Offer #2020-03153

Ingénieur IA : knowledge graph and linked data

The offer description below is in French

Contract type : Fixed-term contract

Level of qualifications required : Graduate degree or equivalent

Fonction : Temporary scientific engineer

About the research centre or Inria department

Le centre Inria Sophia Antipolis - Méditerranée compte 34 équipes de recherche, ainsi que 7 services d'appui à la recherche. Le personnel du centre (500 personnes environ dont 320 salariés Inria) est composé de scientifiques de différentes nationalités (250 personnes étrangères sur 50 nationalités), d'ingénieurs, de techniciens et d'administratifs. 1/3 du personnel est fonctionnaire, les autres sont contractuels. La majorité des équipes de recherche du centre est localisée à Sophia Antipolis et Nice dans les Alpes-Maritimes. **Quatre** équipes sont implantées à Montpellier et deux équipes sont hébergées l'une à Bologne et l'autre à Athènes. Inria est membre fondateur d'Université Côte d'Azur et partenaire de l'I-site MUSE porté par l'Université de Montpellier.

Context

Dans le cadre d'un partenariat entre l'équipe de recherche Wimmics (https://team.inria.fr/wimmics/) et la coopérative Mnémotix (https://www.mnemotix.com/) s'inscrivant dans le contexte du 3IA Côte d'Azur (https://3ia.univ-cotedazur.eu/), la mission consiste à poursuivre le développement de la bibliothèque sémantique open source CORESE (https://project.inria.fr/corese/) en vue de son industrialisation. Le poste sera exercé dans les locaux de INRIA Sophia Antipolis, au sein de l'équipe Wimmics. Quelques déplacements ponctuels pour rencontrer les partenaires peuvent être prévus.

Assignment

La plateforme sémantique CORESE consiste en une base de triplets (*triple store*) et une interface de requêtes implémentant les standards du Web Sémantique (RDF, RDFS, SPARQL) et proposant de nombreuses fonctionnalités innovantes, fruit de plusieurs années de recherches en IA symbolique et ingénierie des connaissances.

Avec l'aide de son architecte et principal développeur, la personne recrutée reprendra le développement de cette bibliothèque open source afin de la porter aux standards actuels de l'industrie et de développer sa communauté. Elle sera également en étroite collaboration avec les ingénieur·e·s de l'équipe Wimmics qui exploitent CORESE dans divers projets de recherche. Enfin une collaboration rapprochée avec l'équipe de la coopérative Mnémotix, et en particulier le titulaire de la chaire 3IA "affiliate", fournira un terrain plus directement ancré dans les problématiques de production.

Plus particulièrement, l'évolution du code sera pilotée par la levée des verrous technologiques identifiés et relevés par le consortium à l'origine de ce projet d'industrialisation. La liste suivante pose les grandes lignes des axes à développer sur les 2 premières années, et sera affinée au cours du projet:

- Persistance et indexation de gros graphes de connaissances (*knowledge graph*) pour leur traitement dans des algorithmes d'IA (raisonnement, apprentissage)
- Mode□les de repre□sentation, d'indexation et algorithmes de raisonnement passant a□ l'e□chelle pour le traitement intelligent de gros graphes de connaissances et de leurs ontologies.
- Gestion de l'e□volution et de temporalite□ de grands graphes de connaissances et de leurs ontologies pour l'efficacite□ et pour l'exploitation et l'analyse en IA de ces e□volutions.
- Me□thode d'IA pour la ge□ne□ration (semi-)automatique de patrons et connecteurs d'import et d'export vers d'autres formats et sources classiques afin de supporter l'he□te□roge□ne□ite□ des donne□es en entre□e des me□thodes d'IA
- Me□thodes d'IA (apprentissage, raisonnement) pour la ge□ne□ration d'alignements et de re□gles et contraintes de validation et le nettoyage des donne□es pour en assurer la qualite□.

La mission comprend également un objectif d'augmentation du nombre de contributeurs et usagers de CORESE. Le code source est actuellement publié sur un dépôt Github, et une partie du temps de travail devra être dédiée à l'amélioration de la documentation, la gestion des tickets, et plus généralement l'animation de la communauté autour du projet CORESE (liste de diffusion, annonces de nouvelles versions, recueil des besoins et suggestions, etc.).

Main activities

Principales activités:

- Analyser et tester le code actuel, et conduire des benchmarks
- Conception et architecture des évolutions selon les grandes lignes précédentes
- Développement et tests du code
- Rédiger la documentation
- Animer une communauté de contributeurs et usagers

Activités complémentaires:

- Organiser des temps de formation ou de présentation à la communauté
- Préparer des démonstrations
- Diffuser les actualités du projet

Skills

Compétences techniques et niveau requis :

- Java, niveau confirmé ou expert
- Maîtrise des standards du web sémantique (RDF, OWL, SPARQL)
- Maîtrise des principaux types de bases de données
- Bonnes notions sur les bases de graphes (RDF et/ou Property Graphs, etc.)

Langues :

- Anglais
- Français

Compétences relationnelles :

- Travail en équipe
- Autonomie
- Capacité de vulgarisation
- Travail sur projet ouvert et open source

Compétences additionnelles appréciées :

- Bonnes notions de l'architecture Web et de ses standards
- Bonnes notions de logique formelle
- Langages d'inférences
- Passage à l'échelle et gros volumes de données

Benefits package

- 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
- Sécurité sociale

General Information

- **Theme/Domain :** Data and Knowledge Representation and Processing
  Software Experimental platforms (BAP E)
- **Town/city :** Sophia Antipolis
- **Inria Center :** Centre Inria d'Université Côte d'Azur
- **Starting date :** 2021-02-01
- **Duration of contract :** 2 years
- **Deadline to apply :** 2021-01-20

Contacts

- **Inria Team :** WIMMICS
- **Recruiter :**
  Gandon Fabien / fabien.gandon@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

Vous avez un intérêt marqué dans la conception et l'écriture de logiciels complexes et vous maitrisez et appréciez le langage Java (ou Scala). Vous êtes très autonome et ne craignez pas de vous attaquer à des domaines légèrement en deçà de votre zone de confort, ni à des projets d'envergure. Vous êtes curieu·x·se et appréciez profondément apprendre continuellement de nouvelles choses. Vous avez idéalement une bonne connaissance des standards du Web Sémantique, ou à minima de bonnes notions sur les bases de graphes. Vous avez déjà des notions ou l'envie d'en apprendre plus sur les logiques de description, ou les systèmes d'inférences, ou vous savez peut-être déjà que le "chaînage avant" peut s'utiliser ailleurs que sur les roues d'une voiture en hiver.

Vous adorez travailler dans un environnement scientifiquement stimulant, où vous communiquez indifféremment en anglais ou en français, et interagissez avec plusieurs collaborateurs ou partenaires aux profils variés. Vous aimez communiquer sur votre travail, vous savez qu'un code non documenté est un code condamné, et vous êtes convaincu·e de l'intérêt de l'open source, et conscient de l'exigence de qualité qui en est son corollaire.

Vous voulez pouvoir dire aux prochains repas de famille d'être le contributeur principal d'un des triple store RDF open source de référence au niveau international.

Si vous vous retrouvez dans la plupart de ces points, n'attendez plus une seconde et envoyez-nous votre candidature.

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.