

**Concept Clearance: Digital Health Equity, Training and Research Consortium: Broadening the Benefit of Artificial Intelligence/Machine Learning (AI/ML) Technologies to Reduce Health Inequities and Enhance Diversity of the AI/ML Workforce**

**Background:** Following consultation with a working group of the Advisory Committee to the Director (ACD), and a series of internal discussions with NIH leadership, we identified a series of major challenges facing more widespread use of Artificial Intelligence/Machine Learning (AI/ML) capabilities, including:

- AI/ML capabilities have enormous potential to drive new medical discovery but can be costly, difficult, and time-consuming to implement and learn.
- The field of AI/ML runs the risk of perpetuating harmful biases in its practice, algorithms, and outcomes without more diversity of both data and researchers.
- Most biomedical studies and datasets still lack diverse representation of study participants from minority populations, and under resourced communities (e.g. some rural areas) leading to inadequate understanding of continued health disparities & inequities.
- Many under-represented communities have the potential to contribute clinical (and other) data, diverse recruitment, and cutting-edge science but may lack the financial, infrastructural, and training support to do so.
- Electronic Health Record (EHR) data can be a great “proving ground” to begin to build capacity and know-how but need to have a path over time to add Social Determinates of Health (SDOH), genetic, imaging, and other data types.

**NIH Strategy:** NIH envisions a multi-year program beginning with FY21 appropriated funds for AI/ML to develop and support a researcher and data network of highly diverse institutions with the capacity and interest to contribute their own data and to leverage AI/ML infrastructure, training, and know-how to conduct the biomedical research that is most important to them. Tackling the complex drivers of health disparities requires innovation, and a transdisciplinary framework that cuts across scientific and organizational silos to integrate multiple disciplines. NIH will actively seek public private partnerships to help achieve the ambitious goals of this program.

**Program Goals:** Establish mutually beneficial partnerships to increase the participation and representation of researchers and communities currently underrepresented in the development of AI/ML research. Research Opportunity Announcement(s) (ROAs) for funding in FY21 will focus on:

- Enabling a coordinated data and computing infrastructure that enhances the interoperability and AI/ML-readiness of data across the network, beginning with Electronic Health Records (EHRs), while preserving privacy and autonomy of control by each institution.
- Enhancing inclusion of groups currently underrepresented in the AI/ML research workforce across all scientific subject areas
- Building capacity and capability across the networked institutions through infrastructure investments and training.
- Fostering AI/ML approaches to redress the challenges of health disparities, health inequities, and minority health - especially with, for, and by researchers and institutions currently underrepresented in AI/ML research
- Supporting scientific research that can use Electronic Health Records (EHRs), connect the social determinants of health (SDOH), address biases and lack of data, and incorporate community-engaged research

- Catalyzing access to high quality diverse data sets both in terms of racial, ethnic, geographic location of participants, as well as types of data sets

### **Program Overview:**

Infrastructure: Create a collaborative network of federated, AI/ML-ready data where data are maintained, governed, and prepared by individual member institutions to preserve privacy and autonomy, while ensuring data interoperability across the network whether the data reside within the member institution or if the institution uses cloud infrastructure. AI/ML tools, applications, and other computations can run over this federated data, with permission from the host institution; and AI/ML models can be created, trained, and tested using this federated data, similar to federated learning. Investments in local institutions will take advantage of advanced AI/ML computing architectures and provide a gateway for the consortium network to build a community that can host and use data in the cloud. Critically, consistent software environments and tools will be used throughout the federation so that the network appears uniform regardless of where the user is in the network, or which computing resources are being used from laptop to data center to clouds; thereby facilitating the sharing of tools as well as training.

A phased approach to a coordinated and network-wide AI/ML effort begins with preparing participating institutions to use their EHR data, conduct collaborations, and run workflows; coordinated through central planning for data interoperability and tools. Later phases will enhance EHR with linkages to, for example, SDOH. A hybrid model of shared and private spaces for data storage and computing will provide opportunities to launch new users into AI/ML, improve data sets such as EHRs and SDOH, pooling and linking relevant data sets, while opening new avenues for AI/ML and health disparities research, and fostering the skill development for a diverse AI/ML workforce.

Training/Partnership: Form trans-disciplinary research partnerships that focus on facilitating regional partnerships to create a “network of networks” that integrates data science research networks with community engagement and clinical research networks. Desired outcomes are to a) enhance the diversity of the AI/ML research workforce for ALL scientific areas; and b) enhance the utilization of AI/ML methodology to advance research to address health disparities. Partnerships will prioritize equity in funding support to emphasize support of trainees, smaller academic institutions and community organizations vs. well-funded academic investigators/institutions. For each stakeholder, the goal should be to build on existing knowledge and facilitate a “pivot” to new area.

AI/ML Research Questions: Establish trans-disciplinary partnerships to build new, synthetic, or leverage existing datasets (EHR, images, genetics, SDOH, and other types of data), to develop and enhance AI/ML algorithms and apply AI/ML to improve healthcare, prevention, diagnosis, treatment, and implementation and intervention strategies. Research themes could include, for example: detection and mitigation of biases, criteria for AI/ML success, role and impact of SDOH and other factors, development of predictive models, and metrics to measure health inequities and disparities.

**Planned Funding:** \$50M/yr starting in FY21 using Other Transaction Authority (OTA).

## Appendix - Complementary Activities:

### Existing NIH FOAs:

- Administrative Supplements to Support Collaborations to Improve the AI/ML-Readiness of NIH-Supported Data ([NOT-OD-21-094](#)). *This opportunity is intended to support collaborations* that bring together in biomedicine, data management, and AI/ML to improve the AI/ML-readiness of data generated from NIH-funded research and shared through repositories, knowledgebases, or other data sharing resources.
- Administrative Supplements for Workforce Development at the Interface of Information Sciences, Artificial Intelligence and Machine Learning (AI/ML), and Biomedical Sciences ([NOT-OD-21-079](#)). The purpose is to *support the development and implementation of curricular or training activities at the interface of information science, AI/ML, and biomedical sciences* to develop the competencies and skills needed to make biomedical data [FAIR](#) (findable, accessible, interoperable, and reusable) and AI/ML-ready

### Examples of NIH-supported partnerships, programs, and initiatives:

#### Programs and activities focused on health equity research and researchers/communities underrepresented in biomedical research

- [NIMHD Research Centers](#)
- [Research Centers in Minority Institutions \(RCMI\)](#)
- [NIH Community Engagement Alliance \(CEAL\)](#)
- [Disparities Elimination through Coordinated Interventions to Prevent and Control Heart and Lung Disease Risk \(DECIPHeR\) Alliance](#)
- [The Institutional Development Award \(IDeA\)](#)
- [Resource Centers for Minority Aging Research \(RCMAR\)](#)
- [Native American Research Centers for Health \(NARCH\)](#)
- [The Institutional Development Award \(IDeA\)](#)
- [Resource Centers for Minority Aging Research \(RCMAR\)](#)

#### Programs and activities in Precision Medicine Data Science and AI/ML

- [NNLM | Network of the National Library of Medicine](#)
- [Bridge to Artificial Intelligence \(Bridge2AI\)](#)
- [Trans-Omics for Precision Medicine \(TOPMed\) Program](#)
- [All of Us](#)
- [NHLBI BioData Catalyst](#) and other NIH Data Sharing Resources
- [National Library of Medicine Data Sharing Resources](#)

#### Clinical Research Networks

- [National COVID Cohort Collaborative \(N3C\)](#)
- [Clinical and Translational Science Awards \(CTSA\) Program](#)
- [Practice Based Research Networks \(AHRQ\)](#)

### Examples of Additional Stakeholder Groups

#### Education and Training

- Minority Serving Institutions
- K-12 (early pipeline)
- R1 academic institutions/health systems

#### Health Systems

- [Federally Qualified Health Centers](#)
- Safety net health systems
- [Indian Health Service \(IHS\)](#)
- [Department of Veterans Affairs](#)
- Local/state health departments

#### Health and Science Related Federal Agencies

- [Indian Health Service \(IHS\)](#)
- [Department of Veterans Affairs](#)
- [Health Resources and Services Administration](#)
- [Patient-Centered Outcomes Research Institute \(PCORI\)](#)
- [Agency for Healthcare Research and Quality's \(AHRQ\)](#)
- [National Science Foundation \(NSF\)](#)

#### Social Service Federal Agencies

- [Housing and Urban Development \(HUD\)](#)
- [Department of Education](#)

#### Community

- Citizen Scientists
- Tribal Nations
- Local/state/regional social service agencies (governmental and non-profit)

#### Tech/industry

- Amazon
- Google
- Intel
- Microsoft