



Offer #2025-08988

Ingénieur R&D en Interprétation et Annotation Textuelle de Code Optimisé (H/F)

The offer description below is in French

Contract type : Fixed-term contract

Level of qualifications required : Graduate degree or equivalent

Fonction : Temporary scientific engineer

Level of experience : Recently graduated

About the research centre or Inria department

Le Centre Inria de l'Université Grenoble Alpes, regroupe un peu moins de 600 personnes réparties au sein de 24 équipes de recherche et 9 services support à la recherche.

Son effectif est distribué sur 3 campus à Grenoble, en lien étroit avec les laboratoires et les établissements de recherche et d'enseignement supérieur (Université Grenoble Alpes, CNRS, CEA, INRAE, ...), mais aussi avec les acteurs économiques du territoire.

Présent dans les domaines du calcul et grands systèmes distribués, logiciels sûrs et systèmes embarqués, la modélisation de l'environnement à différentes échelles et la science des données et intelligence artificielle, le Centre Inria de l'Université Grenoble Alpes participe au meilleur niveau à la vie scientifique internationale par les résultats obtenus et les collaborations tant en Europe que dans le reste du monde.

Context

Dans le cadre d'un projet de recherche innovant, nous recherchons un ingénieur R&D pour une durée d'un an. Passionné par la compréhension fine du code machine et son interprétation, ce poste s'inscrit directement dans la problématique de la thèse de doctorat intitulée "Compréhension et Alignement Sémantique du Code Super-Optimisé pour le Calcul Haute Performance et le Deep Learning".

Notre outil existant, GUS, est une plateforme de débogage de performance qui fournit des métriques et des informations détaillées sur le code assembleur final. Ces données sont cruciales, mais leur interprétation est rendue complexe par les

transformations agressives des compilateurs modernes, notamment en calcul haute performance et deep learning.

Le cadre d'utilisation de GUS est ici focalisé sur la super-optimisation de programmes de deep learning menée dans les projets Holigrail (PEPR IA) et Deepgreen (BPI). Ces projets visent à développer une infrastructure logicielle combinant deux stratégies complémentaires pour améliorer l'efficacité des réseaux de neurones profonds. D'une part, la compression du réseau, qui inclut des techniques telles que la quantisation et la factorisation, modifie la structure et le comportement du réseau afin de réduire sa taille et ses besoins en calcul. D'autre part, l'optimisation de compilation, qui repose exclusivement sur des transformations de programme sans altérer la logique du modèle, permet d'adapter et d'améliorer l'exécution du code pour maximiser ses performances sur l'architecture ciblée.

Assignment

L'objectif principal de ce poste est de rendre les informations brutes de GUS plus lisibles, intelligibles et explorables pour les développeurs, en enrichissant son interface utilisateur par des traitements et des présentations textuelles contextualisées et interactives. Et ce cadre n'est pas restrictif, car il n'est pas exclu d'enrichir les sorties de GUS en les fusionnant avec la sortie avec d'outils comme Perf.

Main activities

Sous la supervision de l'équipe de recherche, l'ingénieur aura les responsabilités clés suivantes :

Conception et implémentation d'annotations textuelles avancées (priorité haute)
Analyser et prioriser les besoins des développeurs pour la compréhension des métriques de performance et des informations d'assembleur issues de GUS.
Développer des mécanismes d'annotation textuelle pour enrichir l'affichage de l'assembleur brut. Ces annotations incluront l'intégration directe de métriques de performance (ex: cycles CPU, cache misses) à côté des instructions concernées.
Mettre en œuvre des techniques pour reconstruire et afficher textuellement des abstractions de haut niveau (le "raising conceptuel" de la thèse). Cela pourrait se traduire par l'identification et l'annotation de structures de contrôle (boucles optimisées), de blocs fonctionnels, ou de patterns d'accès mémoire spécifiques au deep learning.

Établir et afficher textuellement le lien sémantique avec le code source original ("l'alignement conceptuel"). Cela implique des références textuelles, des numéros de ligne, ou des portions de code source annotées directement dans la vue assembleur.

Développement d'une interface textuelle interactive

Concevoir et implémenter des mécanismes d'interaction basés sur le texte pour naviguer au sein des annotations et des différentes couches d'information. Par

exemple, des liens cliquables dans le texte pour “déplier” des détails, des filtres pour afficher ou masquer certains types d’annotations, ou des fonctions de recherche contextuelle.

L’objectif est de permettre une exploration guidée et approfondie des données sans quitter la vue textuelle.

Documentation et veille applicative

Documenter de manière exhaustive les nouvelles fonctionnalités, les formats d’annotations et les choix techniques.

Effectuer une veille ciblée sur les approches innovantes d’analyse de texte pour le code, de rétro-ingénierie textuelle et de présentation d’informations techniques complexes.

Skills

- Compétences techniques indispensables :
 - Excellente maîtrise du langage C/C++ (pour l’intégration avec le cœur de GUS et la manipulation du code machine).
 - Solide compréhension des bases de l’assembleur et du fonctionnement interne des compilateurs (chaîne d’optimisation).
 - Forte aptitude à la manipulation, l’analyse et la présentation de données textuelles structurées.
 - Expérience en développement d’interfaces utilisateur en mode texte (CLI enrichies) ou utilisant des frameworks légers permettant l’interactivité dans un environnement non graphique.
- Compétences techniques appréciées :
 - Connaissance des outils de profilage de performance (Perf, Intel VTune, etc.).
 - Notions en analyse de binaire (outils comme Ghidra, IDA Pro) et/ou en ingénierie inverse.
 - Compréhension des enjeux liés au calcul haute performance ou au deep learning.
 - Expérience avec des bibliothèques de parsing ou de manipulation de graphes de dépendance.
- Qualités personnelles :
 - Rigueur et précision dans l’analyse et l’annotation des informations techniques.
 - Autonomie et capacité à prendre des initiatives pour proposer des solutions innovantes.
 - Curiosité pour les problématiques d’optimisation de code et d’interprétation de sa sémantique.
 - Excellentes capacités de communication écrite pour documenter et expliquer des concepts complexes.

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

Remuneration

A partir de 2 692 € brut mensuel selon expérience et diplômes.

General Information

- **Theme/Domain** : Architecture, Languages and Compilation Scientific computing (BAP E)
- **Town/city** : Grenoble
- **Inria Center** : [Centre Inria de l'Université Grenoble Alpes](#)
- **Starting date** : 2025-10-01
- **Duration of contract** : 12 months
- **Deadline to apply** : 2025-07-10

Contacts

- **Inria Team** : [CORSE](#)
- **Recruiter** :
Huard Guillaume / guillaume.huard@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

- **Formation** : Diplôme d'ingénieur ou Master 2 en informatique, avec une spécialisation ou un intérêt marqué pour la compilation, l'architecture des systèmes informatiques, l'analyse de code et le traitement ou la présentation

de données textuelles complexes.

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.