# Offer #2023-06854

# PhD Position F/M Reliability and Security of Large Foundation Models

**Contract type :** Fixed-term contract

**Level of qualifications required :** Graduate degree or equivalent

**Fonction :** PhD Position

## About the research centre or Inria department

The Inria Rennes - Bretagne Atlantique Centre is one of Inria's eight centres and has more than thirty research teams. The Inria Center is a major and recognized player in the field of digital sciences. It is at the heart of a rich R&D and innovation ecosystem: highly innovative PMEs, large industrial groups, competitiveness clusters, research and higher education players, laboratories of excellence, technological research institute, etc.

## Assignment

Large Foundation Models (LFMs) are cutting-edge technology for natural language processing, object detection and segmentation, and audio and multimodal processing, outperforming any available machine learning technique. LFMs, such as OpenAI GPT-4, Google ViT, and Meta LLaMA, have gained public attention with their unprecedented accuracy. Given the superior performance of LFMs, they are being deployed in safety-critical and mission-critical applications, including space exploration and self-driving cars. Improving LFMs' security and reliability is crucial to enable dependable real-time safety-critical systems.

Large and complex accelerators like Graphics Processing Units (GPUs) are ideal for deploying LFMs in safety-critical applications. However, GPUs integrated into safety-critical systems must meet specific constraints, including real-time execution and high classification/detection accuracy, even in harsh environments. It is imperative to evaluate whether these critical requirements are met when undesirable events, such as radiation-induced faults and electromagnetic hardware attacks, disrupt correct hardware execution and modify the expected results of the LFMs.

This Ph.D. aims to identify hardware and software vulnerabilities in LFM-based systems and propose error mitigation techniques.

## Main activities

The Ph.D. student will characterize the impact of radiation-induced faults and electromagnetic hardware attacks on system reliability and security on GPUs for vision, language processing, and multimodal LFMs. The results will be combined with software simulation data to identify effective hardening solutions. The Ph.D. student will work on developing new fault tolerance approaches tailored for LFMs. Standard fault tolerance techniques may introduce unacceptable overhead. We will conduct a comprehensive fault propagation analysis to propose efficient and effective hardening methods.

The Ph.D. student will participate in international experiments and internships at laboratories like Rutherford Appleton Laboratory in the UK and Los Alamos National Laboratory in the USA. In addition, the student will participate in conferences and international projects. This can help them to develop their research skills and network with other professionals in their field.

## Skills

**Required technical skills:**

- Good knowledge of computer architectures and embedded systems
- Good programming knowledge (C/C++, python)
- Basics of Machine Learning (pytorch/tensorflow)
- Experience in fault tolerant architectures is a plus
- Experience with hardware design is a plus

Candidates must have a Master's degree (or equivalent) in Computer Science, Computer Engineering, or Electrical Engineering.

**Languages**: proficiency in written English and fluency in spoken English is required.

**Relational skills**: the candidate will work in a research team, where regular meetings will be set up. The candidate has to be able to present the progress of their work in a clear and detailed manner.

**Other valued appreciated**: Open-mindedness, strong integration skills and team spirit.

**Most importantly, we seek highly motivated candidates.**

## Benefits package

- Subsidized meals
- Partial reimbursement of public transport costs
- Possibility of teleworking (90 days per year) and flexible organization of working hours
- Partial payment of insurance costs

## Remuneration

monthly gross salary amounting to 2082 euros for the first and second years and 2190 euros for the third year

## General Information

- **Theme/Domain :** Architecture, Languages and Compilation
  System & Networks (BAP E)
- **Town/city :** Rennes
- **Inria Center :** Centre Inria de l'Université de Rennes
- **Starting date :** 2024-06-01
- **Duration of contract :** 3 years
- **Deadline to apply :** 2024-05-24

## Contacts

- **Inria Team :** TARAN
- **PhD Supervisor :**
  Kritikakou Angeliki / angeliki.kritikakou@irisa.fr

## About Inria

Inria is the French national research institute dedicated to digital science and technology. It employs 2,600 people. Its 200 agile project teams, generally run jointly with academic partners, include more than 3,500 scientists and engineers working to meet the challenges of digital technology, often at the interface with other disciplines. The Institute also employs numerous talents in over forty different professions. 900 research support staff contribute to the preparation and development of scientific and entrepreneurial projects that have a worldwide impact.

**Warning** : you must enter your e-mail address in order to save your application to Inria. Applications must be submitted online on the Inria website. Processing of applications sent from other channels is not guaranteed.

## Instruction to apply

Please submit online : your resume, cover letter and letters of recommendation eventually

For more information, please contactangeliki.kritikakou@irisa.fr

**Defence Security :**
This position is likely to be situated in a restricted area (ZRR), as defined in Decree No. 2011-1425 relating to the protection of national scientific and technical potential (PPST).Authorisation to enter an area is granted by the director of the unit, following a favourable Ministerial decision, as defined in the decree of 3 July 2012 relating to the PPST. An unfavourable Ministerial decision in respect of a position situated in a ZRR would result in the cancellation of the appointment.

**Recruitment Policy :**
As part of its diversity policy, all Inria positions are accessible to people with disabilities.