



Offre n°2021-03252

Stage : Construction d'un vocabulaire de brevets by means of deep learning

Type de contrat : Stage

Niveau de diplôme exigé : Bac + 5 ou équivalent

Fonction : Stagiaire des fonctions support

Niveau d'expérience souhaité : De 3 à 5 ans

Contexte et atouts du poste

Le Startup Studio Inria propose un stage pour le projet qatent au sein de l'équipe de recherche ALmanaCH.

Il s'agit de la **Création d'un vocabulaire technique multimots par croisement du vocabulaire extrait de Wikipédia et d'autres ressources scientifiques avec un corpus de brevets.**

Mission confiée

L'essentiel pour réussir

- Il est important d'être opérationnel en python et outils TAL d'apprentissage profond (Spacy, Pytorch, ...)
- Aussi : Vous êtes passionné(e) par l'étude de la langue et par le traitement automatique de la langue

Début du contrat idéal : début mars.

Principales activités

- Lire <https://stackoverflow.com/questions/61218518/count-frequency-of-multi-word-terms-in-large-texts-with-python/61293305#61293305>
- Extraction de phrases définitives, analyse syntaxique, extraction de termes et d'hyponymes par simplification de syntagmes. Lire <https://towardsdatascience.com/unsupervised-synonym-harvesting-d592eaaf3c15>
- Ajout de vocabulaire multi-mots par fréquence et entropie.
- Utilisation des informations portant sur les classes de brevets (p. ex. CPC) pour la définition d'un vocabulaire dépendant du domaine technique.
- Compléter le vocabulaire à l'aide de plongements multi-mots
- Développement d'une taxonomie de terminologie
- Intégration de la taxonomie dans un API Django

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

- Ville : Paris
- Centre Inria : [Centre Inria de Paris](#)
- Date de prise de fonction souhaitée : 2021-03-01
- Durée de contrat : 6 mois
- Date limite pour postuler : 2021-02-28

Contacts

- Équipe Inria : INCUB-PRO (DGD-I)
- Recruteur :
Sagot Benoit / 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.

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.