



Offer #2024-08452

Data engineer

Contract type : Fixed-term contract

Renewable contract : Yes

Level of qualifications required : PhD or equivalent

Fonction : Temporary scientific engineer

Corps d'accueil : Ingénieur de Recherche (IR)

Level of experience : From 5 to 12 years

Context

[Software Heritage](#) is a universal software source code archive project, whose aim is to recover, preserve for the very long term and share all publicly available source code, together with its development history (e.g., as stored in version control systems). The Software Heritage archive already contains over 19 billion unique source files and 4.2 billion commits, retrieved from over 300 million software development projects. The Software Heritage initiative, hosted by the Inria Foundation, is an entirely free software (FOSS) and non-profit project.

Assignment

We are looking for an experienced Big Data-oriented software engineer. The ideal candidate will have significant interest and experience in large-scale data processing and exploitation architectures, including storage, indexing and retrieval.

You can consult a more detailed list of our current projects on the Software Heritage Roadmap 2024 (<https://docs.softwareheritage.org/devel/roadmap/roadmap-2024.html>)

Main activities

- Setting up a data processing architecture (a la Spark)
- Design and modeling of Big Data architectures
- Implementation of solutions based on defined architectures
- Set up Big Data pipelines

Skills

The ideal candidate will have experience in Big Data development and architecture, preferably in an open-source context. We expect self-organization and autonomy skills commensurate with the candidate's experience. Participation in existing FOSS projects in any capacity (developer, community organizer, technical writer, etc.) is an added advantage.

The following skills are expected:

- Mastery of a large-scale data processing system (e.g. Apache Spark, Flink, or Hadoop)
- Fluent software development skills (basics in Rust and Python)
- Good level of English (written and spoken)
- Use of Git
- Use of continuous integration tools (e.g. Gitlab and/or Jenkins)

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 (after 6 months of employment) 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

Remunerating based on diploma and professional experience

General Information

- **Town/city** : Paris
- **Inria Center** : [Siège](#)
- **Starting date** : 2025-03-01
- **Duration of contract** : 3 years
- **Deadline to apply** : 2025-01-10

Contacts

- **Inria Team** : 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

Knowledge and experience of the following will be considered an asset:

- Experience in data processing on a scale of tens of terabytes or even petabytes
- Experience with Cassandra and Kafka
- Knowledge of Java
- Knowledge of Kubernetes
- Data visualization

Software Heritage is a complex technical architecture, based on many different technologies, which continues to evolve. We do not expect candidates to master all of them, but rather to be open to discovery and learning. Prior knowledge of one or more of the above-mentioned subjects will help in the process of getting to grips with the project, but we encourage you to apply whatever your level of experience in these technologies.

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 and cover letter required

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.