



Offre n°2021-04141

Doctorant F/H Corpus, méthodes et ressources pour la transcription automatique des documents manuscrits patrimoniaux contemporains

Type de contrat : CDD

Niveau de diplôme exigé : Bac + 5 ou équivalent

Autre diplôme apprécié : Double profil sciences humaines et informatique

Fonction : Doctorant

Niveau d'expérience souhaité : De 3 à 5 ans

Contexte et atouts du poste

Dans le cadre d'un partenariat avec l'Université de Montréal, l'objectif est de développer un projet doctoral en co-direction entre Paris et Montréal visant à systématiser, au plan théorique et au plan pratique, ce qui fait la spécificité des sciences humaines comme domaine d'application de la reconnaissance des écritures manuscrites.

Mission confiée

Poursuivre un projet doctoral qui s'intègre bien aux différents projets de l'équipe ALMAAnaCH dans le domaine de la reconnaissance automatique d'écriture manuscrite et qui permette des avancées scientifiques dans le domaine de la compréhension des difficultés inhérentes au travail sur des documents patrimoniaux ainsi que des perspectives (vérités de terrain, métriques, protocoles d'apprentissage, évolution des plates-formes) permettant d'y remédier.

Principales activités

Effectuer un état de l'art sur l'état des recherches en paléographie numérique

Formaliser une description des écritures représentées dans les corpus des XIXe et XXe siècles

Proposer une classification permettant de mesurer les faiblesses d'un jeu de données ou d'un modèle de transcription en fonction des catégories d'écritures présentes dans un document

Offrir des pistes permettant de standardiser une telle classification et d'intégrer ces informations dans les métadonnées des corpus d'entraînement ou édités

Tester les performances des modèles de transcription obtenus par l'intermédiaire des logiciels actuellement disponibles en fonction de la constitution des corpus

Compétences

Très bonnes connaissances des enjeux du déploiement de méthodes numériques en contexte patrimonial

Connaissance fine des méthodes de reconnaissance automatique d'écriture manuscrite

Capacités avancées de programmation

Langues: excellente maîtrise orale et écrite de l'anglais

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

Informations générales

- **Thème/Domaine** : Langue, parole et audio
Production, traitement et analyse des données (BAP D)
- **Ville** : Paris
- **Centre Inria** : [Centre Inria de Paris](#)
- **Date de prise de fonction souhaitée** : 2021-11-01
- **Durée de contrat** : 3 ans
- **Date limite pour postuler** : 2021-10-24

Contacts

- **Équipe Inria** : [ALMANACH](#)
- **Directeur de thèse** :
Romary Laurent / Laurent.Romary@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.