



Offre n°2024-08309

## CDD Ingénieur de recherche - Coordinateur de projet scientifique pour le défi COLaF (H/F)

Type de contrat : CDD

Niveau de diplôme exigé : Thèse ou équivalent

Fonction : Ingénieur scientifique contractuel

### Contexte et atouts du poste

À travers le projet COLaF (Corpus et Outils pour les Langues de France), [Inria](#) a pour objectif de contribuer au **développement de corpus et d'outils libres** pour le **français** et les autres **langues de France**, en étroite collaboration avec des [partenaires](#) académiques et institutionnels.

Le périmètre de COLaF inclut à la fois :

- les **données textuelles** ([ALMAnaCH](#), Centre Inria de Paris),
- les **données de parole** et de langue des signes ([MULTISPEECH](#), Centre Inria de l'Université de Lorraine).

COLaF vise à couvrir la diversité du français et des langues de France :

- il a pour ambition de couvrir un échantillon aussi divers que possible: français de France et d'ailleurs, langues régionales, créoles à base française (y compris hors de France), langues autochtones, langues d'immigration, langue des signes française;
- l'ensemble des axes de variation seront étudiés, au-delà de l'état standard de la langue, et notamment les langues de spécialité, la diachronie, les états non standard (réseaux sociaux, langue d'apprenants, etc.).

Les travaux au sein du projet couvrent notamment l'acquisition et structuration de textes à partir de sources non textuelles (livres, enregistrements audio, etc.), la classification par langues et par variétés linguistiques de gros volumes de textes (en lien étroit avec le projet [OSCAR](#)), le développement de modèles d'annotation et de transformation (traduction, normalisation, synthèse vocale, génération de langue des signes) au service du développement de corpus et de l'exploitation des ressources nouvellement créées.

COLaF est un [DEFI](#) Inria porté par [Benoît Sagot](#) (responsable de l'équipe-projet ALMAnaCH) et [Slim Ouni](#) (responsable de l'équipe-projet MULTISPEECH).

### Mission confiée

La personne recrutée sera en charge de la coordination de l'équipe ALMAnaCH autour du DEFI et de la coordination avec les partenaires internes (MULTISPEECH) et externes (scientifiques et institutions spécialistes de langue de France, institutions patrimoniales, etc.).

Elle participera à la constitution de corpus, à la gestion des corpus (en particulier OSCAR) et aura à sa charge la gestion des ingénieur-e-s.

Il est aussi attendu de la personne concernée de participer à la vie scientifique de l'équipe.

### Principales activités

Principales activités :

- Identifier les acteurs avec lesquels l'équipe ALMAnaCH est amenée à collaborer dans le cadre du DEFI, prendre contact avec eux, gérer ces partenariats.
- Organiser les événements du DEFI en partenariat avec MULTISPEECH.
- Gérer le projet dans son quotidien.
- Identifier, produire et maintenir des corpus en collaboration avec les ingénieur-e-s, doctorant-e-s et chercheur-se-rs permanent-e-s du projet.

Activités secondaires:

- Conduire une recherche autour du traitement automatique des langues ou des humanités computationnelles.
- Participer à la vie de l'équipe, par delà le DEFI.

## Compétences

Compétences techniques et niveau requis :

- Expérience académique en algorithmes de traitement automatique des langues
- Des compétences en humanités numériques et/ou en linguistique (descriptive, historique) seraient des plus
- Une expérience dans une équipe de recherche serait également un plus

Langues : français natif ou courant ; anglais professionnel ; la maîtrise d'une langue de France autre que le français serait un plus.

## 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
- Aménagement du temps de travail (après 12 mois d'ancienneté)
- 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

## Informations générales

- **Thème/Domaine** : Langue, parole et audio  
Ingénierie logicielle (BAP E)
- **Ville** : Paris
- **Centre Inria** : [Centre Inria de Paris](#)
- **Date de prise de fonction souhaitée** : 2025-02-01
- **Durée de contrat** : 2 ans
- **Date limite pour postuler** : 2025-01-10

## Contacts

- **Équipe Inria** : [ALMANACH](#)
- **Recruteur** :  
Sagot Benoit / [Benoit.Sagot@inria.fr](mailto:Benoit.Sagot@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

Se sentir à l'aise dans un environnement de dynamique scientifique et à l'interface entre recherche en TAL et développement de ressources linguistiques, mais également aimer apprendre et relever des défis complexes et avoir le goût du travail bien fait sont des qualités essentielles pour réussir cette mission.

**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.

