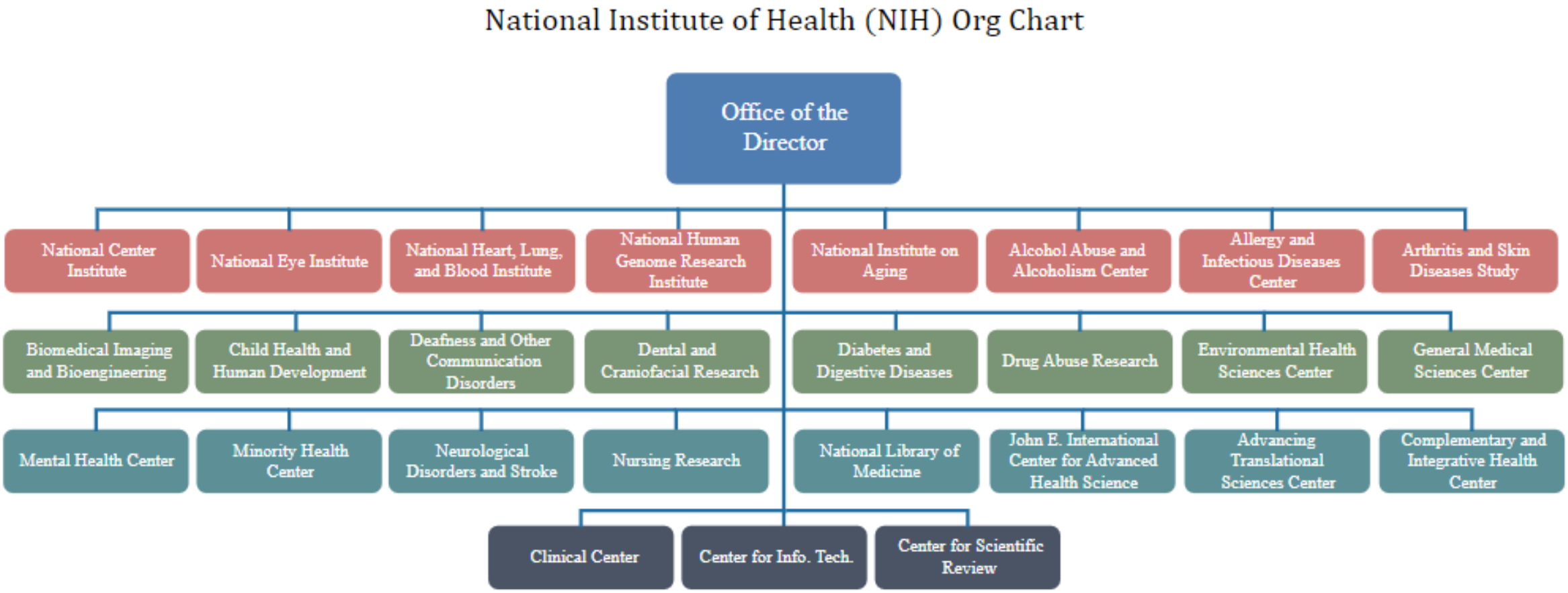
National Institutes of Health
Office of Data Science Strategy

Connecting Data, Enhancing Software, and Creating a Digital Data Health Ecosystem

Susan K. Gregurick, PhD

March 22, 2022

National Institutes of Health Institutes, Centers, and Offices



NIH Strategic Plan for Data Science

Support common infrastructure and architecture for specialized platforms.

Leverage industry resourced by adopting and adapting from other fields to use in biomedical research.

Enhance the biomedical data-science research workforce through improved programs and novel partnerships.

Enhance data sharing, access, and interoperability such that NIH-supported data resources are FAIR.

Ensure the security and confidentiality of data in accordance with NIH requirements and applicable law.

With community input, develop, promote, and refine data standards, including standardized data vocabularies and ontologies.

Improve the ability to capture, curate, validate, store, and analyze clinical data for biomedical research.

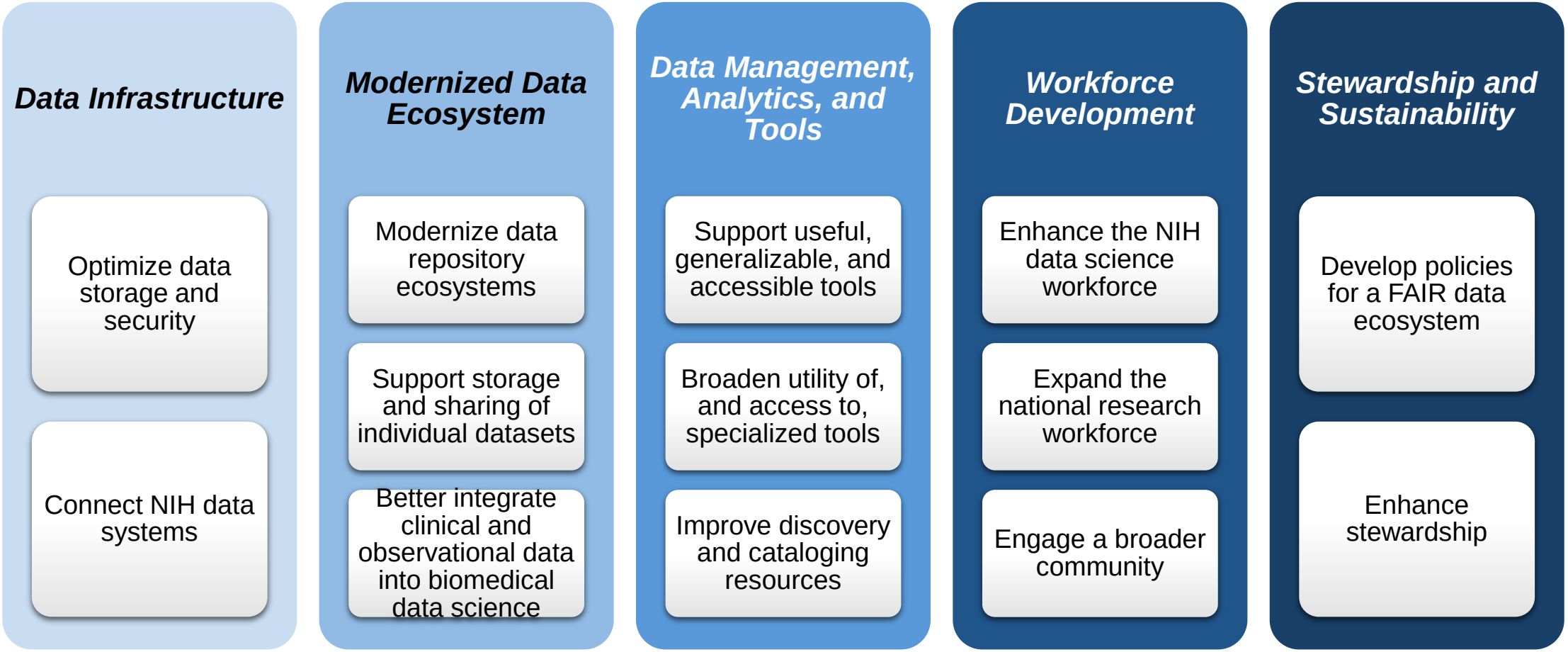
Office of Data Science Strategy

The NIH **Office of Data Science Strategy (ODSS)**, in the Office of the Director:

- Provides **leadership and coordination** on the strategic plan for data science.
- Develops and implement NIH's vision for a **modernized** and **integrated** biomedical data ecosystem.
- Enhances a **diverse and talented** data science workforce.
- **Builds strategic partnerships** to develop and disseminate advanced technologies and methods.

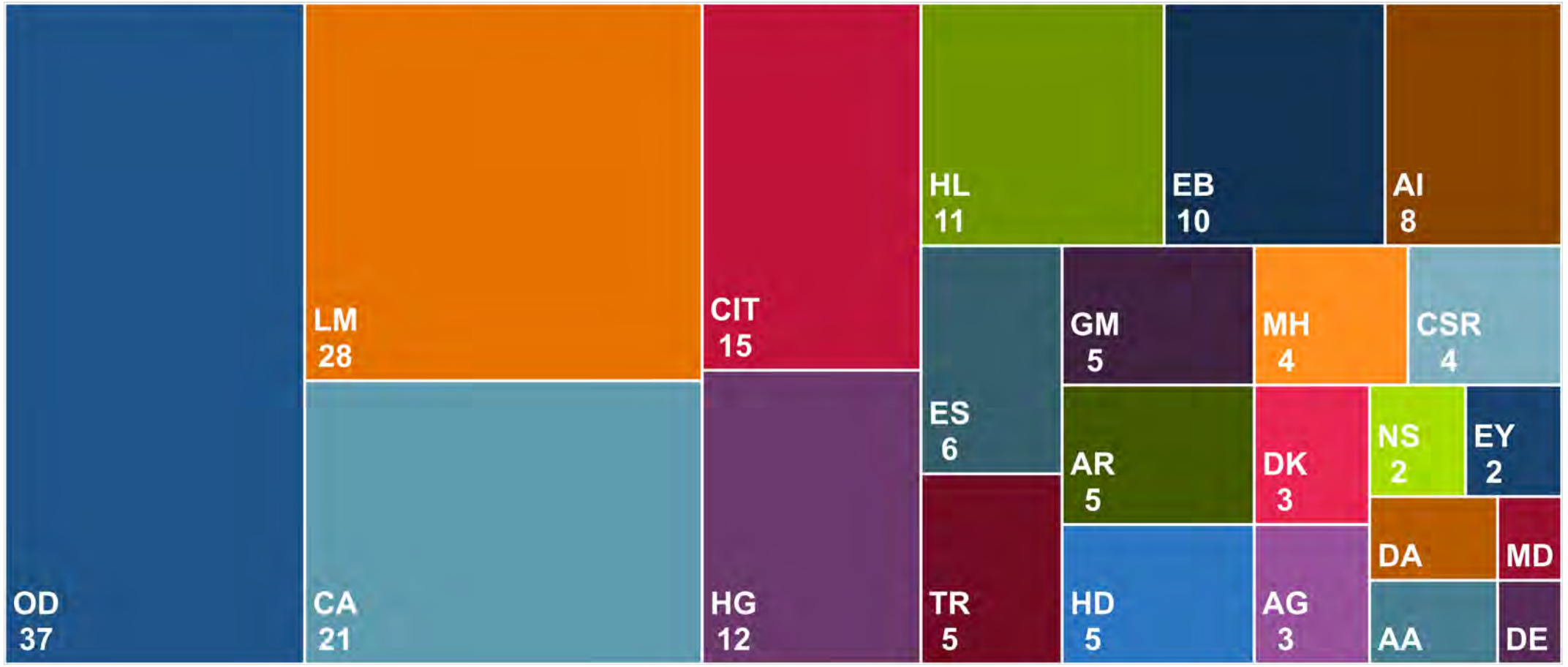
.

Strategic Plan for Data Science: Goals and Objectives



Catalyzing Data Science Across NIH

More than 190 NIH staff from 23 ICOs contributed to these activities



Agenda for Today's Talk

- **Artificial Intelligence/Machine Learning Consortium to Advance Health Equity and Researcher Diversity**
- **Prioritizing Ethics in AI**
- **Bridge2AI**
- **Smart and Connected Health**
- **Digital Twins**
- **The Science and Technology Research Infrastructure for Discovery, Experimentation, and Sustainability (STRIDES)**
- **Data Repository Ecosystem**
- **Workforce Development**

The Harvard Gazette

NATIONAL & WORLD AFFAIRS

For Native Americans, COVID-19 is 'the worst of both worlds at the same time'

The Skywalk over the Grand Canyon was closed by the Hualapai tribe to increase social distancing measures during the pandemic.

Stan Lawrimore/Unsplash

The Striking Racial Divide in How Covid-19 Has Hit Nursing Homes

Homes with a significant number of black and Latino residents have been twice as likely to be hit by the coronavirus as those where the population is overwhelmingly white.

Rural America Could Be the Region Hardest Hit by the COVID-19 Outbreak

Black Americans Face Alarming Rates of Coronavirus Infection in Some States

Data on race and the coronavirus is too limited to draw sweeping conclusions, experts say, but disparate rates of sickness — and death — have emerged in some places.

Many Who Need Testing For COVID-19 Fail To Get Access

April 3, 2020 · 5:00 AM ET

COVID-19 in Prisons and Jails in the United States

Laura Hawks, MD[1,2]; Steffie Woolhandler, MD, MPH[2,3]; Danny McCormick, MD, MPH[1,2]

» Author Affiliations | Article Information

JAMA Intern Med. Published online April 28, 2020. doi:10.1001/jamainternmed.2020.1856

Artificial Intelligence/Machine Learning Consortium to Advance Health Equity and Researcher Diversity (AIM-AHEAD)

Goals

- to enhance the **participation** and **representation** of researchers and communities currently underrepresented in the development of artificial intelligence and machine learning (AI/ML) models

- to address health disparities and inequities using AI/ML

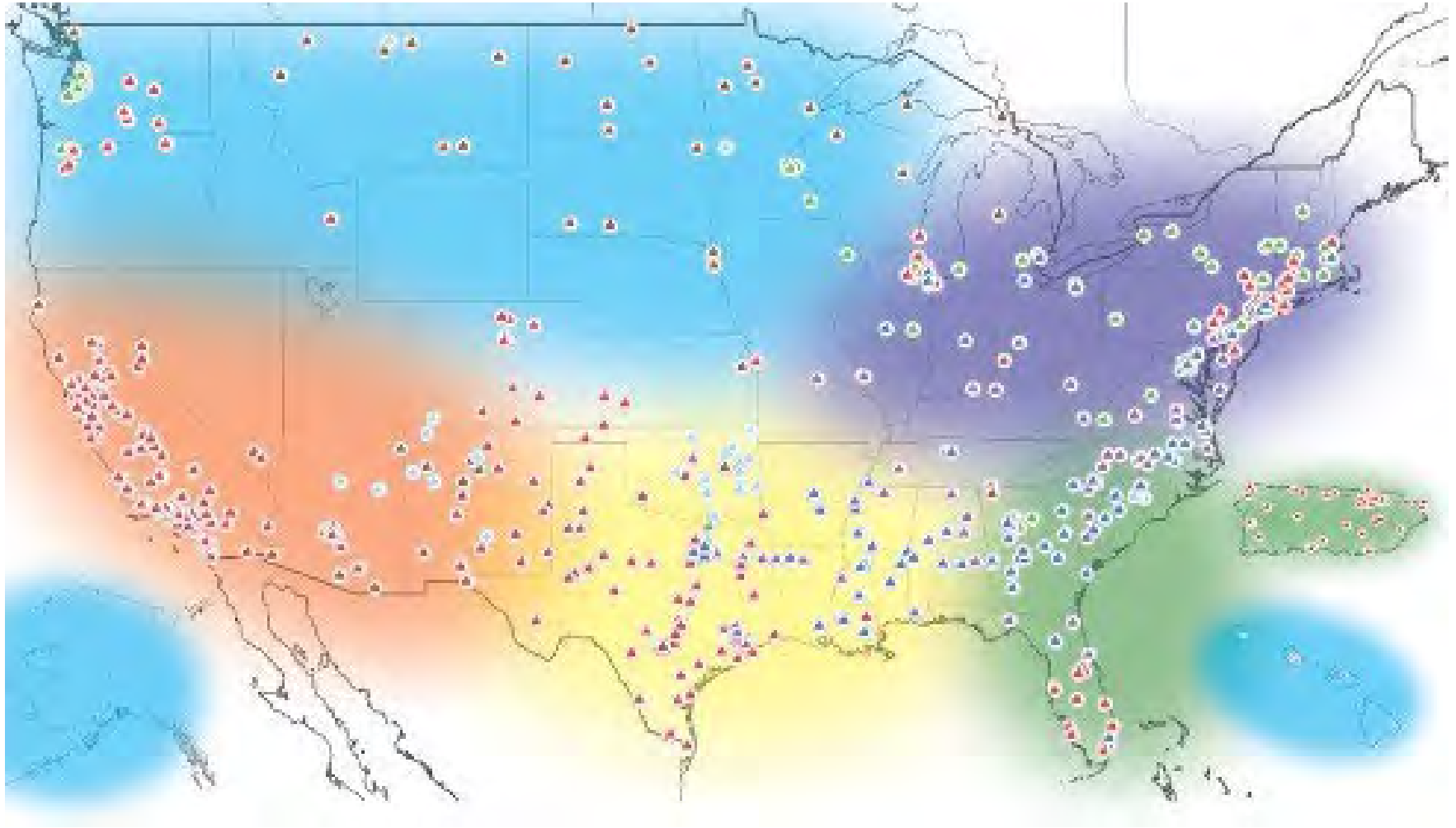
- to improve the capabilities of this emerging technology, beginning with the use of electronic health record (EHR) and extending to other diverse data

https://aim-ahead.net/
https://datascience.nih.gov/artificial-intelligence/aim-ahead

The AIM-AHEAD Program

Leadership/Admin Core:	Provide overall leadership and management including, recruiting consortium members and establishing trusted relationships with populations affected by health disparities.
Data Science Training Core:	Identify training needs and gaps, as well as develop training and workforce development resources to address AI/ML and health disparities research.
Data and Research Core:	Determine the AI/ML and health disparities research use cases that will drive the design of the data and computing infrastructure and associated data sharing and training offerings, as well as prepare and support data resources for AI/ML applications and health disparity research.
Infrastructure Core:	Engage with the largest, most diverse array of MSIs to assess institution needs for infrastructure that will provide inclusive capacity for AI/ML and health disparities research as well as design, develop, and implement appropriate data infrastructure(s).

The AIM-AHEAD Program, a Hub and Spoke Model



https://aim-ahead.net/

Leadership Core
University of North Texas Health Science Center in Fort Worth

Regional Hubs
Vanderbilt University Medical Center
University of Houston
University of North Texas Health Science Center in Fort Worth
University of Colorado-Anschutz Medical Center in Aurora
University of California, Los Angeles
Meharry Medical College in Nashville, Tennessee
Morehouse School of Medicine in Atlanta, Georgia
Johns Hopkins University in Baltimore, Maryland

Data Science Training Core
- Howard University in Washington, D.C.

Infrastructure Core

- National Alliance Against Disparities in Patient Health in Woodbridge, Virginia
- Harvard Medical School in Boston, Massachusetts
- Vanderbilt University Medical Center in Nashville, Tennessee

Data and Research Core
- OCHIN in Portland, Oregon

Next steps

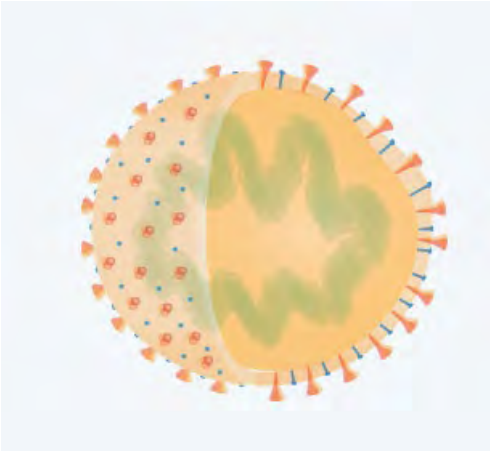
NIH has established the AIM-AHEAD Coordinating Center with the essential expertise in AI/ML and health disparities research, data science training, and data and computing infrastructure

Next steps:

- The AIM-AHEAD Coordinating Center will develop and implement a fair, equitable, and transparent process for recruiting consortium members

- Updates will be posted on the AIM-AHEAD Coordinating Center website https://aim-ahead.net/

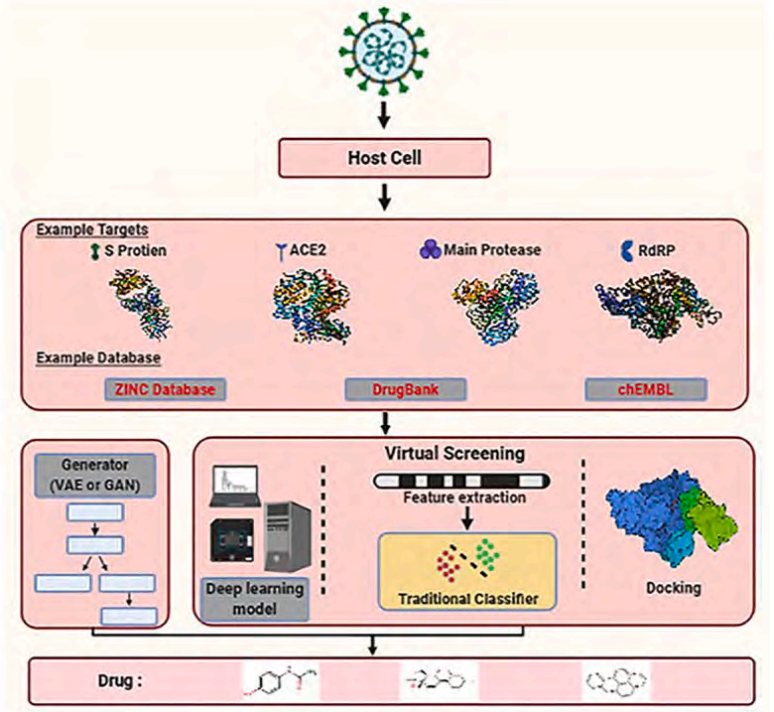
Accelerating us of Artificial Intelligence to address COVID-19



LitCovid is a curated literature hub for tracking up-to-date scientific information about the 2019 novel Coronavirus.

The White House partnered with AI research institutions to mine scientific literature to better understand Covid-19.

Leveraging AI to understand the structure of the novel coronavirus and to expedite drug discovery



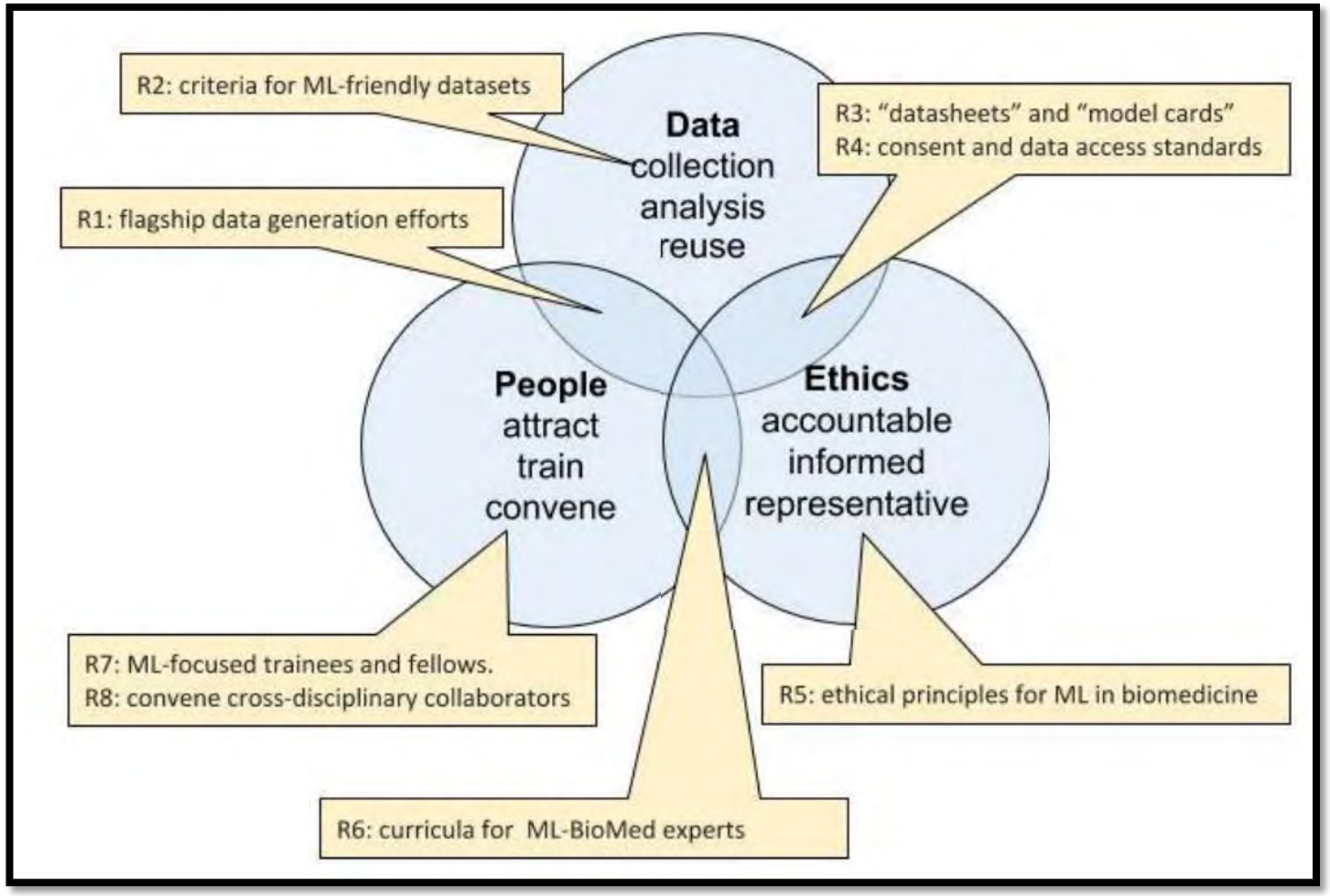
Artificial intelligence-driven drug repurposing and structural biology for SARS-CoV-2
K. Prasad, V. Kumar, Curr Res Pharmacol Drug Discov. 2021; 2: 100042. doi: 10.1016/j.crphar.2021.100042

Prioritizing Ethics in Biomed-AI

Rajpurkar, P., Chen, E., Banerjee, O. *et al.* AI in health and medicine. *Nat Med* **28,** 31–38 (2022). https://doi.org/10.1038/s41591-021-01614-0

Biomedical AI: Visions for an **ETHICAL** Future

NIH ACD AI Working Group Recommendations:



- Outlined opportunities to fuse AI/ML with exponential increase in biomedical data
- Ethics was identified as equally important to Data and People, reflecting the primary importance of infusing ethical thinking into AI/ML use in biomedical research

Collaboratively Envisioning AI and Ethics in Biomedical Research

The NIH is bringing together a diverse cross-section of scientists, social scientists, ethicists, advocates, legal scholars, communicators, and artists interested in the social implications of technology to

Forge new collaborations among these cross-disciplinary groups

Identify important areas of consideration at the intersection of artificial intelligence (AI) and machine learning (ML), biomedicine, and ethics.

Generate creative strategies to solve ethical dilemmas in biomedical AI/ML

Collaboratively Envisioning AI and Ethics in Biomedical Research

Micro Lab #1
December 15th, 2021

Engaged participants to provide collaborative expertise, explore, and assess the opportunities and challenges

Micro Lab #2
January 12th, 2022

Gathered a diverse group of stakeholders to map the current landscape of expertise needed

Micro Lab #3
January 26th, 2022

Welcomed an expanded community from cross-disciplinary backgrounds to pose exciting, innovative, and far-reaching research questions and imagine the future opportunities and challenges of AI/ML ethics.

Innovation Lab: A Data Ecosystems Approach to Ethical AI for Biomedical and Behavioral Research
March 14-18th, 2022

Provided an intensive, interactive, and idea-generating workshop designed to forge new collaborations among cross-disciplinary participants challenged to develop social and technical approaches to defining and implementing ethics across the AI data ecosystem.

Collaborations to Advance Ethical Use of AI/ML

New in FY22: NOT-OD-22-065

Advancing the Ethical Development and Use of AI/ML in Biomedical and Behavioral Sciences

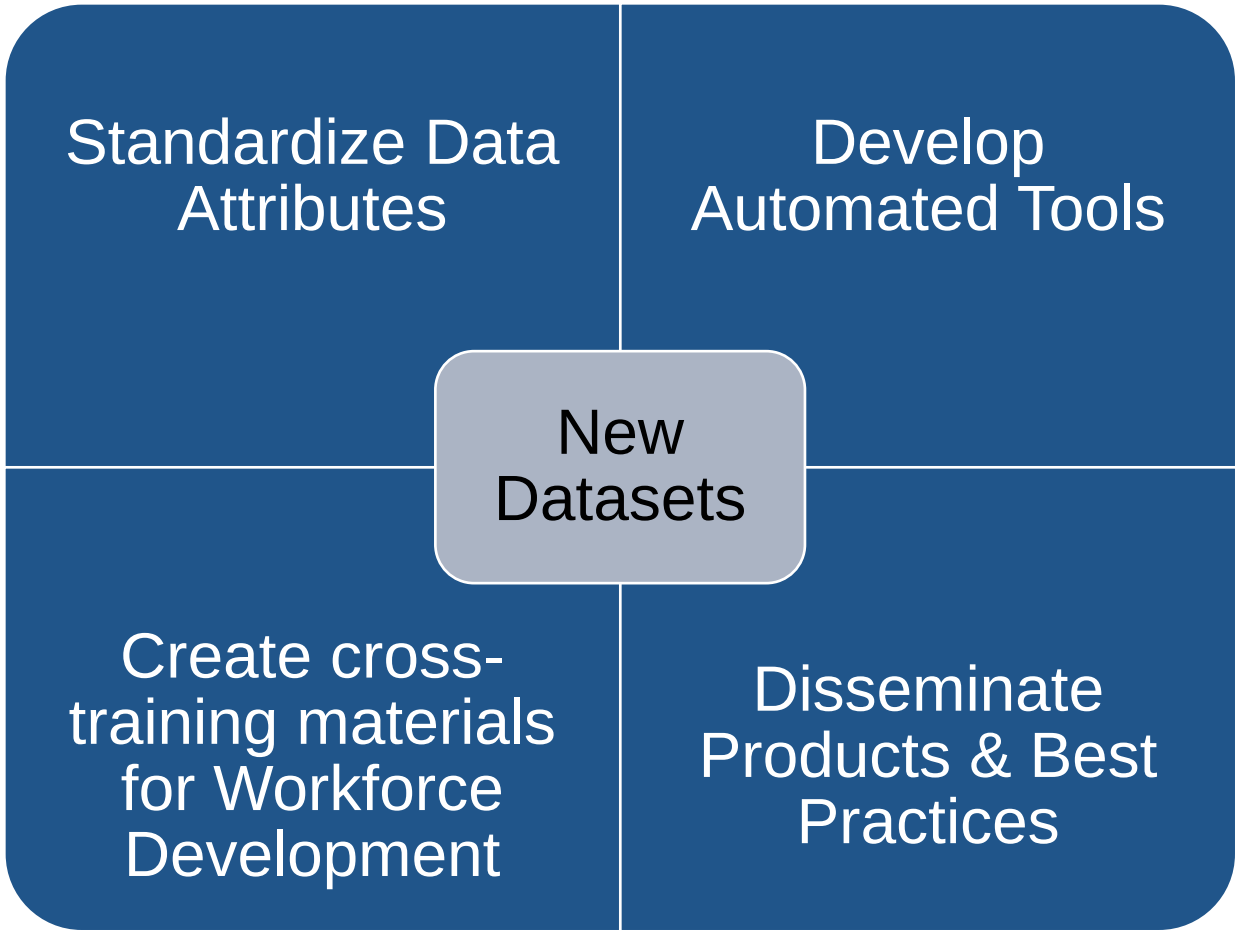
ODSS will support collaborations that bring together expertise in ethics, biomedicine, data collection, and AI/ML to advance the understanding, tools, metrics, and practices for the ethical development and use of AI/ML in biomedical and behavioral sciences.

These collaborations are intended to generate **new understanding, practices, tools, techniques, metrics, or resources that will aid *others*** in making ethical decisions throughout the development and use of AI/ML, including the collection and generation of data as well as the reuse of data and models by others. Research products developed under this NOSI will be shared and made broadly reusable.

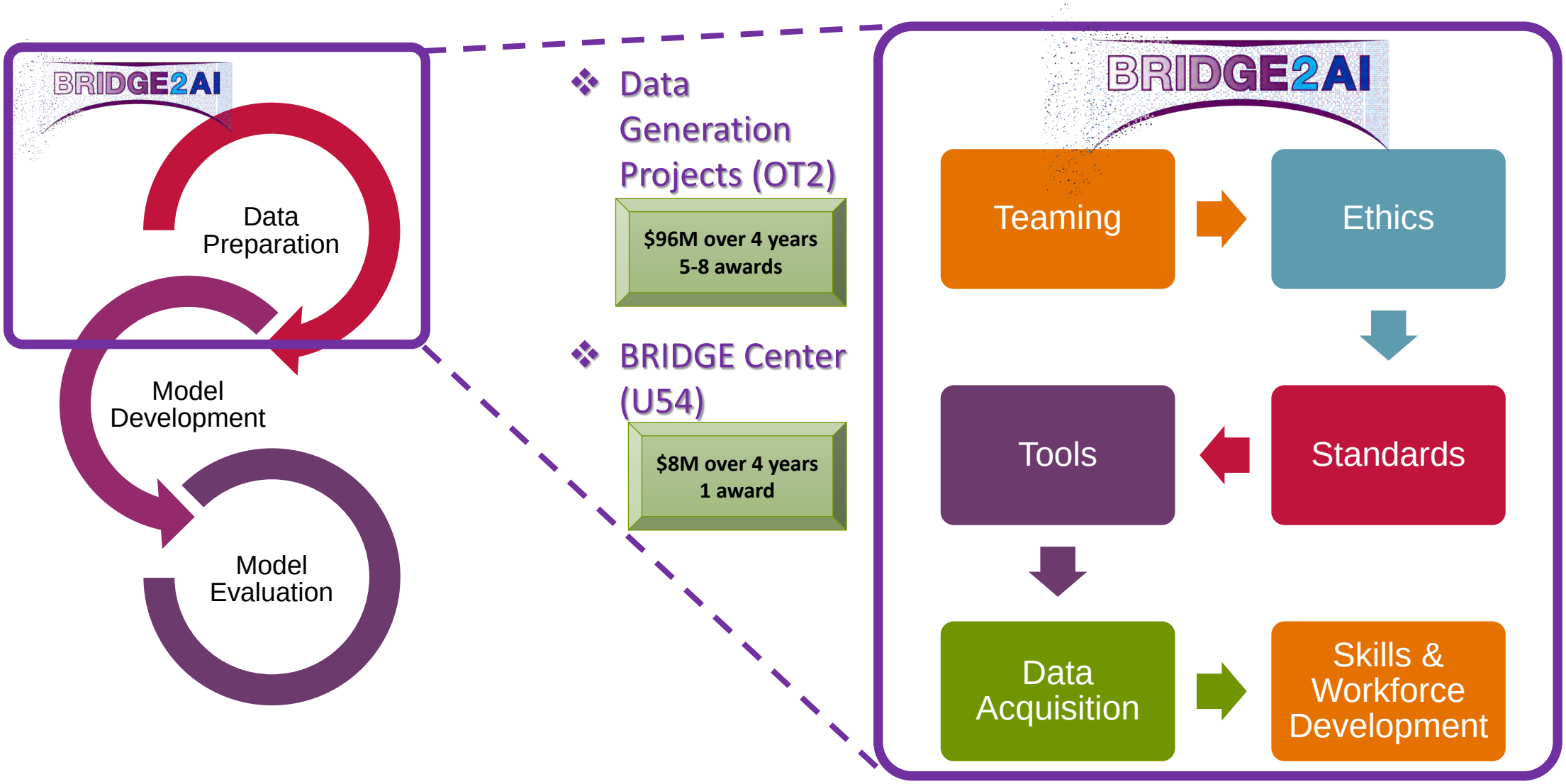
Due April 01, 2022

Bridge2AI

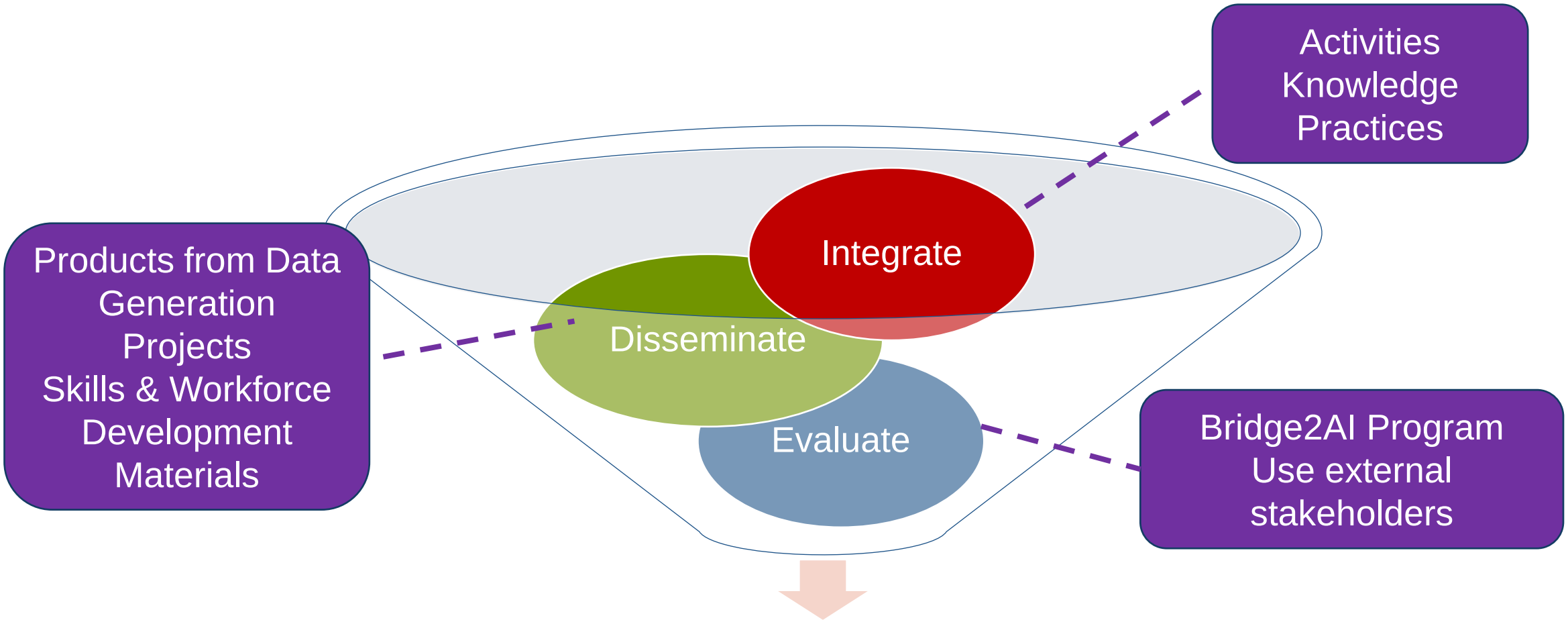
- Use biomedical and behavioral research grand challenges to generate **flagship data sets**
- Emphasize **ethical** best practices
- **Prepare** AI/ML-friendly data
- Promote **diverse teams**



Preparing the Data



BRIDGE Center



Best Practices for AI/ML in Biomedical and Behavioral Research

Who needs to come together?

- Biomedical Scientists
- Behavioral Scientists
- Social Scientists
- Clinicians
- Economists
- Ethicists
- Philosophers
- Anthropologists

- Diverse career stages
- Diverse social and cultural backgrounds
- Academic, industrial, and technical backgrounds
- Diverse communities and institutions

- AI/ML Experts
- Statisticians
- Computer Scientists
- Data Scientists
- Mathematicians
- Bio-informaticists
- Engineers
- Team Science Experts…

What will Bridge2AI Produce?

- Novel, complete, trustworthy datasets - ethically sourced, following FAIR principles, motivated by biomedical and behavioral grand challenges
- Tools to accelerate the creation of data sets for AI/ML analysis (intelligent annotators, metadata-filling instruments)
- Community evaluation of datasets -- **culture change** to embrace data preparation -- for AI/ML analysis
- Interdisciplinary AI/ML-Biomedical and Behavioral Research Community

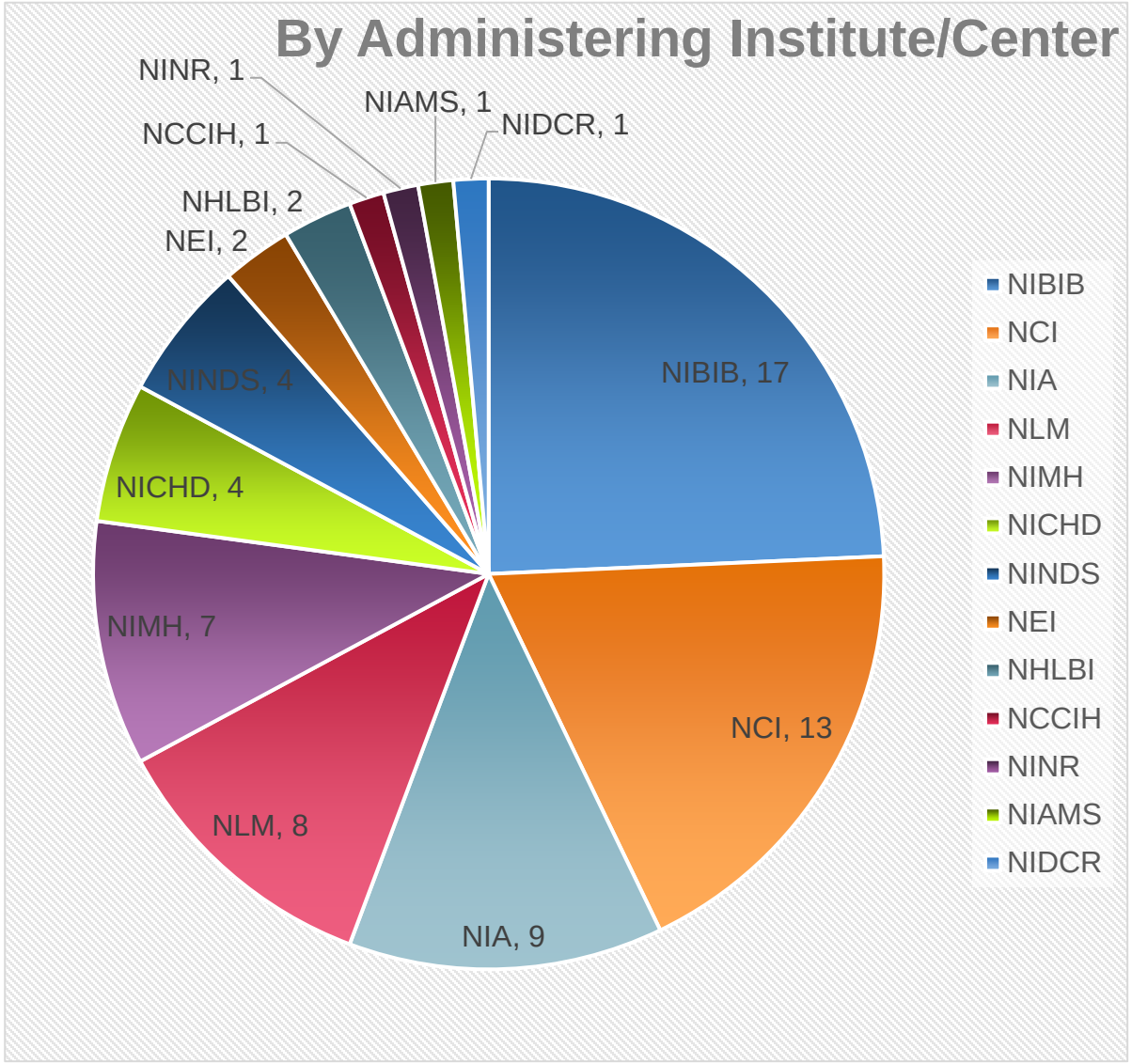

Software as an Important Part of a Data Ecosystem

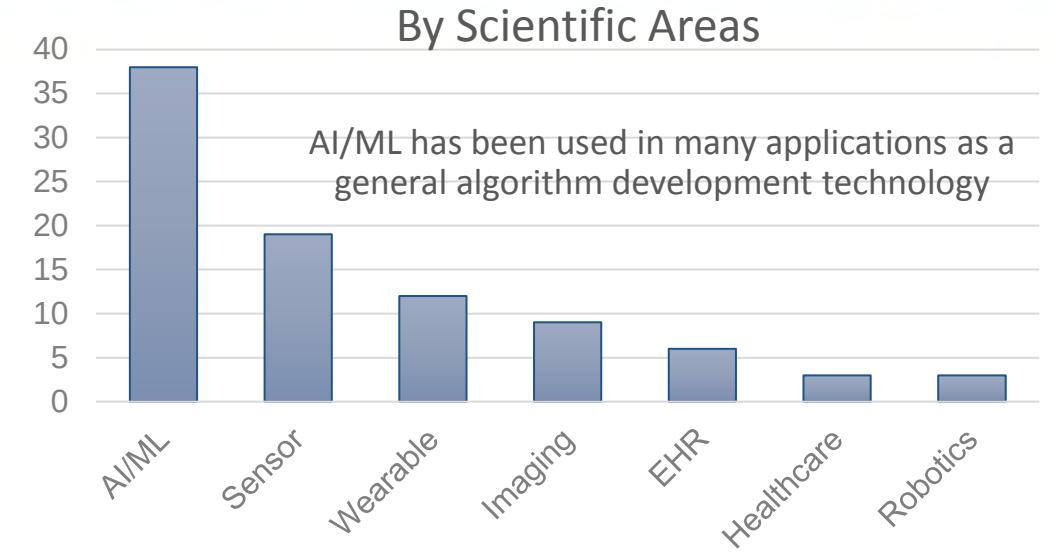
Smart and Connected Health (SCH)

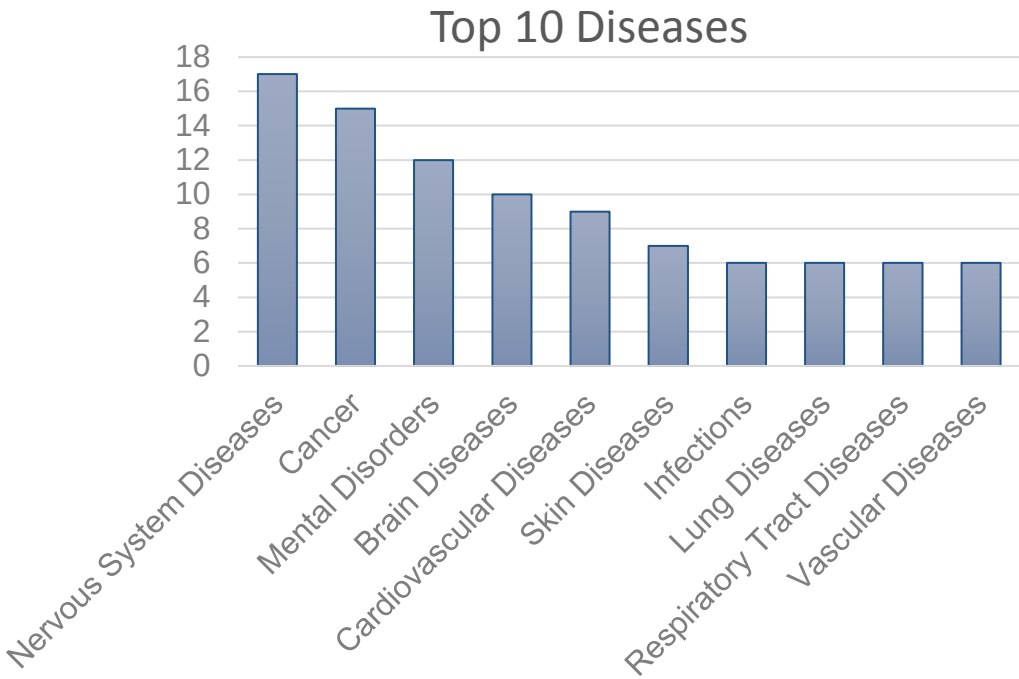
Accelerate innovations in computer and information science and engineering to support the transformation of health and medicine

NOT-OD-21-011

Smart and Connected Health - NIH Funded Awards



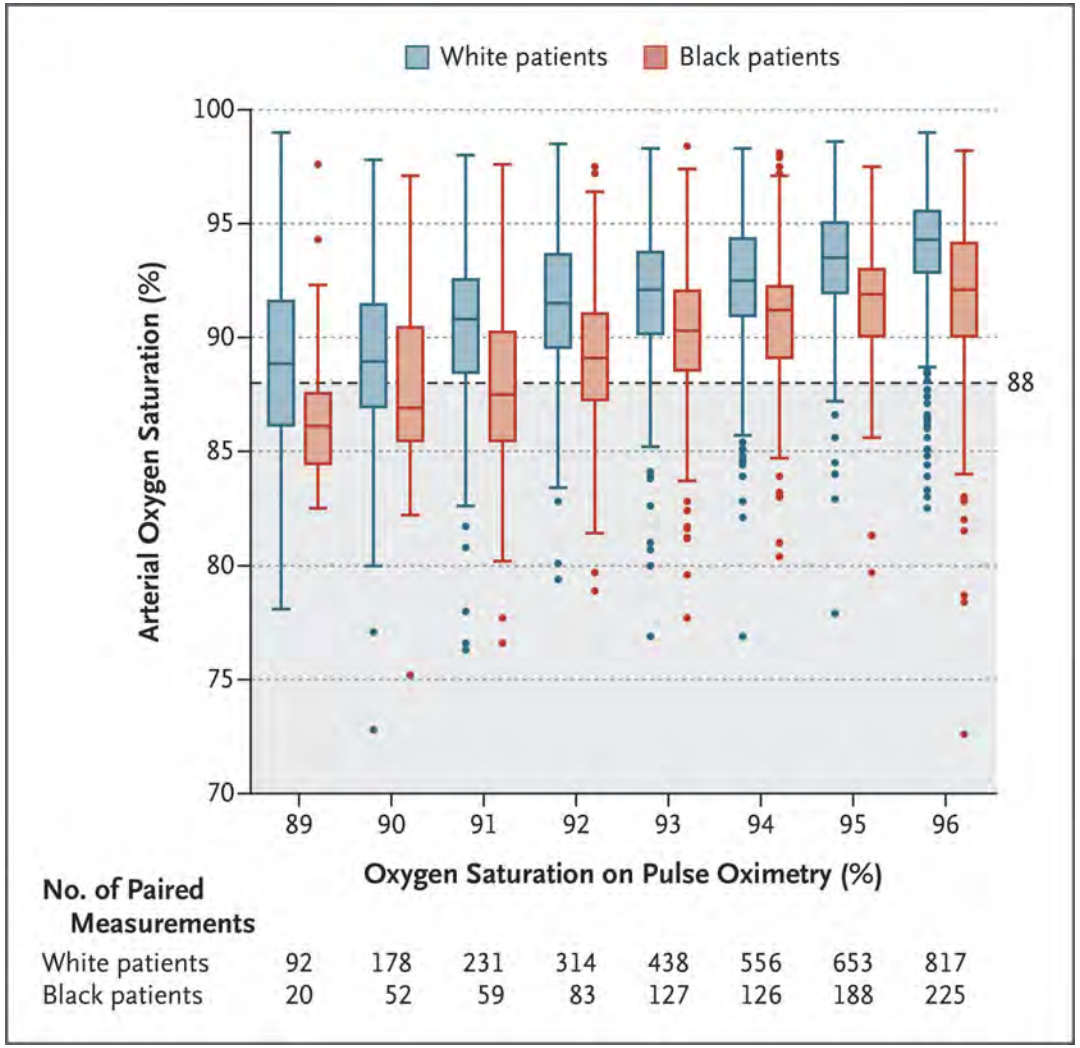




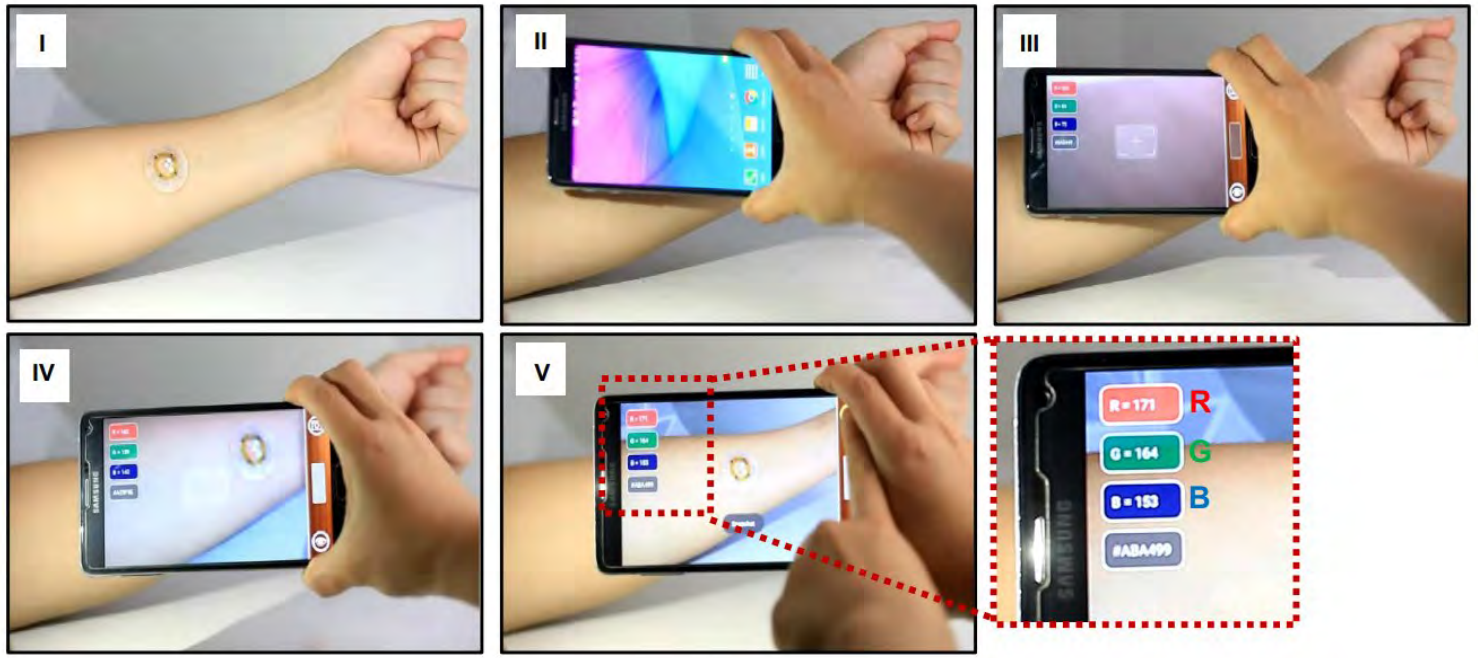
Racial Bias in Pulse Oximetry Measurement

Leveraging Clinical Time Series to Learn Optimal Treatment of Acute Dyspnea, EHR data analysis

- Published on December 17, 2020, N Engl J Med 2020; 383:2477-2478, DOI: 10.1056/NEJMc2029240
- **Conclusion:** Black patients had nearly three times the frequency of occult hypoxemia that was not detected by pulse oximetry as White patients because original pulse oximetry was developed without considering racial diversity.



Research Driven Innovation and Invention



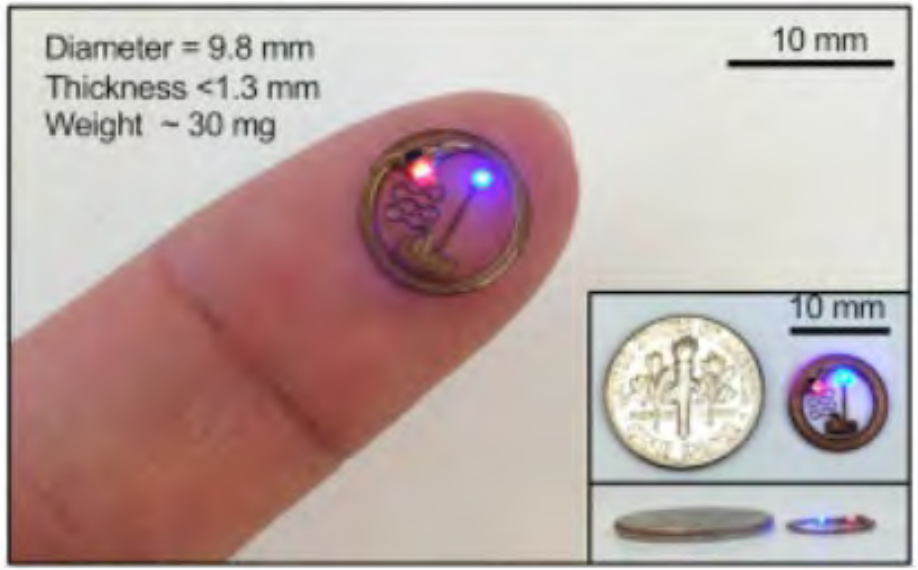
Developing soft, wearable microfluidic device for the capture, storage, and colorimetric sensing of sweat

Interface Monitoring System to Promote Residual Limb Health for Sensor Development

•Published in November 2016, Science Translational Medicine 8(366):366ra165-366ra165, DOI:10.1126/scitranslmed.aaf2593

•A soft, wearable microfluidic device to obtain quantitative values for sweat rate, total sweat loss, pH, and concentration of chloride and lactate.

Research Driven Innovation and Invention



Picture of a completed device

Images of a recovered mouse from surgery

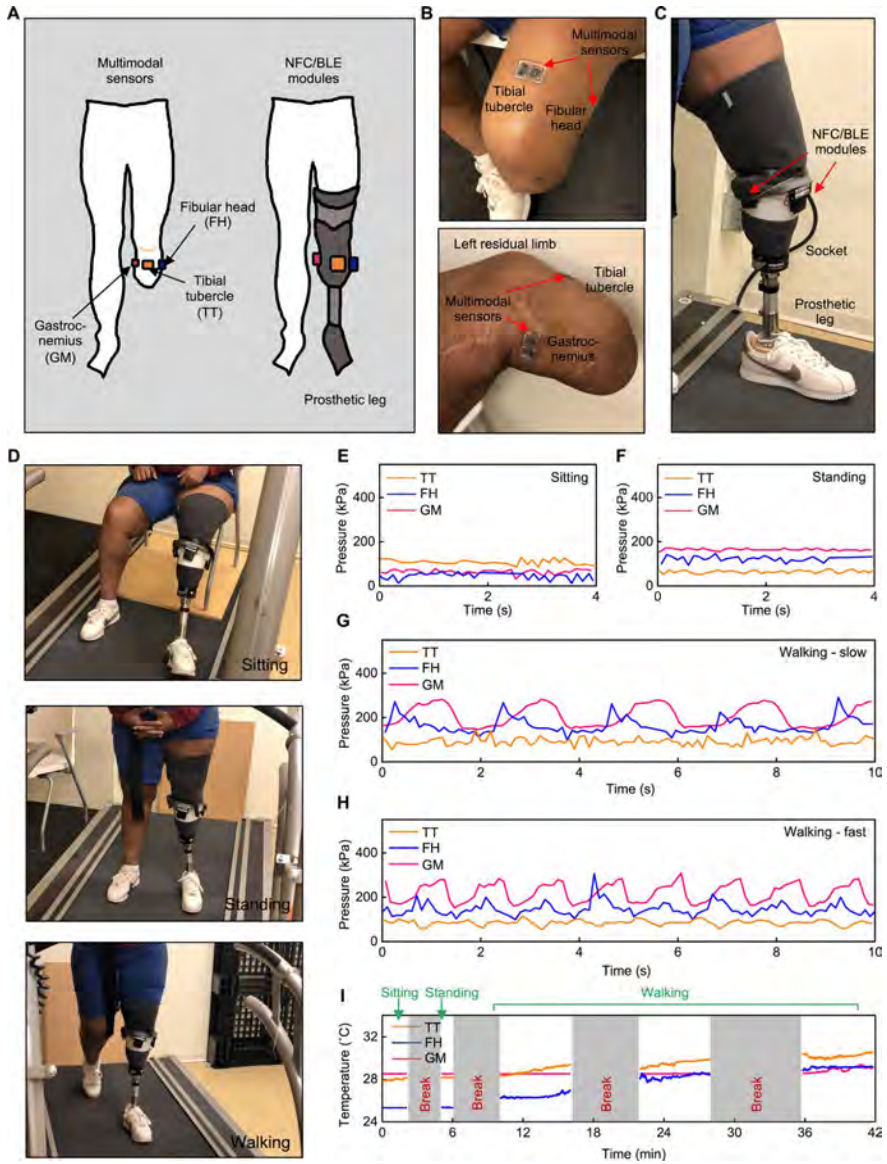
Flexible Near-Field Wireless Optoelectronics as Subdermal Implants for Broad Applications in Optogenetics

Interface Monitoring System to Promote Residual Limb Health" Sensor Development

- Published in Neuron Volume 93, Issue 3, 8 February 2017, Pages 509-521.e3

- **Results:** Designs and Operational Features of Thin, Flexible, Millimeter-Scale Devices for Wireless, Programmed Delivery of Light into Biological Tissues for Optogenetic Experiments.

Research Driven Innovation and Invention



Wireless sensors for continuous, multimodal measurements at the skin interface with lower limb prostheses

- Published in SCIENCE TRANSLATIONAL MEDICINE, 16 Dec 2020 Vol 12, Issue 574, DOI: 10.1126/scitranslmed.abc4327.

- Kwak and colleagues developed pressure and temperature sensors to monitor the interface between a prosthesis and residual limb.

Enhancing Biomedical Software for Greater Reuse

Developed best practices for open software development & sharing

Improving reuse and effectiveness of NIH-developed software for open science

Encouraged new collaborations between biomedical and clinical scientists and software engineers

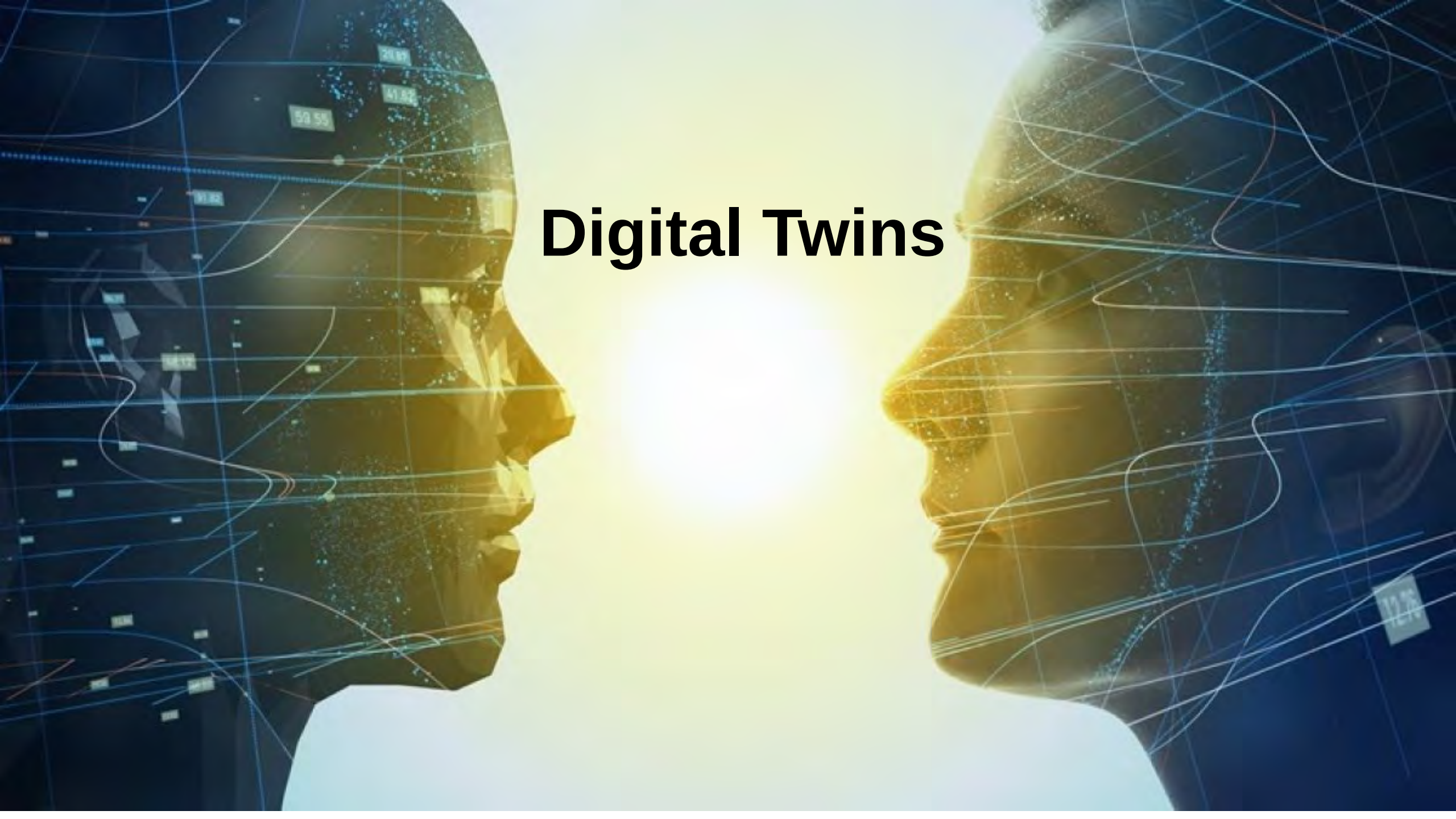
Notice of Special Interest (NOSI): Administrative Supplements to Support Enhancement of Software Tools for Open Science (NOT-OD-22-068) **Due April 13!**

Outcomes from FY20 Awards

- **28 grants** were awarded software engineering funds and collectively produced **120 scientific publications.**
- A Texas grantee ported advanced biomolecular modeling software to the Azure cloud to study COVID. Their community outreach includes a dynamic "graphic novel" presented on Twitter for LatinXChem.

42 New FY21 Awards

NCATS | NIA | NIBIB | NIDA | NIAMS | NIMH | NIDCD | NIDCR | NIAID | NIEHS | NIGMS | NHLBI | NCI | NINDS | NICHD | NLM | NIDDK


Digital Twins

Uses of Digital Twins

Digital twins are currently being pursued across domains

- **Structural digital twins** have shown promise in virtual health monitoring, certification, and predictive maintenance.[1]
- In **healthcare and biomedical sciences**, digital twins of human beings/patients may advance medical assessment, diagnosis, personalized treatment and clinical practice.[2]
- Digital twins of individual **students** offer a path to personalized education.[3]
- **Smart cities** enabled by digital twins and Internet of Things (IoT) devices may revolutionize urban planning, resource allocation, sustainability and traffic optimization.[4]
- Digital twins of **earth's atmosphere** may enhance weather and climate predictions.[5]
- Digital twins of **products and complex production systems** may enable increases in productivity and efficiency, and allow companies to make predictions about future performance, improve operations and productivity, and reduce the risk of unplanned downtime.[6]

[1] Bergthor Björnsson et al., "Digital twins to personalize medicine," *Genome medicine* 12, no. 1 (2020).
[2] Hernandez-Boussard et al., “Digital twins for predictive oncology will be a paradigm shift for precision cancer care,” *Nature Medicine* (2021).
[3] Witold Kinsner, "Digital Twins for Personalized Education and Lifelong Learning" (paper presented at the 2021 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), 2021).
[4] Maryam Farsi et al., *Digital twin technologies and smart cities* (Springer, 2020).
[5] Paul Voosen, "Europe builds ‘digital twin’ of Earth to hone climate forecasts," (American Association for the Advancement of Science, 2020).
[6] Fei Tao et al., "Digital twin in industry: State-of-the-art," *IEEE Transactions on Industrial Informatics* 15, no. 4 (2018).

Critical Research Opportunities

- Handling of data, models, and decisions that cross **multiple scales**
- **Constraint** of digital twins from operational data
- Integration of **complex data streams** within digital twins;
- Verification, validation, and uncertainty quantification for predictive digital twins
- Level or spectrum of **complexity and uncertainty** can be incorporated into a digital twin
- Capacity for **artificial intelligence/machine learning and big data** to facilitate and/or enhance digital twin technologies
- Importance of **user experience and interfacing techniques** (e.g., human-computer interactions, interfacing with biologics and/or micro-scale systems) in meeting the needs of digital twins across domains
- **Use—and possible expansion—of existing system technologies** (e.g., hardware, software) in facilitating the implementation of digital twins

Toward Building a Cancer Patient “Digital Twin”

Digital Twin: Dynamic, ultrahigh fidelity simulation of a physical system to make predictions through virtual experiments

Patient-tailored care incorporating multiomic, clinical, environmental and social data

Convergence of advanced computing and oncology across spatiotemporal scales

U.S. DEPARTMENT OF ENERGY | NIH NATIONAL CANCER INSTITUTE

https://events.cancer.gov/cbiit/dtwin2020

Enhance the biomedical data-science research workforce through improved programs and novel partnerships.

STRIDES Initiative (*The Science and Technology Research Infrastructure for Discovery, Experimentation, and Sustainability*) provides:

- State-of-the-art data storage and computational capabilities
- Training and education for researchers
- Innovative technologies such as artificial intelligence and machine learning
- Professional engineering and technical support

Partnerships with: Microsoft Azure Google Cloud aws

The STRIDES Initiative aims to help NIH and its institutions accelerate biomedical research by reducing barriers in utilizing commercial cloud services. This initiative aims to harness the power of the cloud to accelerate biomedical discovery. NIH and NIH-funded researchers can take advantage of STRIDES benefits.

Enroll Now

Gain access to

- Discounts on partner services
- Professional services consultations
- Access to training
- Potential collaborative engagements

>163
Petabytes of Data

201M
Compute Hours

>693
NIH & NIH-funded Research Programs/ Projects

$28M
Cost Savings

>4081
People Trained

https://datascience.nih.gov/strides

Leveraging STRIDES to enable FAIR Data Sharing

From managing over 163 petabytes of data in the cloud to enabling new data workflows to…

Managing Petabytes of Data

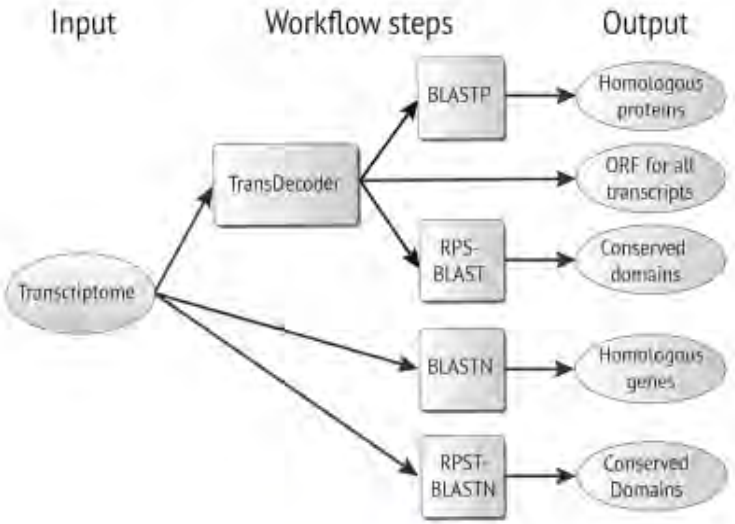
Figure 1. Growth of SRA over the last decade. Through September 2021 Public Access contains approximately 25.6 Petabase …

Nucleic Acids Res, Volume 50, Issue D1, 7 January 2022, Pages D387–D390, https://doi.org/10.1093/nar/gkab1053

OXFORD UNIVERSITY PRESS

Enabling FAIR, efficient, innovative workflows

Figure 2: Schema of the transcriptome annotation workflow.



Gigascience, Volume 10, Issue 2, February 2021, giaa163, https://doi.org/10.1093/gigascience/giaa163

OXFORD UNIVERSITY PRESS

Success Story: Transcriptome Annotation

Objective: To establish an estimate of the cost and compute time needed for the execution of multiple BLAST runs in a cloud environment.

Findings: Public cloud providers are a practical alternative for advanced computational biology experiments

The BLAST alignments required to annotate a transcriptome with ~500,000 transcripts can be processed in <2 hours with a compute cost of ~$200–$250.

The choice of cloud platform is dependent on details and requirements of the cloud providers

- accessibility for institutional use
- the technical knowledge required for effective use of platform services
- availability of open-source frameworks such as APIs to deploy the workflow

Transcriptome annotation in the cloud: complexity, best practices, and cost
Alvarez et al. *GigaScience*, Volume 10, Issue 2, February 2021, giaa163, https://doi.org/10.1093/gigascience/giaa163

In FY 22: Cloud Lab Pilot

Reducing barriers to entry

- Provide NIH and NIH-funded groups an easy route to access the cloud so they can quickly evaluate its utility for their project without having to make major time or financial commitments

Technical development

- Allow experienced teams access to the cloud environment(s) so they can prototype new architectures, and/or evaluate new software/hardware combinations in a cloud environment

Training

- Provide access to the cloud, simplifing access to tools and environments that can be used for training purposes

Final NIH Policy for Data Management and Sharing

(NOT-OD-21-013)

Release Date: **October 29, 2020 |** Effective Date: **January 25, 2023**

NIH requires researchers to prospectively plan for how scientific data will be preserved and shared through submission of a Data Management and Sharing Plan

Submission of a Data Management and Sharing Plan outlining how scientific data and any accompanying metadata will be managed and shared, taking into account any potential restrictions or limitations.

Plan is part of the budget Justification section of the application for extramural awards and as part of the technical evaluation for contracts

The DMS Policy applies to all research, funded or conducted in whole or in part by NIH, that results in the generation of scientific data. This includes research funded or conducted by extramural grants, contracts, Intramural Research Projects, or other funding agreements regardless of NIH funding level or funding mechanism. **The DMS Policy does not apply to research and other activities that do not generate scientific data, including training, infrastructure development, and non-research activities.**

Implementing FAIR Data Sharing

NIH strongly encourages
open access data sharing repositories
as a first choice

Scaled implementation options for sharing datasets

Datasets up to **2 gigabytes**

PubMed Central

- Stores publication-related supplemental materials and datasets directly associated publications.

Datasets up to **20*gigabytes**

Generalist Repositories

- Datasets associated with publications or otherwise and links to PubMed.

High priority datasets **petabytes**

Cloud Partners
(STRIDES Program)

- Store and manage large scale, high priority NIH datasets.

Optimized Funding for NIH Data Repositories and Knowledgebases

- Data resources are important research tools
- Historically funded through research grants
- Funding mechanism should be optimal for type of resource
- **End goal:** researcher confident in data and information integrity

- **Solution:** New Funding Announcement for data repositories and knowledgebases
- Resource plan requirement

Scientific Impact

Community Engagement

Quality of Data and Services and Efficiency of Operations

Governance

PAR-20-089 and PAR-20-097

Data Repository (DR) & Knowledgebase (KB) Program

An NIH program to support investigator-initiated, sustainable data resource development driven by critical research needs

- Fill a scientific need or gap
- Encourage adoption of good data management practices
- Engage the research community to contribute and use data
- Govern data life-cycle and preservation

In 2020-2021: 29 applications reviewed & 7 awarded

BioPortal

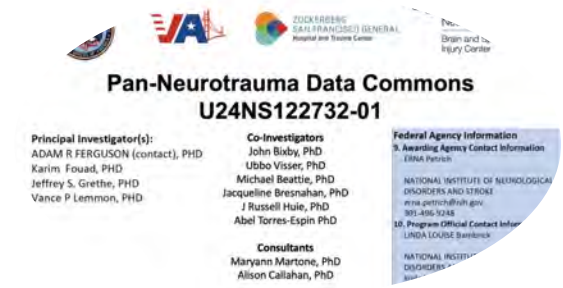
UniProt

ICTV

ctd

RED FLY

SHARC
SOUTHERN HIV & ALCOHOL RESEARCH CONSORTIUM


Pan-Neurotrauma Data Commons
U24NS122732-01
Principal Investigator(s):
ADAM R FERGUSON (contact), PHD
Karim Fouad, PHD
Jeffrey S. Grethe, PHD
Vance P Lemmon, PHD
Co-Investigators
John Bixby, PhD
Ubbo Visser, PhD
Michael Beattie, PhD
Jacqueline Bresnahan, PhD
J Russell Huie, PhD
Abel Torres-Espin PhD
Consultants
Maryann Martone, PhD
Alison Callahan, PhD
Federal Agency Information
9. Awarding Agency Contact Information
ERNA Patrich
NATIONAL INSTITUTE OF NEUROLOGICAL DISORDERS AND STROKE
10. Program Official Contact Infor
LINDA LOUISE Bambrick

NEW: The Generalist Repository Ecosystem Initiative

Solicit applications from generalist repositories working together to:

- Implement consistent capabilities (NOT-OD-21-016)
- Create better access to & discovery of NIH funded data
- Conduct outreach & train on FAIR data practices
- Engage the research community

Expected Outcomes

- Make data sharing easier
- Improve discoverability
- Increase reproducibility of research
- Encourage secondary use of data

The Dataverse Project · DRYAD · figshare · MENDELEY DATA · OSF · Vivli CENTER FOR GLOBAL CLINICAL RESEARCH DATA

(New) Data Curation Network – Event Series (ODSS, NLM)

https://datacurationnetwork.org/

Event 1: Kick-off Webinar for Researchers (date – Apr 6 2022)

Role of Librarians at Universities, Service offerings, Curation Resources

Event 2: Virtual (half-day) Workshop for Program Officers (~Jul 2022)

DMPs – Review/Evaluation, & Metrics of Review

Event 3: Virtual (half-day) Workshop for Curators (~ Oct 2022)

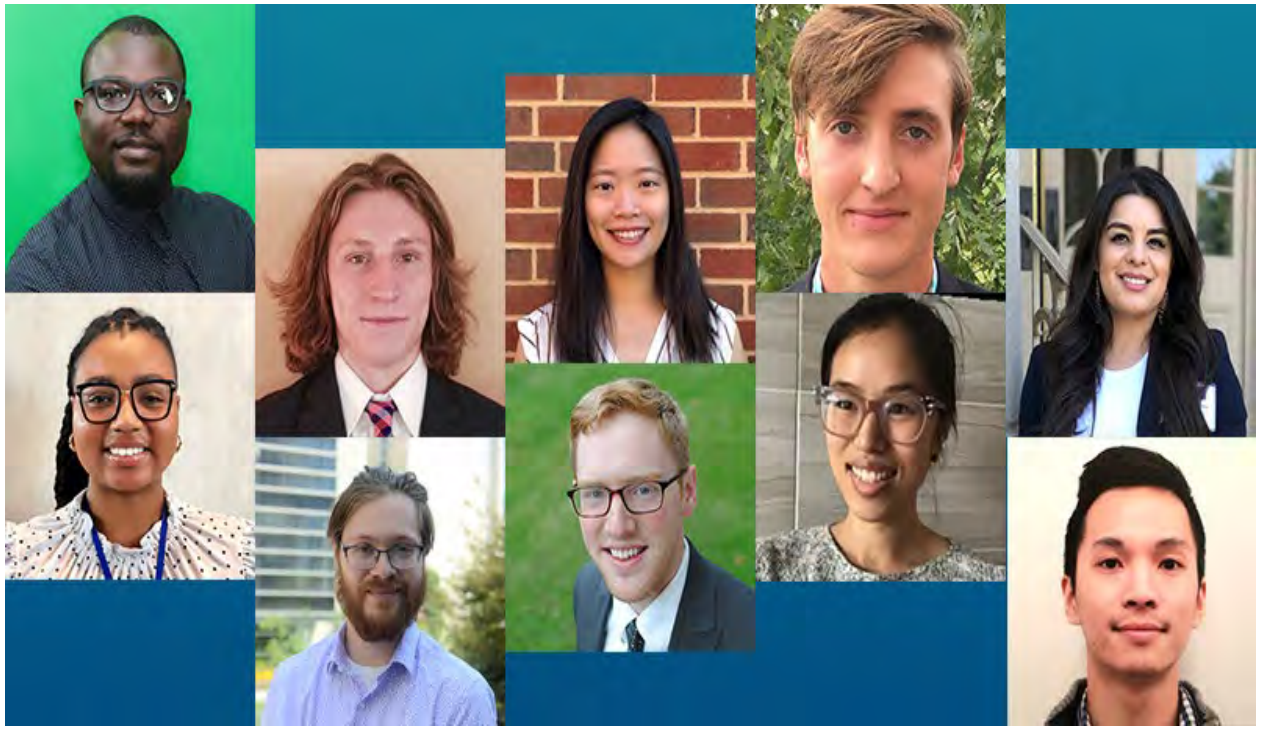
Train librarians, repository owners, others on curation of data type/format

Event 4: In-person Workshop for Curators (2 day) (~ Feb 2023)

Train librarians/curators on biomedical data types/formats – BYOD workshop

Make **Self-Paced Training Content** Available to Researchers, POs, Repository Owners & Other Curators

Graduate Data Science Summer Program



- Led by Office of Intramural Training and Education as part of Summer Internship Program
- Hosting **14** master's-level students for 10-week summer program
  - First cohort had 10 students
- Pilot driven by discussion with local universities consortium
- All interns placed in intramural research labs

https://www.training.nih.gov/data_science_summer

NIH Data and Technology Advancement (DATA) National Service Scholar Program

- One- or two-year **national service sabbatical** in high-impact NIH programs
- Experts in **Data science and technology** to advance NIH mission
- About 5 fellows each year, with **Third cohort** starting in 2022

Announcement to be posted Soon!

Example Projects include:

- Eye Health Data Interoperability
- AI-Ready Data for Pandemic Preparedness
- Automating Consumer Health Information
- Wearables Predicting Clinical Outcomes
- Multi Modal Cancer Data
- AI/ML for Genomics
- AI/ML for Medical Image and Clinical Data
- Accelerating Medicines Partnership

2021 Data Scholars and Projects

Dr. Anne Deslattes Mays
Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD)
Cross-project use cases for Kids First Data Resource & INCLUDE Data Hub

Dr. Priyanka Ghosh
National Center for Biotechnology Information in the National Library of Medicine (NLM)
Scalable, advanced search methods in Sequence Read Archive (SRA).

Dr. Lara Clark
National Institute of Environmental Health Sciences (NIEHS)
Software/code, documentation, tutorials, manuscripts, and outreach for environmental health.

Dr. Jaleal Sanjak
National Center for Advancement of Translational Science (NCATS)
Data integration applied across 7,000 rare diseases

Dr. John Gachago
Office of Data Science Strategy (OD)
Electronic health records for health disparities research and ethical use of machine learning/artificial intelligence (ML/AI) techniques.

Dr. Ansu Chatterjee
All of Us Research Program (OD)
Multimodal data integration and record linkages, and ethical data science and machine learning for research on large biomedical databases.

April Data Sharing and Reuse Seminar

The central nervous system (CNS) is intrinsically complex and CNS injury impacts molecules, cells, circuits, cognition, mood and behavior, and all organ systems.

To manage this complexity the fields of traumatic brain injury and spinal cord injury have launched community-driven data sharing initiatives and ecosystems for making research data FAIR and AI-ready.

Dr. Ferguson will provide an overview of these efforts and describe how machine intelligence on pooled data is revolutionizing precision care in these epidemiologically large, yet poorly understood areas of biomedicine.

Adam R. Ferguson, M.S., Ph.D. will present "Data Sharing and Machine Intelligence for Translational CNS Injury Research" at the monthly Data Sharing and Reuse Seminar on **April 8 at 12 p.m. ET.**

https://datascience.nih.gov/news/april-data-sharing-and-reuse-seminar-2022

Special Thanks

- **Enabling Cloud Data Storage and Compute: STRIDES** (Andrea Norris, Nick Weber and CIT Cloud Services team), **SRA** (Steve Sherry, Rodney Brister, Ravinder Eskandary and NCBI Team; Sarah Yang), **Hari Shroff**
- **Creating Integrated Cloud Platforms: NCPI** (Valentine Di Francesco, Ken Wiley, Tanja Davidsen, Alastair Thomson, Valerie Cotton, James Coloumbe, Haluk Resat, Chris Kinsinger, Huiqing Li, George Papanicolaou, Erika Kim, Ian Fore, Allen Dearry, Asiyah Lin, Natalie Kucher, Jon Kaltman, Chip Schwartz) **RAS** (Rebecca Rosen, Jeff Erickson, and CIT Deloitte team), and BioTeam
- **Implementing FAIR Data Sharing: FAIR Data Repositories Team** (Jennie Larkin, Ishwar Chandramouliswaran, Lisa Federer, Maryam Zaringhalam, Yanli Wang, Belinda Seto, Brian Haugen), **Lifecycle Metrics Team** (Dawei Lin, Kim Pruitt, Fenglou Mao, Elaine Collier, Christine Melchior, Minghong Ward, Anupama Gururaj, Noffisat Oki, Charles Schmitt), **Joint Repositories Tiger Team** (Michelle Heacock, Susan Wright, Matthew McAuliffe)
- **Software in the Data Ecosystem: Software Supplements Team** (Heidi Sofia, Mike Conway, Anthony Kirilusha, Jane Lockmuller, Todd Merchak, Andrew Weitz), **Smart and Connected Health Team** (Dana Wolff-Hughes, Fenglou Mao, Partha Bhattacharyya)
- **Actioning Biomed-ML (Artificial Intelligence):** Laura Biven, Alyssa Tonsing-Carter, **AI Workforce Team** (Michele Ferrante, Carol Shreffler, Jennifer Collins, James Gao, Anthony Kirilusha, Elizabeth Ginexi, Alison Gammie, Asif Rizwan) **AIM-AHEAD Team** (Brittany Chao, Nicole Redmond, Dina Paltoo, Chanel Press, Christian Evans, Erin Walker)
- **Building a Data Science Workforce:** Allissa Dillman, Jackie Cattell, Natasha Hurwitz, Jess Mazerik, Phil Ryan, Coding it Forward mentors

ODSS works with more than 200 individuals across nearly all NIH Institutes and Centers, as well as with multiple support contractors. We appreciate the contributions each and every person makes to our data science activities!

Office of Data Science Strategy

www.datascience.nih.gov

A modernized, integrated, FAIR biomedical data ecosystem