AI/ML AND VIRTUAL HUMAN PLATFORMS FOR THREAT AGENT HAZARD ASSESSMENT AND MEDICAL COUNTERMEASURE DISCOVERY AND DRUG DEVELOPMENT

Ai Assisted Multi-omics Data Harmonization And Integration

Mark Weston Netrias, LLC **Mohammed Eslami** Netrias, LLC **Alexander Verbitsky** Netrias, LLC **George Zheng** Netrias, LLC

Analysis of any data in biology requires representation and integration of the biological context in which that data was generated. This integration process is currently a manual, cumbersome, labor intensive process. Our research bridges two domains - machine learning and data integration - with state-of-the-art assistive technology that significantly accelerates the data integration process, leading to faster turnarounds and discovery loops. This technology has a broad appeal across the DTRA CBD S&T spectrum, but is particularly relevant to complex, multi-omics data integration efforts, such as Multi-Organ-on-Chip platforms. Term alignment and standardization is needed to ensure datasets can be combined and made ready for Artificial Intelligence/Machine Learning (AI/ML) algorithms.

Despite advancements in ontologies and workflow tooling, data integration and metadata harmonization remain a significant challenge for multi-omics data integration. We present the Active Discovery Engine (ADE), a Netrias platform that uses machine learning techniques to perform AI-assisted data harmonization and integration. We will showcase examples that highlight term alignment, automated data joining, and custom analytics integration to produce AI/ML ready data.

Bench scientists, engineers, data scientists, and decision makers will gain an enhanced understanding of the modern challenges of data harmonization and how we address them with our AI assisted data integration technology. Together, we will remove the manual burden of data integration on the individual researcher and empower organizations to bring together disparate, heterogeneous data to accelerate the application of AI and machine learning predictive techniques.

This material is based upon work supported by the Defense Advanced Research Projects Agency (DARPA) under Contract No. 140D0420C0032.