



Offer #2025-08872

Post-Doctoral Research Visit F/M [DRI Campaign] - Exploring Socio-Cultural Bias in a Multilingual Context

Contract type : Fixed-term contract

Renewable contract : Yes

Level of qualifications required : PhD or equivalent

Fonction : Post-Doctoral Research Visit

Context

Context

The goal of this project is to address key issues in Large Language Models (LLMs), particularly cultural biases stemming from Western-centric training data. These models often underperform or exhibit prejudice in non-English and especially South American contexts due to limited resources for bias detection. We propose to define sociologically grounded notions of social bias that can be computationally identified and measured. This involves data collection, annotation, and adapting existing datasets. Once biases are defined, we will detect them through model behavior analysis and apply cutting-edge fact-editing techniques to adjust the model's internal weights, mitigating harmful stereotypes while enhancing culturally relevant knowledge. Our focus is on general, multilingual methods, with a key application to Latin American languages and cultural contexts.

Is regular travel foreseen for this post ?

Regular travels to Chile are planned and fully funded.

Computing resources

Access to high-level Inria HPC clusters is granted. Access to the Jean Zay and Adastra HPC cluster will be sought.

Assignment

Assignments :

With the help of the project team, and especially Djamé Seddah, the recruited person will be taken to develop techniques able to ease the exploration and the mitigation of socio-cultural biases in current Large Language Models. An important part of the project will be devoted to the development of new benchmarks able to assess the exact sensitivity of models to culturally-based biases.

Collaborations :

The person recruited will work in close liaison with (i) other members of the Almanach team involved in the SaLM project (<https://salmproject.github.io/>) working on adjacent domains of biases detection and neutralization, (ii) members of Universidad de Chile (led by Valentin Barrière) and Inria Chile (led by Luis Marti, Nayat Sanchez Pi)

Main activities

- Establishing the state of the art of relevant techniques to the project
- Software development
- Conducts of experiments
- Preparation of reports and publications
- Release as open source the resulting software and resources

Skills

Technical skills and level required :

Excellent level of python + all tools needed for data science, deep learning and natural language processing. Experience with large scale experiments

Languages :

Fluent English. Knowledge of Spanish, especially the variant spoken in Latin-America as well as their regional socio-cultural contexts would be a plus.

Relational skills :

Team worker, autonomous and eager to disseminate results.

Other values appreciated : Enthusiasm, team spirit.

Benefits package

- Subsidized meals
- Partial reimbursement of public transport costs
- Leave: 7 weeks of annual leave + 10 extra days off due to RTT (statutory reduction in working hours) + possibility of exceptional leave (sick children, moving home, etc.)
- Possibility of teleworking and flexible organization of working hours
- Professional equipment available (videoconferencing, loan of computer equipment, etc.)
- Social, cultural and sports events and activities
- Access to vocational training
- Social security coverage

Remuneration

Salary: 2927€ (Brut) per month

General Information

- **Theme/Domain :** Language, Speech and Audio
- **Town/city :** Paris
- **Inria Center :** Centre Inria de Paris
- **Starting date :** 2025-11-01
- **Duration of contract :** 12 months
- **Deadline to apply :** 2025-06-01

Contacts

- **Inria Team :** ALMANACH
- **Recruiter :**
Seddah Djame / Djame.Seddah@inria.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.

The keys to success

Autonomy, dynamism and thoroughness.

A real interest in language and its challenges.

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

Deadline for applying: June 1st 2025

Join with your application:

- Detailed CV with a description of the PhD and a complete list of publications with the two most significant ones highlighted
- A Motivation letter
- 2 letters of recommendations
- Passport copy

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.