



Offer #2022-05326

Post-Doctoral Research Visit F/M Software Heritage: large-scale empirical analysis of open source code artifacts

Contract type : Fixed-term contract

Renewable contract : Yes

Level of qualifications required : PhD or equivalent

Fonction : Post-Doctoral Research Visit

Context

This postdoc position is open at Inria in the context of the [Software Heritage](#) project, on the topic of **large-scale empirical analysis of open source code artifacts**, such as source code files and commits, as captured by state-of-the art distributed version control systems (VCSs).

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.

Software Heritage is a unique initiative to build the universal archive of software source code, catering for the needs of research, industry and society as a whole.

Assignment

Assignments:

With the help of the Software Heritage team, the recruited person will work on the analysis of the Software Heritage graph dataset (see the article online at <https://dl.acm.org/citation.cfm?id=3341907>), the largest publicly available corpus of Free and Open Source Software (FOSS) development history.

Main activities

To this end, the recruited person will:

- Exploit the compressed representation of the Software Heritage graph (see the article online at <http://dx.doi.org/10.1109/SANER48275.2020.9054827>) to conduct empirical analyses on subsets of interest of the Software Heritage archive.
- Propose, implement, and validate experimentally novel ways of exploiting the Software Heritage archive to conduct similar analyses in the future.
- Produce and curate open datasets mixing and matching software artifacts from Software Heritage and related data sources.

The research activity will take place in a multidisciplinary team involving computer scientists and industry leaders with the purpose of analyzing free/open source software at the scale of Software Heritage. It will also involve documentation of the work and results in the form of conference proceedings and journal papers, and the presentation of the results at scientific meetings.

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

General Information

- **Town/city** : CRI de Paris
- **Inria Center** : [Siège](#)
- **Starting date** : 2023-01-01
- **Duration of contract** : 12 months
- **Deadline to apply** : 2023-03-15

Contacts

- **Inria Team** : DGD-I (DGD-I)
- **Recruiter** :
Dupre Laurence / Laurence.Dupre@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

The ideal candidate will have obtained a PhD degree within the last 3 years, in computer science, applied math, computational science, or a related technical field, and will have proven experience in:

- conducting experiments in the field of empirical software engineering
- mining software artifacts from free/open source software repositories

It is expected that the candidate will have proven consensus builder in a highly collaborative environment and excellent written and oral communication skills.

Will be considered a plus:

- experience with graph compression/analysis techniques
- experience with machine learning and "big code" analysis
- participation in Free/Open Sources Software development

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

CV et lettre de motivation

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.