In partnership with EU consortium, GE HealthCare to advance AI in healthcare through synthetic data

Access to data is fundamental to AI product development, guiding the process from initial creation to final deployment. However, access to sensitive patient data is often restricted due to privacy, ethical and regulatory constraints, often leading to a bottleneck in the development cycle. Addressing these challenges requires a concerted effort from healthcare providers, policymakers, clinical experts and technology developers.

The generation and use of synthetic data that mimics real-world data offers a promising solution to these difficulties. The creation of synthetic datasets has recently gained momentum and could play a pivotal role in overcoming the challenges of data scarcity, bias or lack of generalizability without compromising privacy and data security requirements. Despite the promise of synthetic data, it also comes with unresolved questions, such as ensuring data accuracy and reliability, maintaining privacy or identifying the best synthetic data generation methods for the various potential application scenarios.

GE HealthCare is leading a multidisciplinary collaboration of 32 medtech/pharma industry and academia partners from 16 countries to address this issue by providing a 360° vision on how to advance healthcare applications through synthetic data use. The SYNTHIA consortium aims to evaluate and deliver proven methods, standards and frameworks to build reliable tools for synthetic data generation and their use in the development, training and validation of AI algorithms. It brings together expertise across multiple domains from both private and public contributors to address the challenges of algorithm development with synthetic data along legal, ethical and regulatory considerations, while exploring methods to increase the availability of high-quality training datasets.

The research efforts will center around six diseases: two solid tumors (lung cancer and breast cancer), two blood cancers (multiple myeloma and diffuse large B-cell lymphoma), one neurodegenerative disease (Alzheimer's disease) and one metabolic disease (type 2 diabetes). As such, the SYNTHIA consortium will tackle the critical need for privacy-preserving data solutions in healthcare by developing validated tools and methods for synthetic data generation across various data types, including laboratory results, clinical notes, genomics, imaging and mobile health data.

Within GE HealthCare, research teams specializing in MR, X-ray and women's health are collaborating closely with their colleagues from the Science and Technology Organization. Together, they are working with the other corporate and academic consortium partners to further their collective goals. The European GE HealthCare MR research team will focus on leveraging synthetic data to advance MR applications and techniques in breast cancer imaging.

The Health Research Institute Hospital La Fe (IIS La Fe) is the academic lead for SYNTHIA coordinated by Leonor Cerdá-Alberich, PhD, Principal Investigator and Head of Computing and Artificial Intelligence.

Aligned with GE HealthCare's AI strategy, SYNTHIA and its pioneering consortium of world-leading experts will pave the way to harnessing synthetic data for accelerating patients' access to novel drugs, tools and devices, maximizing the development of precision medicine while maintaining sustainability of EU healthcare systems. Ⓢ

"SYNTHIA is shaping the future of healthcare by delivering validated, regulatory-compliant synthetic datasets for clinical applications and accelerating market-ready solutions while eliminating barriers in data sharing, scarcity and bias."

Leonor Cerdá-Alberich, PhD