Contribution ID: **163**   Type: **Oral Presentation**

Development of an LLM-Based System for Automatic Code Generation from HEP Publications

Friday, 20 March 2026 09:00 (22 minutes)

Ensuring the reproducibility of physics results is a critical challenge in high-energy physics (HEP). In this study, we aim to develop a system that automatically extracts analysis procedures from HEP publications and generates executable analysis code capable of reproducing published results, leveraging recent advances in large language models (LLMs).

Our approach employs open-source LLMs to accurately extract event selection criteria, definitions of physical quantities, and other relevant information described in scientific papers. The system also traces referenced publications when necessary, enabling the construction of a selection list that remains faithful to the original analysis. Based on this extracted information, the system automatically generates analysis code, which is executed on ATLAS Open Data to validate the reproducibility of the published results.

Specifically, we utilize proton-proton collision data recorded in 2015 and 2016 by the ATLAS experiment and released as open data. This dataset allows direct comparison between our automatically generated analysis results and those reported in the literature. By comparing with manually developed analysis code used as a baseline, we evaluate the current performance of open-source LLMs in terms of code quality, computational efficiency, and physics validity.

This research contributes to the advancement of reproducible science in HEP and explores the potential of LLM-driven automation in physics analysis workflows. In the long term, our study envisions research environments where physicists can perform data analysis through natural-language interaction, and where automated verification and review support improves the reliability and accessibility of scientific publications. In this presentation, we report the status of our prototype system and initial performance evaluation results.

Primary authors: TANAKA, Junichi (University of Tokyo); SAITO, Masahiko (ICEPP, The University of Tokyo); KISHIMOTO, Tomoe (KEK)

Presenter: SAITO, Masahiko (ICEPP, The University of Tokyo)

Session Classification: Artificial Intelligence (AI) - IV

Track Classification: Track 10: Artificial Intelligence (AI)