

Building a FAIR-Compliant Platform for AI-Ready Data in Particle Accelerators

Wednesday, March 6, 2024 1:50 PM (20 minutes)

In the rapidly evolving field of particle accelerator technology, Machine Learning (ML) shows great potential in optimizing accelerator performance and predictive maintenance. However, the success of these applications often depends on high-quality, real-time data sources. This paper introduces the plan and status of building an innovative machine learning data acquisition platform, specifically designed for continuously generating machine learning datasets applicable in the field of particle accelerators. The platform adheres to the FAIR principles, i.e., the Findability, Accessibility, Interoperability, and Reusability of the data. It also will ensure data integrity, consistency, and reliability during data collection, ultimately generating an AI-ready dataset. By utilizing advanced data collection networks, edge computing, and storage technologies, the platform achieves real-time data capture, preprocessing, and annotation. Importantly, the platform will include a variety of algorithms that intelligently select and store data points based on real-time operational states and the needs of machine learning models, and provides appropriate annotations for the data. This research will not only offer a powerful data support tool for the operation and maintenance of particle accelerators but also provide new perspectives and methods for the industrial sector on how to effectively acquire and manage data for machine learning applications.

Primary Keyword

Secondary Keyword

Tertiary Keyword

Primary authors: LU, Xiaohan; JIAO, Yi (Institute of High Energy Physics)

Presenter: LU, Xiaohan

Session Classification: Infrastructure / Deployment Workflows

Track Classification: Infrastructure / Deployment Workflows