The City College of New York
Office of Research

CCNY College-wide Research Vision (CRV) Initiative Concept White Paper

Artificial Intelligence for Health Equity and Diversity (AIHED)

Project Team:

Principal Investigators: Ashiwel Undieh, Jie Wei, Maria Lima, Karen Hubbard, Reza Khanbilvardi.
Department/School: Medicine, Engineering, Architecture, Colin Powell, Science Division, CCNY
Other Key Personnel: Noel Manyindo, Victoria Frye (Medicine); Akira Kawaguchi, Zhigang Zhu, Jeff Garanich, Bingmei Fu (Engineering); Kevin Foster (Powell School), Ahu Aydogan (Architecture).

Project Concept Description : (Maximum 2 pages)

Keywords: Health equity, Health disparities, Chronic diseases, Artificial intelligence, Machine learning, Big data, Community participatory research, Researcher diversity, Health apps.

Objective: The overall goal of AIHED is to establish a translational research program focused on development and deployment of augmented intelligence technologies to advance health equity and the diversity of healthcare providers and researchers in New York City and State. AIHED will establish a collaborative partnership among CCNY researchers across schools and divisions to mobilize and integrate a diversity of talent, expertise, resources and approaches toward enhancing the participation of researchers and communities currently underrepresented in the development and deployment of artificial intelligence technologies geared toward addressing health disparities and inequities. During the phase of CRV support, the program will focus on teaming, building of requisite skills and capabilities, enlisting critical infrastructure for pilot projects in AI for health, proof-of-concept demonstrations, and work toward garnering external resources to enhance continuity and success of the program's mission.

The specific objectives of the program are: (1) To conduct multi-modal electronic health record (EHR) data collection, management, visualization, analysis and outcome measures using de-biased and defined case scenarios; (2) To develop prototype AI algorithms through studies that include outcomes addressing health disparities; (3) To foster collaborative team development of pilot AIHED projects based on multi-modal clinical/biomedical data, including EHR and social and environmental determinants of health (SEDOH), and other well-curated data from partnering organizations; and (4) To organize research skills training modules that will broaden the diversity of the talent of researchers and innovators engaged in AIHED-type research and applications at CCNY and across CUNY and the City.

Approach: Following will be our approaches to tackling each of the above objectives:

Obj.1: We will explore and examine current algorithms and data feeds to carefully classify and cluster health disparities from curated data to gain deeper insights into bias phenomena, their ramifications, and standardization strategies that could be implemented to prevent or minimize data and algorithm biases that tend to perpetrate health disparities.

Obj. 2: We will establish and exploit trusted partnerships with corporations marketing AI tools and with institutions and organizations that have a core mission to serve underserved groups impacted by health disparities, and through these partnerships we will gain access to resources and data sources to feed and evolve our own algorithms and test beds. These multimodal studies will then drive the design of pilot projects focusing on data curation or deployable product development.

Obj. 3: To foster development of pilot AIHED projects, program members will collect and curate multi-modal clinical/biomedical data including HER and genomics data, and SEDOH (including CUNY-CREST surveillance data) and develop collaborative projects for case studies and pilot products.

Obj. 4: To organize training programs that will broaden the diversity of the talent of researchers and innovators engaged in AIHED research and applications, we will leverage CCNY's existing DSE and MTM graduate programs to introduce AIHED modules encompassing course electives and capstone projects focusing on SEDOH, the EHR, AIHED application development, and community participatory research enlistment to enhance our mission success.

Expected Outcomes: The expected outcomes include standards for unbiased AIHED data collection, AIHED algorithms developed, results of EHR/SDH case studies, pilot results from AIHED product development projects, community sensitization to AIHED research and product deployment plans.

Expected Products: Results of case studies and pilot projects supported under this CCNY initiative are expected to generate at least two high-impact publications per year. These outcomes will also form the basis for at least two external grant proposals to support scale up of data collection and algorithm development as well as application deployment for use in communities impacted by health disparities.
Merits: Artificial intelligence (AI) has created landscape changes in a broad array of fields and applications due to dramatic advances in capabilities of hardware such as GPU and HPC systems, machine-learning/deep-learning algorithms, and increasingly available data digitization such as the EHR. The expanding AI revolution is yet to substantially impact the healthiness of New York residents, particularly members of minority and underrepresented communities. This project will bring together productive CCNY researchers and teachers in AI, medical science, and SEDOH to address healthiness and health disparities with state-of-the-art techniques. Our project will also help train underrepresented and underserved researchers via an integrated and inclusive AIHED curriculum, so that these young professionals can take advantage of the AI revolution and skills to expand their reach and careers.
Impact: We expect that the research will generate system prototypes using AI/ML techniques, well curated, debiased, multi-modal EHR data sets standardized to fairly include underserved communities, Electronic design automation tools to help harness EHR and other health data, and accessible tutorials and learning aids to disseminate our findings and inform future research, development and deployment.

Milestones: Milestones are not linear as most objectives will proceed concurrently over the 3 years.
Months 1-36: Teaming and partnership development. Recruit additional 3-6 faculty/fellows to create a 10-15 researcher Program encompassing diverse backgrounds, expertise, and capabilities relevant to AIHED. Develop internal cohesiveness of team by sharing strategy, tactics and responsibilities and by implementing activities to strengthen trust and technical capabilities such as open sharing of research interests and methods, planning of regular program meetings, workshops and seminars, and developing joint pilot projects. Recruit 5-7 corporations or organizations as Program partners and external advisors, e.g., IBM Health, Amazon Health, St. Barnabas Hospital, Northwell health, etc.
Months 3-12: Develop training curricula, materials, and systems to incorporate AIHED concepts, methods and objectives into MTM Biodesign and DSE Capstone projects. Deliverables: Curriculum, teaching materials and lab/practice activities to be piloted and refined in Years 2-3.
Months 3-15: Data sourcing and curation for algorithm development. Negotiate access rights to data from multiple sources such as EHR and SEDOH databases to support case and pilot studies.
Months 6-18: Develop validated strategies for improving quality, diversity, balance, debiasing and interpretation of data used for AIHED research and applications.
Months 6-27: Develop AI algorithms using case studies that include outcomes for health disparities, exploring and examining algorithms to carefully classify, cluster and characterize health disparities and their correlates from curated data in order to gain deeper insights into the phenomena of bias and strategies to eliminate bias in data used for AIHED algorithms and product designs.
Months: 13-36: Use multi-modal EHR data collection, management, visualization, analysis and their associated health disparities research to develop case studies that will drive the design of data sharing and training offerings as well as future projects.
Months 18-36: Pilot the collaborative development of AIHED projects and products involving multi-modal clinical/biomedical data, including EHR, SEDOH, and other well-curated data from internal and partner sources to demonstrate feasibility and power of AI applications to improve the health of individuals and especially of underserved communities.

Budget (Maximum Budget $200K):

Personnel Costs: (please list key positions and estimate budget required)

- Key Personnel (PI, Co-PI, senior personnel): 0.5 mo summer salary for 8 PIs/Co-I's = $72,000.
- Research staff: We estimate 1 fulltime post-doc/ data scientist at $80,000
- Administration staff: We estimate 1 part-time admin/college assistant at $12,000
- Students: We estimate engaging 4 PT MS research assistants @ $7,000 each = $28,000

OTPS Costs: Supplies, data access subscriptions, domestic travel, etc., estimated @ $8,000.
Total estimated budget from the above is $200,000/year = $600,000 for the 3-year project.