



Offer #2024-07756

PhD Position F/M A multi-modal language model for Earth observation

Contract type : Fixed-term contract

Level of qualifications required : Graduate degree or equivalent

Fonction : PhD Position

About the research centre or Inria department

Inria is a national research institute dedicated to digital sciences that promotes scientific excellence and transfer. Inria employs 2,400 collaborators organised in research project teams, usually in collaboration with its academic partners.

This agility allows its scientists, from the best universities in the world, to meet the challenges of computer science and mathematics, either through multidisciplinary or with industrial partners.

A precursor to the creation of Deep Tech companies, Inria has also supported the creation of more than 150 start-ups from its research teams. Inria effectively faces the challenges of the digital transformation of science, society and the economy.

Context

This PhD offer is funded by the GEO-ReSeT ANR project, representing a collaboration between Inria (team [EVERGREEN](#), Montpellier) and Université de Paris Cité (team [LIPADE](#), Paris).

Leveraging the large amounts of available geo-spatial data from different sources, the [GEO-ReSeT](#) (Generalized Earth Observation with Remote Sensing and Text) project has the objective to learn a rich representation of any geo-spatial location and convey a semantic representation of the information, by improving on existing models and providing a better experience to the end users. By using location on the Earth's surface as the common link between different modalities, a geo-spatial foundation model would be able to incorporate a variety of data sources, including remote sensing imagery, textual descriptions of places, and other generic features.

Such a foundation model has the potential to open a set of all new possibilities in terms of Earth observation applications, by allowing for few or zero-shot solutions to classical problems such as land-cover and land-use mapping, target detection, and visual question answering. It will also be useful for a wide range of applications with a geo-spatial component, including environmental monitoring, urban planning and agriculture.

By leveraging several data modalities, this foundation model could provide a comprehensive and accurate understanding of the Earth's surface, enabling informed decisions and actions. This will be particularly valuable for new potential users in sectors such as journalism, social sciences or environmental monitoring, who may not have the resources or expertise to collect their own training datasets and develop their own methods, thus moving beyond open Earth observation data and democratizing the access to Earth observation information.

Assignment

The work to be conducted during the proposed PhD thesis will contribute to the ambition of the GEO-ReSeT ANR project by linking textual descriptions of places (e.g., collected from heterogeneous online sources, such as news articles or search engine results), to their approximate geo-location, a task known as geoparsing.

This text-location link will then be used in combination with other geospatial data modalities, with a focus on remote sensing data from sensors such as Sentinel-1 and -2, in order to train multi-modal models that are aware about the way in which people describe locations.

This will be done by first combining information stemming from different databases containing geographic named entities, such as Open Street Map, Wikipedia and gazetteers, such that geographic points or polygons can be linked to each named entity.

In a second step, a Natural Language Processing (NLP) pipeline will be developed to obtain the most likely geographic named entities that are referred to in any piece of text that describes a place.

With respect to existing Named Entity Recognition (NER) methodologies, in order to avoid restricting us to cases where entities' names appear exactly as in the databases or gazetteers, we will leverage pre-trained Large Language Models (LLM) to resolve ambiguities and gather evidence towards the most

likely entities that are being described in the text. Such an approach will be trained and validated by using the cases that do match the names in the gazetteer.

We will then move on, in collaboration with the rest of the GEO-ReSeT consortium, to train a multi-modal large language model (MMLLM) that will serve as a foundation model for Earth observation tasks.

This model will finally be evaluated on several agro-environmental tasks.

Main activities

- Description of the state-of-the-art in unstructured text geoparsing, with a focus on approaches leveraging LLMs.
- Collection of a database of geographic named entities linked to their geographic footprint (e.g. point or polygon). Collection of a database of unstructured online text that is likely to contain a reference to a geographic location.
- Development of an NLP pipeline to link each piece of geographic text to its likely geographic footprint.
- Participate in the design and training of a multi-modal large language model (MMLLM) using remote sensing and geoparsed text.
- Evaluation of the final model on two of the following case studies at a national or continental scale: ecosystem type mapping, crop type mapping or land-use mapping.

Skills

- Python programming.
- Deep Learning with Python (preferably with Pytorch).
- Experience with NLP.
- Experience with GIS would be a plus.

Benefits package

- Subsidized meals
- 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
- Social, cultural and sports events and activities
- Access to vocational training

Remuneration

Gross Salary per month: 2010€ brut per month (year 1 & 2) and 2190€ brut per month (year 3)

General Information

- **Theme/Domain** : Language, Speech and Audio Information system (BAP E)
- **Town/city** : Montpellier
- **Inria Center** : [Centre Inria d'Université Côte d'Azur](#)
- **Starting date** : 2024-10-01
- **Duration of contract** : 3 years
- **Deadline to apply** : 2024-06-30

Contacts

- **Inria Team** : [EVERGREEN](#)
- **PhD Supervisor** :
Marcos Gonzalez Diego / diego.marcos@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

We are looking for someone with strong competences in Python programming and Deep Learning, ideally with experience with geospatial data and NLP. A strong motivation towards using these skills for tackling problems related to environmental monitoring is appreciated.

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

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.