



Contribution ID: 114

Type: **Early Career Scientist**

Large Language Models for Particle Physics Experiments

Particle Physics experiments rely on large code bases. These include proprietary code for common programming languages, as well as code and instructions for specialized programs and languages. Working with these, as well as developing, maintaining, and creating such code is part of the work of particle physicists.

AI driven Large Language Models like Chat GPT-4 have recently made great leaps in not only processing and generating natural language, but they have also shown themselves adept at handling code. They can provide solutions to common coding problems, generate working code snippets, explain parts of given code, and can introduce users interactively to new programming languages.

We propose here to establish a service that provides particle physics experiments with the ability to train and use Large Language Models on their code bases to aid in the development and use of the relevant code. Additionally, this service should maintain Large Language Models that are explicitly trained on commonly used particle physics specific software like Monte Carlo Event generators and pertinent specialized domain knowledge to service as an interactive knowledge repository.

This service would allow to operate experiments using such services at greater efficiency, would speed up R&D and data analysis efforts, while also lowering the learning curve new members of experiments.

Primary authors: WEBER, Christian (Brookhaven National Laboratory); Dr ZHIVUN, Elena (Brookhaven National Laboratory)