Contribution ID: **432** Contribution code: **WEPD082** Type: **Poster Presentation**

Toward particle accelerator machine state embeddings as a modality for large language models

Wednesday, 24 September 2025 16:30 (1h 30m)

Understanding and diagnosing the state of a particle accelerator requires navigating high-dimensional control system data, often involving hundreds of interdependent parameters. We propose a novel multimodal embedding framework that jointly learns representations of machine states from both numerical control system readouts and natural language descriptions. This enables the translation of complex machine conditions into human-readable summaries while maintaining fidelity to the underlying physical system. The obtained embeddings are subsequently adapted to an open-weights large language model via cross-attention conditioning. We demonstrate a first implementation trained on European XFEL machine state data. This work covers the embedding model architecture, training methodology, and presents initial examples demonstrating the model's capabilities in action. Due to the general concept of machine state, the model can be easily adapted to other facilities and control system environments.

Funding Agency

Footnotes

Author: MAYET, Frank (Deutsches Elektronen-Synchrotron DESY)

Co-authors: SULC, Antonin (Lawrence Berkeley National Laboratory); TENNANT, Chris (Thomas Jefferson National Accelerator Facility); HELLERT, Thorsten (Lawrence Berkeley National Laboratory)

Presenter: HELLERT, Thorsten (Lawrence Berkeley National Laboratory)

Session Classification: WEPD Posters

Track Classification: MC13: Artificial Intelligence & Machine Learning