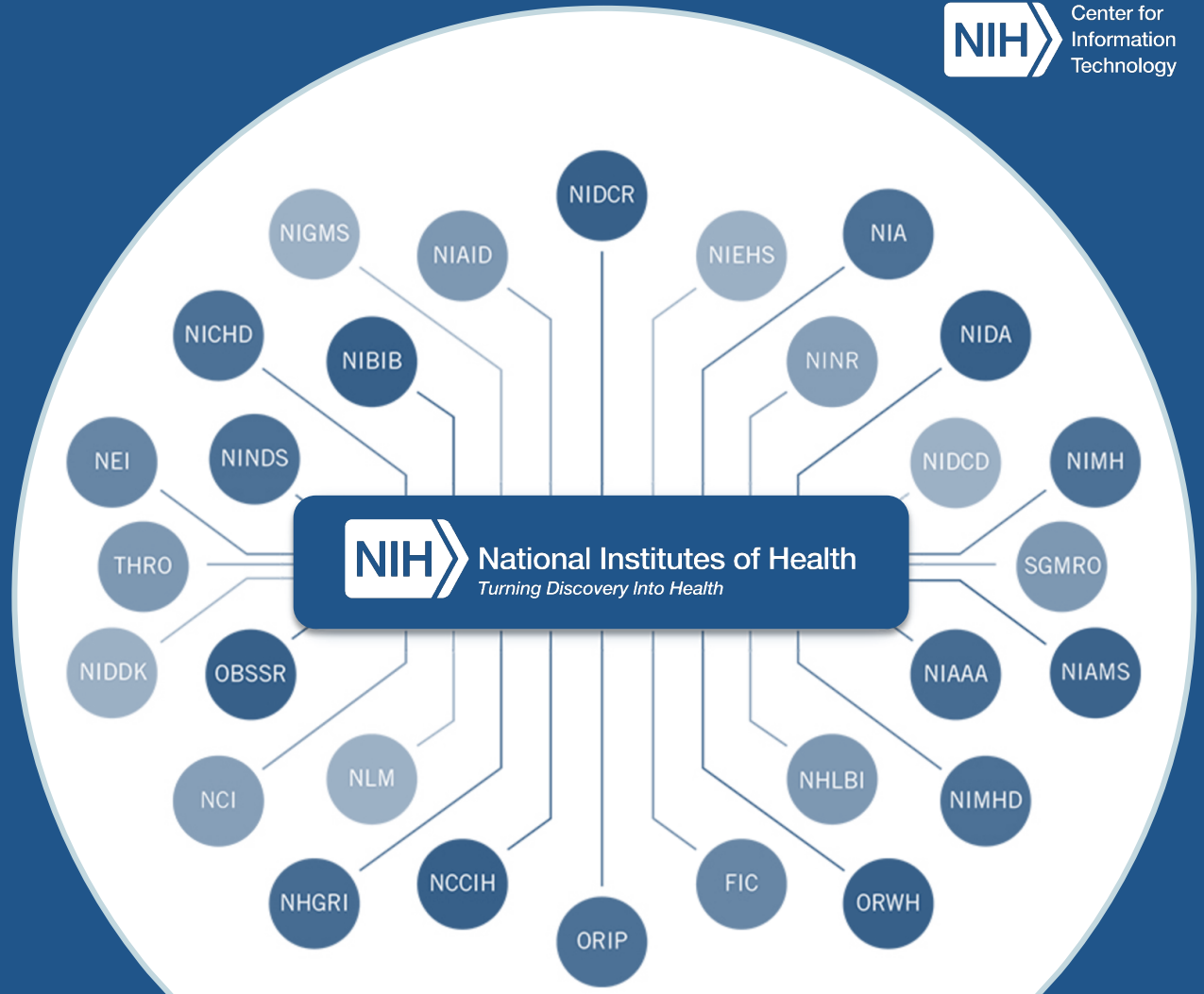
NIH Center for Information Technology

NIH National Institutes of Health
Office of the Chief Information Officer

NIH National Institutes of Health
Office of Data Science Strategy

NIH National Library of Medicine

NIH Center for Information Technology



NIH Drives Biomedicine and Contributes to IC-Specific Cyberinfrastructure.

NIH Center for Information Technology

AI Platforms Strategy

We have a heavy use of **foundation models** (e.g., ChatGPT, others) and generative AI.

There is a **community of practice** for AI at the NIH and many area-specific strategic planning efforts.

There are **many needs**, such as secure/private AI environments, workspaces to bring disparate data together, etc.

This is an important area for discussion and **advice is needed** based on the work group's experience and needs.

NIH Center for Information Technology

How You Can Help

We need a **roadmap** for Cyberinfrastructure and AI/data science platforms.

Thus, we need **your input** on our strategy and direction.

This will have a wide impact for **both intramural and extramural** and, I expect, will be welcomed.

Center for Information Technology

Today's Information Technology,
Tomorrow's Cyberinfrastructure

Thank You!

Charge to the ITCC Working Group

- Provide critical input on NIH IT and cybersecurity governance;
- Articulate high-priority areas for NIH investment in cyberinfrastructure, cybersecurity, AI, and data science;
- Validate or provide additional insight on proposed strategies to support new data generating tools and platforms;
- Evaluate established or proposed platforms for platform interoperability;

Charge to the ITCC Working Group (cont.)

- Identify gaps, challenges, and opportunities on issues related to national biomedical cyberinfrastructure;
- Evaluate and potentially expanding cloud-based tools to expand or enhance access to rich datasets and advanced computational infrastructure, tools, and services;
- Identify and promoting solutions to challenges in training, recruitment, and broadening the technical workforce in biomedical research; and
- Provide insight and feedback on risk-based cybersecurity protections.

Process, Deliverables, and Timeframe

- The ITCC Working Group of the ACD will:
  - Seek broad input from IT and scientific communities in public and private sectors;
  - Include perspectives from those with experience in information technology, cyberinfrastructure, and cybersecurity in a research or clinical environment;
  - Establish a workgroup of 1-2 internal members and 6-10 external members (external to NIH and the ACD);
  - Hold quarterly or bi-annual 90-minute meetings to provide feedback to efforts made on strategies and direction; and
  - Deliver an interim report within 12 months; a final report within 18-24 months.