Applicant: Rūta Lomsargytė
Mentor: Peter Elmer
Project Duration: 8 Weeks
Proposed Start Date: 1st July 2025

IRIS-HEP Fellows Program Project Proposal

Development of an LLM-Based Co-Pilot Tool for CMS Analysts

Project description

In the recent years, Large Language Models (LLMs) have been helping users access complicated technical information and generate code, potentially saving time. In the context of CERN, CMS analysts often interact with specific software, internal documents and other computing tools. However, the internal documentation is not accessible to mainstream LLMs, therefore the information they output is not as accurate as it could be.

The aim of the project is to explore the possibilities of building an LLM-based co-pilot tool that assists CMS analysts. The tool would provide helpful answers to technical questions, assist in writing or understanding code, and retrieve relevant technical information from internal CMS sources. To reduce the need for fine-tuning and to optimize the use of computing resources, retrieval-augmented generation (RAG) technique will possibly be implemented. The project will focus on using internal CMS data while making sure that usage follows CERN policies.

Project goals

- Build an assistant tool that can:
  - Find and show relevant parts of CMS documentation.
  - Answer common questions about CMS analysis.
  - Help understand or write code snippets.
- Use CMS specific materials, including:
  - CMS Twikis, user guides, and other documents.
  - Code examples.
- Ensure safe and efficient usage by:
  - Using open-source or self-hosted large language models.

  - Implementing retrieval-augmented generation (RAG) technique.
  - Making sure the tool follows CERN data privacy policies.

Preliminary timeline

Weeks 1–2:

- Determine common tasks performed by CMS analysts that could benefit from LLM assistance.
- Set up the basic environment and create a plan for the rest of the project.

Weeks 3–5:

- Build a tool capable of understanding and responding to basic queries.
- Make the internal CMS documents accessible to the tool.
- Test the tool and correct based on feedback.

Weeks 6–8:

- Improve the tool's ability to retrieve and output relevant information.
- Prepare a final demonstration and documentation of the tool.
- Draw conclusions for possible future improvements.

References

[1] CERN openlab Technical Workshop 2025, "Development of Large Language Models at CERN," Mar. 2025. [Online]. Available: `https://indico.cern.ch/event/1440389/contributions/6364348/attachments/3024469/5339708/LLMs@CERN_OpenLabWS.pdf`

[2] CMS Collaboration, "CMS Software Documentation – The CMS Workbook," [Online]. Available: `https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBook`