



Offre n°2025-08804

Optimisation des interactions avec les LLMs dans un environnement Python et expérimentation avec les approches symboliques

Type de contrat : Convention de stage

Niveau de diplôme exigé : Bac + 5 ou équivalent

Fonction : Stagiaire des fonctions support

A propos du centre ou de la direction fonctionnelle

Depuis 2018, Inria joue un rôle crucial en tant que support à l'action de l'Etat dans le domaine de l'intelligence artificielle, en collaboration étroite avec la Stratégie Nationale pour l'IA. Pour piloter cette mission, Inria a créé le Programme IA, qui s'inscrit dans le dispositif de l'Agence de Programme confié par l'État à Inria, et qui coordonne des dispositifs innovants transverses en IA : recherche, transfert et appui aux politiques publiques.

Dans le cadre de la stratégie d'accélération IA et en particulier dans l'axe « développer des infrastructures numériques de pointe et stimuler les partenariats public-privé et accroître l'impact de la recherche en IA », Inria a proposé la mise en œuvre d'un projet d'accélération visant à soutenir la mise en place et le développement d'une plateforme souveraine ouverte et interopérable de bibliothèques logicielles de l'IA pour les entreprises françaises, et son passage à l'échelle européenne.

Cette initiative est une des priorités d'Inria dans le cadre de l'élaboration de la stratégie d'accélération IA et dans l'objectif de contribuer à la souveraineté technologique française et à l'impact économique sur le tissu industriel français.

L'impact attendu de ces mesures sur l'industrie française est notamment de permettre à l'ensemble des entreprises françaises mais également aux acteurs publics et institutionnels d'accélérer leur transformation numérique à l'aide de l'IA, sans être dans une relation captive avec les acteurs dominants, de par la disponibilité d'un socle technologique générique et d'un ensemble de dispositifs de transfert associés.

Dans le cadre de cette ambition, la mise en œuvre de la mesure repose sur le développement et la diffusion d'une suite cohérente de plateformes de logiciels pour l'IA et les données reposant sur trois actions :

- L'action 1 se concentre sur l'interopérabilité des données, proposant une architecture basée sur les standards du Web (Corese).
- L'action 2, le "Data Wrangling," aborde la préparation des données, visant à réduire le temps manuel nécessaire dans les projets de science des données (skrub).
- L'action 3 se focalise sur l'apprentissage et l'exécution de modèles, en optimisant la bibliothèque scikit-learn pour faciliter l'analyse des données.

Contexte et atouts du poste

Dans le cadre du projet P16, nous explorons les interactions entre les grands modèles de langage (LLMs) et les environnements de développement Python, avec un accent particulier sur les enjeux d'interprétabilité et d'interfaçage avec les graphes de connaissances. L'objectif est d'identifier des cas d'usage concrets et d'expérimenter des approches pertinentes pour renforcer la robustesse, la précision et la fiabilité des outils exploitant les LLMs à des fins de prototypage

Mission confiée

Le projet vise à **optimiser la génération et la validation des réponses des LLMs** en développant des mécanismes de vérification d'incohérences, de gestion d'erreurs et d'ajustement contextuel automatique. Il s'agira également d'**expérimenter des modèles hybrides** combinant LLMs et algorithmes explicites pour améliorer l'interprétabilité et la robustesse des réponses.

Par ailleurs, un **pipeline d'interfaçage** entre les environnements de développement Python et les LLMs devra être conçu, garantissant la compatibilité des formats d'entrée/sortie et permettant de chaîner de manière modulaire les opérations de pré-traitement, de post-traitement et d'optimisation.

Enfin, le projet prévoit **l'intégration de graphes de connaissances** pour enrichir les interactions, en explorant l'alignement entre représentations vectorielles et structures symboliques, et en testant des scénarios combinant raisonnement logique et génération de texte afin d'améliorer la pertinence et la fiabilité des réponses produites.

Principales activités

Le ou la stagiaire aura pour mission de :

- **Optimiser la génération et la validation des réponses des LLMs :**
 - Développer des mécanismes de vérification des réponses des LLMs, de détection d'incohérences et d'ajustement automatique en fonction du contexte et des contraintes.
 - Intégrer des stratégies de gestion d'erreurs.
 - Expérimenter des modèles hybrides combinant LLMs et algorithmes plus explicites pour améliorer l'interprétabilité et la robustesse.
- **Concevoir un pipeline d'interfaçage entre environnements de développement Python et LLMs**
 - Assurer une compatibilité fluide entre les formats d'entrée/sortie des LLMs et les outils de traitement des données.
 - Mettre en place des pipelines modulaires permettant de chaîner différentes opérations de pré-traitement, de post-traitement et d'optimisation des interactions avec les LLMs.
- **Intégrer les graphes de connaissances pour enrichir les interactions avec les LLMs :**
 - Expérimenter des approches d'alignement entre représentations vectorielles des LLMs et structures symboliques des graphes (embedding de graphes, requête hybride SPARQL/LLM).
 - Tester des scénarios d'usage combinant raisonnements logiques et génération de texte pour renforcer la pertinence et la fiabilité des réponses.

Compétences

- Formation en informatique, intelligence artificielle, science des données ou domaine connexe.
- Solides compétences en Python et expérience avec les APIs des LLMs.
- Connaissance des graphes de connaissances et du web sémantique (RDF, SPARQL, OWL).
- Intérêt pour les approches hybrides combinant apprentissage machine et représentation symbolique.

Avantages

- Restauration subventionnée
- Transports publics remboursés partiellement
- Congés: 7 semaines de congés annuels + 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

Rémunération

Gratification de 4.35€ par heure

Informations générales

- **Ville :** Le Chesnay
- **Centre Inria :** [Siège](#)
- **Date de prise de fonction souhaitée :** 2025-06-01
- **Durée de contrat :** 6 mois
- **Date limite pour postuler :** 2025-05-31

Contacts

- **Équipe Inria :** Prog-IA (DG)
- **Recruteur :**
Arunraja Emilie / emilie.arunraja@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.

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

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.