CERFACS

JOB OFFER – STAGE

HPC Software Engineering Intern – AI for HPC

OFFER INFORMATION

Reference: COOP-25-AD-03
Team: COOP

Location: 42 Avenue Gaspard Coriolis – 31057 Toulouse

Supervisors:

- Antoine Dauptain
- Thibault Marzlin

Gratification: 800€ net per month - M2 level or last year at engineering school

Period: 6 months - from: 09/02/2026

Key words: High Performance Computing (HPC), Large language models (LLMs), Programming, Artificial Intelligence (AI), Database

CERFACS

Cerfacs is a private research, development, transfer and training center for modeling, simulation and high-performance computing. Cerfacs designs, develops and proposes innovative software methods and solutions to meet the needs of its partners in the aeronautics, space, climate, environment and energy sectors. Cerfacs trains students, researchers and engineers in simulation and high-performance computing.
Cerfacs works closely with its seven partners: Airbus, Cnes, EDF, Météo France, Onera, Safran et TotalEnergies.

HOSTING TEAM - COOP

The COOP (Computing and Operational Practice) team aims to accelerate the adoption of best-practise techniques for scientific software on modern high-performance computing (HPC) architectures. Key research and development activities include: software portability, efficiency and scalability; mesh manipulation, adaptation and partitioning; industrial technology transfer; code analytics; machine learning and artificial Intelligence in HPC.

CONTEXT

The COOP team is at the forefront of pushing the boundaries of scientific computing by integrating artificial intelligence into the heart of our workflows. We've seen how modern HPC systems provide immense computational power, but their complexity presents a significant barrier to optimization. The sheer scale and specialized nature of these systems, with their unique architectures, massive parallelization, and intricate memory hierarchies, mean that traditional development methods are often slow and inefficient.

Tackling these challenges manually is a formidable task. Code porting, optimization, and debugging for different HPC platforms are time-consuming processes that require deep expertise and are prone to error. This complexity slows down scientific discovery and engineering innovation.

We believe that Large Language Models (LLMs) can be a transformative tool in this domain. This internship is designed to explore this potential. The intern will work on practical applications of LLMs to address the very real

and complex problems we face daily in AI for HPC, from automating code translation and optimization to creating intelligent assistants that can diagnose and resolve performance issues. He/She will help us build the next generation of tools that empower scientists and researchers to leverage the full power of HPC.

MISSION

The internship focuses on exploring and implementing cutting-edge AI techniques to address core challenges in high-performance computing. The intern will investigate how large language models (LLMs) can be leveraged to assist with complex tasks that are currently performed manually.

The work will begin with a survey of existing LLMs and their potential applications for HPC code. Next, a promising approach will be selected and implemented, focusing on one or more of the following areas: automating code-to-code translation, assisting with performance optimization, or creating an intelligent error troubleshooting tool.

This mission is not just about writing code, it's about building a solid foundation for future development. The intern will be responsible for creating and improving datasets, experimenting with various models, and fine-tuning them to achieve optimal results. The contributions will be critical to demonstrating the tangible value of AI in making scientific computing more efficient and accessible.

DESIRED PROFILE

- Academic background: Currently pursuing a Master's degree (M2) or engineering school in Computer Science, Data Science, or a related field.
- Essential Programming Skills: Proficiency in Python, C/C++ or Fortran, with a good understanding of machine learning frameworks like PyTorch or TensorFlow.
- Solid understanding of machine learning and large language models (LLMs). Basic knowledge of HPC concepts or parallel programming would be a plus.
- Proactive, analytical skills, autonomous and a passion for tackling complex challenges.

WHAT WE OFFER AT CERFACS

- Broad access to technology, a rich interpersonal environment, in-house skills recognized nationally and internationally.
- An inclusive and equitable work environment.
- A structure accessible to people with disabilities.
- Possibility of benefiting from 1.83 days of reduced working hours per month, linked to your choice of a 39-hour rather than 35-hour working week.
- 50% reimbursement of public transport costs.

HOW TO APPLY ?

To apply, please send your CV and covering letter to dauptain@cerfacs.fr and mazlin@cerfacs.fr , applications are open until 15/01/2026.

See you soon at CERFACS!