

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

Dina N. Paltoo, PhD, MPH
Assistant Director, Scientific Strategy and Innovation
Immediate Office of the Director, NHLBI



Major Challenges Identified

- AI/ML capabilities can be costly, difficult, and time-consuming to implement and learn.
- The field runs the risk of perpetuating harmful biases in its practice, algorithms, and outcomes without more diversity of both data and researchers.
- Biomedical studies and datasets lack diverse representation leading to inadequate understanding of continued health disparities and inequities.
- Many under-represented communities have the potential to contribute data, diverse recruitment, and cutting-edge science but may lack financial, infrastructural, and training support.
- 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.



FY21 Appropriations Language:

*Artificial Intelligence/Big Data. **Advancing life sciences is increasingly dependent on data computation and infrastructure, machine learning (ML), and collaborative scientific initiatives. NIH is to be commended for leveraging the potential of ML to accelerate the pace of biomedical innovation ... The agreement includes \$105,000,000 to support the agency's efforts, including \$50,000,000 to expand the number of ML-focused grants and \$55,000,000 for ODSS ... There is a growing consensus in the research community that more training is needed for the use of FHIR in clinical and biomedical research, and the recommendation supports expanded training, including for underrepresented and underserved groups.***



NIH Strategy

- Multi-year program beginning with FY21 appropriated funds
- Develop and support a researcher and data network of highly diverse institutions. Institutions will:
 - Contribute Data
 - Leverage AI/ML infrastructure, training, and know-how
 - Conduct the AI/ML biomedical research that is most important to them and aligned with the NIH mission
- Actively seek public private partnerships to help achieve the ambitious goals of this program



Digital Health Equity, Training and Research Consortium

Goal:

Establish mutually beneficial and coordinated partnerships to increase the participation and representation of researchers and communities currently underrepresented in the development of AI/ML models and enhance the capabilities of this emerging technology, beginning with EHR data.

Focus:

- Establish a coordinated data and computing infrastructure
- Enhance inclusion of groups underrepresented in the AI/ML research workforce and participation across all scientific subject areas
- Redress the challenges of health disparities, health inequities, and minority health
- Support research questions that can use EHRs, connect SDOH, address biases and lack of data, develop predictive models, and incorporate community-engaged research
- Catalyze access to high quality diverse data sets



Digital Health Equity, Training and Research Consortium



Infrastructure

Form a federated data network where data are maintained, governed, and prepared by individual member institutions to preserve privacy and autonomy while ensuring data interoperability across the network.



Training/Partnership

Facilitate regional multi-disciplinary partnerships to create a “network of networks” that integrates data science research networks with community engagement and clinical research networks to form mutually beneficial collaborations and support engagement of underrepresented scientists across the career pipeline.



AI/ML Research Questions

Foster trans-disciplinary partnerships to build new, synthetic, or leverage existing datasets (EHR and other types of data), to develop and enhance AI/ML algorithms and apply AI/ML approaches to address health inequities and disparities, to improve healthcare, prevention, diagnoses, treatments, and intervention and implementation strategies.



Strategy for Infrastructure



Key Attributes

Part I: A collaborative network of federated data. Data will remain under the control of the site either on-prem or in the cloud to preserve privacy and autonomy while promoting data interoperability.

Part II: AI/ML applications and tools can run over federated data with permission from the sites. AI/ML models can be tested using data from the sites, similar to federated AI-learning.

Part III: Infrastructure appears seamless to the data user – similar interface to data and tools independent of where the user is in the federation, or what computing resource is used. Institutional investments will take advantage of advanced AI/ML architectures and provide a gateway to the cloud.



Strategy for Training/Partnership



Key Attributes

Partnership

- Convene regional networks to bring together data science, community engagement, and clinical research
- Facilitate integration of underrepresented researchers across the career pipeline to foster a "developmental network"
- Prioritize equity in funding support

Training

- Identify general and specialized curricular components
- Build on existing knowledge of each stakeholder to pivot to new areas
- "Train the trainer" approach to enhance capacity building



Strategy for AI/ML Research Questions



Key Attributes

- Form **research partnerships** that will **build datasets** and **leverage existing data/resources** to use AI/ML, develop AI/ML **algorithms and approaches**:
 - Datasets (e.g., EHR, images, -omics, SDOH, other types of data)
 - NIH-supported programs, initiatives, studies, platforms/repositories
 - Ensure equitable access and sharing of best practices
- Engage **stakeholders** to identify use cases, pilot projects, research interests
 - Potential **research areas**: detect and mitigate biases in EHR and other datasets, criteria for AI/ML success, role and impact of SDOH and other factors on health, metrics to measure health disparities and inequities, and predictive models to prevent, treat and implement healthcare strategies



Next Steps

- **Stakeholder Engagement – May/June**
- **Refining Initiative – June/July**
- **Publish ROAs – Summer 2021**
- **Award – September 2021**

Thanks to the Concept Team

Partha Bhattacharyya

Laura Biven

Brittany Chao

Michael Chiang

Dina Demner-Fushman

Joshua Denny

Eric Dishman

Deborah Duran

James Gao

Jordan Gladman

Rebecca Goodwin

Susan Gregurick

Clem McDonald

Michele McGuirl

Dina Paltoo

Chanel Press

Nicole Redmond

Neha Shah

Rob Star

Lawrence Tabak

Alastair Thomson

Xujing Wang

Monica Webb Hooper

Shannon Zenk

