



Offre n°2025-08693

Data engineer

Type de contrat : Fixed-term contract

Contrat renouvelable : Oui

Niveau de diplôme exigé : PhD or equivalent

Fonction : Temporary scientific engineer

Corps d'accueil : Ingénieur de Recherche (IR)

Niveau d'expérience souhaité : From 5 to 12 years

Contexte et atouts du poste

[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.

Mission confiée

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->

Principales activités

- 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

Compétences

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)

Avantages

- 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

Rémunération

Remunerating based on diploma and professional experience

Informations générales

- **Ville** : Paris
- **Centre Inria** : [Siège](#)
- **Date de prise de fonction souhaitée** : 2025-03-01
- **Durée de contrat** : 3 years
- **Date limite pour postuler** : 2025-06-30

Contacts

- **Équipe Inria** : DGD-I
- **Recruteur** :
Dupre Laurence / Laurence.Dupre@inria.fr

A propos d'Inria

Inria est l'institut national de recherche dédié aux sciences et technologies du numérique. Il emploie 2600 personnes. Ses 215 équipes-projets agiles, en général communes avec des partenaires académiques, impliquent plus de 3900 scientifiques pour relever les défis du numérique, souvent à l'interface d'autres disciplines. L'institut fait appel à de nombreux talents dans plus d'une quarantaine de métiers différents. 900 personnels d'appui à la recherche et à l'innovation contribuent à faire émerger et grandir des projets scientifiques ou entrepreneuriaux qui impactent le monde. Inria travaille avec de nombreuses entreprises et a accompagné la création de plus de 200 start-up. L'institut s'efforce ainsi de répondre aux enjeux de la transformation numérique de la science, de la société et de l'économie.

L'essentiel pour réussir

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.

Attention: Les candidatures doivent être déposées en ligne sur le site Inria. Le traitement des candidatures adressées par d'autres canaux n'est pas garanti.

Consignes pour postuler

CV and cover letter required

Sécurité défense :

Ce poste est susceptible d'être affecté dans une zone à régime restrictif (ZRR), telle que définie dans le décret n°2011-1425 relatif à la protection du potentiel scientifique et technique de la nation (PPST). L'autorisation d'accès à une zone est délivrée par le chef d'établissement, après avis ministériel favorable, tel que défini dans l'arrêté du 03 juillet 2012, relatif à la PPST. Un avis ministériel défavorable pour un poste affecté dans une ZRR aurait pour conséquence l'annulation du recrutement.

Politique de recrutement :

Dans le cadre de sa politique diversité, tous les postes Inria sont accessibles aux personnes en situation de handicap.