Offre n°2025-08693

Ingénieur de données H/F

Type de contrat : CDD

Contrat renouvelable : Oui

Niveau de diplôme exigé : Thèse ou équivalent

Fonction : Ingénieur scientifique contractuel

Corps d'accueil : Ingénieur de Recherche (IR)

Niveau d'expérience souhaité : De 5 à 12 ans

Contexte et atouts du poste

Software Heritage est un projet d'archive universelle du code source logiciel, dont l'objectif est de récupérer, préserver dans le très long terme et partager la totalité du code source qui est publiquement disponible, avec l'historique de son développement (p.ex., comme stocké dans des systèmes de gestion de version). L'archive de Software Heritage contient déjà plus de 19 milliards des fichiers source uniques et 4,2 milliards de commits, récupérés à partir de plus de 300 millions de projets de développement logiciel. L'initiative Software Heritage, hébergée par la Fondation Inria est un projet entièrement logiciel libre (FOSS) et à but non lucratif.

Mission confiée

Nous recherchons un·e ingénieur·e expérimenté·e logiciel orienté·e Big Data. Le·a candidat·e idéal·e sera une personne ayant un un intérêt et une expérience significatifs pour les architectures de traitement et d'exploitation de données à grande échelle, incluant le stockage, l'indexation et l'extraction.

Vous pouvez consulter une liste détaillée de nos projets en cours sur la Roadmap 2024 de Software Heritage (https://docs.softwareheritage.org/devel/roadmap/roadmap-2024.html)

Principales activités

– Mise en place d'une architecture de traitement de données (à la Spark)
– Conception et modélisation d'architectures Big Data
– Implémentation de solutions sur la base des architectures définies
– Mise en place de pipelines big data

Compétences

Le candidat idéal doit avoir de l'expérience dans le développement et l'architecture Big Data, de préférence dans le contexte du logiciel libre. Nous attendons des compétences d'auto-organisation et d'autonomie proportionnelles à l'expérience du candidat. La participation à des projets FOSS existants, à quelque titre que ce soit (développeur·se, organisateur communautaire, rédacteur technique, etc.) constitue un atout supplémentaire.

Les compétences suivantes sont attendues:

– Maîtrise d'un système de traitement de données à grande échelle (par exemple Apache Spark, Flink, ou Hadoop)
– Pratique courante du développement logiciel (Bases en Rust et Python)
– Bon niveau d'anglais (écrit et oral)
– Utilisation de Git
– Utilisation d'outils d'intégration continue (par exemple Gitlab et/ou Jenkins)

Avantages

- Restauration subventionnée
- Transports publics remboursés partiellement
- Congés: 7 semaines de congés annuels + 10 jours de RTT (base temps plein) + possibilité d'autorisations d'absence exceptionnelle (ex : enfants malades, déménagement)
- Possibilité de télétravail (après 6 mois d'ancienneté) et aménagement du temps de travail
- Équipements professionnels à disposition (visioconférence, prêts de matériels informatiques, etc.)
- Prestations sociales, culturelles et sportives (Association de gestion des œuvres sociales d'Inria)
- Accès à la formation professionnelle

Rémunération

En fonction du diplôme et de l'expérience professionnelle

Informations générales

- **Ville :** Paris
- **Centre Inria :** Siège
- **Date de prise de fonction souhaitée :** 2025-03-01
- **Durée de contrat :** 3 ans
- **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'e?orce ainsi de répondre aux enjeux de la transformation numérique de la science, de la société et de l'économie.

L'essentiel pour réussir

La connaissance et l’expérience des éléments suivants seront considérées comme un atout :

– Expérience dans le traitement de données à l’échelle de dizaines de tera-octets, voire peta-octets
– Expérience avec Cassandra et Kafka

– Connaissances en Java
– Connaissances Kubernetes
– Data visualization

Software Heritage est une architecture technique complexe, basée sur de nombreuses technologies différentes et qui continue à évoluer. Nous n'attendons pas que les candidat·es en maîtrisent l'intégralité, mais plutôt qu'iels soient disponibles à la découverte et à l'apprentissage. Une connaissance préalable de l'un ou plusieurs des sujets mentionnés auparavant sera une aide dans le processus de prise en main du projet, mais nous vous encourageons à candidater quelque soit votre niveau d'expérience dans ces 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 et lettre de motivation obligatoire

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.