# FICHE DE POSTE

**Libellé de Fonctions : Ingénieur(e) en humanités numériques et traitement automatique des langues**

**Catégorie : A**

**Corps : IGE / IGR**

**BAP (le cas échéant) :**

**Emploi-type (le cas échéant) :**

*Les activités qui composent la fiche de poste sont appelées à évoluer en fonction des connaissances du métier et des nécessités de service*

## Qui sommes-nous ?

Sorbonne Université est une université pluridisciplinaire de recherche créée au 1er janvier 2018 par regroupement des universités Paris-Sorbonne et UPMC.

Déployant ses formations auprès de 54 000 étudiants dont 4 700 doctorants et 10 200 étudiants étrangers, elle emploie 6 300 enseignants, enseignants-chercheurs et chercheurs et 4 900 personnels de bibliothèque, administratifs, technique, sociaux et de santé. Son budget est de 670 M€.

Sorbonne Université dispose d'un potentiel de premier plan, principalement situé au cœur de Paris, et étend sa présence dans plus de vingt sites en Île-de-France et en régions.

Sorbonne Université présente une organisation originale en trois facultés de lettres, de sciences & ingénierie et de médecine qui disposent d'une importante autonomie de mise en œuvre de la stratégie de l'université dans leur périmètre sur la base d'un contrat d'objectifs et de moyens. La gouvernance universitaire se consacre prioritairement à la promotion de la stratégie de l'université, au pilotage, au développement des partenariats et à la diversification des ressources.

## Présentation de la structure

### L'ObTIC :

L'Observatoire des textes, des idées et des corpus (ObTIC) est une équipe-projet de la Faculté des Lettres à Sorbonne Université, dédiée aux humanités numériques et rattachée administrativement au CELLF UMR 8599. Faisant suite à l'ancien LabEx OBVIL, l'ObTIC s'appuie sur une expertise acquise dans la production et l'édition des données, dans la conception et l'expérimentation des outils de fouille et d'analyse textuelle pour les chercheurs en lettres et sciences humaines et sociales, et dans le développement de nouvelles méthodologies de recherche, fondées sur le numérique, au sein de ces disciplines.

### Localisation :

Maison de la Recherche, 28 rue Serpente, 75006 Paris
SCAI, campus Jussieu, 4 place Jussieu 75005 Paris
Datalab, Bibliothèque François-Mitterrand, Quai François Mauriac, 75706 Paris

**Début du contrat** : 01/09/2024
**Durée** : 12 mois renouvelable
**Rémunération** : Selon expérience à partir de 2268 € bruts mensuels

**Missions et activités principales**

**Mission :**
Au sein de notre équipe interdisciplinaire, la personne recrutée sera chargée de développer des projets en traitement automatique du langage (TAL) et en apprentissage automatique. Elle jouera un rôle important dans l'analyse textuelle, la modélisation de données et l'implémentation de modèles de langage à grande échelle pour améliorer l'accessibilité, la compréhension et la valorisation de vastes collections de données textuelles. En collaboration étroite avec des chercheurs en sciences humaines, littérature et autres domaines, elle contribuera à l'avancement des connaissances et à la diffusion scientifique.

Cette mission comprendra les activités principales suivantes :

- Mettre en œuvre des modèles de TAL pour l'analyse et la catégorisation de données textuelles.
- Utiliser des techniques de machine learning et deep learning pour l'extraction d'informations, la compréhension de texte et la génération de contenu.
- Contribuer à des projets de data science en appliquant des méthodes statistiques, de visualisation de données, et d'apprentissage automatique pour interpréter et valoriser les données.
- Participer à la vie scientifique de l'équipe, en animant des ateliers et des séminaires sur les technologies TAL et l'analyse de données, et en contribuant au site web de l'équipe.
- Gérer les aspects techniques des projets, incluant l'installation d'outils d'analyse textuelle, le versionnage et la mise à disposition des données pour les utilisateurs.

**Compétences:**
Nous recherchons une personne de formation supérieure en traitement automatique des langues, science des données, informatique ou domaine connexe (Ingénieur ou Master), avec ces compétences :

- Expérience solide en TAL et en utilisation de bibliothèques et frameworks spécialisés (NLTK, spaCy, TensorFlow, PyTorch, Transformers…).
- Compétences en machine learning et deep learning, avec une capacité à appliquer ces techniques à des problématiques d'analyse textuelle.
- Maîtrise de la programmation en Python et familiarité avec les outils de gestion de versionnage, des bases de données et des outils et technologies du Web (API, JS…).
- Bonne connaissance des principes de l'analyse de données, y compris la statistique, la visualisation de données et l'interprétation des résultats.
- Capacité à travailler en équipe interdisciplinaire et à communiquer efficacement avec des non-spécialistes pour transformer des besoins de recherche en solutions techniques.
- Intérêt pour les sciences humaines et sociales et volonté de contribuer à des projets à l'intersection de la technologie et de la recherche académique.

**Encadrement : NON**
**Nb agents encadrés par catégorie : … A - … B - … C**

*Dans le cadre de vos fonctions, vous pourrez être amené à dispenser des formations internes en lien avec votre expertise métier.*

| Connaissances et compétences* |
|---|

**Connaissances transversales requises :**

- Organisation et fonctionnement de la recherche et de l'enseignement supérieur en France.
- Organisation et fonctionnement de Sorbonne Université.
- Réglementation applicable à son domaine d'activité professionnelle.
- Statuts, règlements et procédures de gestion des personnels des établissements publics d'enseignement supérieur et de recherche.

**Savoir-faire:**

- Capacité à hiérarchiser et à prioriser les tâches en tenant compte des contraintes et des échéances
- Capacité à proposer des solutions adaptées aux besoins
- Capacité à conduire et accompagner un projet
- Capacité à structurer et rédiger des notes techniques ou des documents de synthèse…
- Capacité à assurer de la bonne circulation de l'information au sein du service

**Savoir être :**

- Bon relationnel et sens du travail en équipe
- Disponibilité et réactivité
- Adaptabilité et autonomie
- Organisation et rigueur
- Confidentialité

**Conditions particulières d'exercice :** néant**.**

*Conformément à l'annexe de l'arrêté du 18 mars 2013 (NOR : MENH1305559A)

## Candidature :

Le dossier de candidature (CV et lettre de motivation) sera adressé au plus tard le 20/05/2024 à :

- Glenn Roe, responsable de l'équipe (glenn.roe@sorbonne-universite.fr)
- Motasem Alrahabi, coordinateur scientifique (motasem.alrahabi@sorbonne-universite.fr)