FICHE DE POSTE

Libellé de Fonctions : Ingénieur en Traitement Automatique du Langage

Catégorie : A

Corps : IGE/IGR

BAP (le cas échéant) : D

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

Missions et activités principales

Nom du projet: Apprentissage Profond pour l'Analyse Informatisée de la Subjectivité et des Emotions dans les troubles psychotiques émergents APAISE

Contexte du projet:
Ce projet interdisciplinaire s'inscrit dans la continuité des recherches en psychiatrie sur les troubles psychotiques émergents. Financé par la Fondation de France, il est le fruit d'une collaboration entre Sorbonne Université, l'INSERM et le CHRU de Brest. Son objectif est la détection précoce et l'intervention précoce pour les patients à risque de développer un trouble psychotique afin de permettre une meilleure prise en charge des patients et un ralentissement de leur évolution vers une chronicité. Il existe des travaux qui proposent des analyses linguistiques dans le cadre des troubles psychotiques émergents. Ils permettent de repérer des anomalies particulièrement discriminantes pour prédire l'évolution des patients sur différents niveaux: prosodie, syntaxe, sémantique, vocabulaire, discours, dialogue. Ces études omettent cependant la dimension subjective de l'expression langagière, étudiée à travers les émotions, les sentiments et les perceptions, qui, lorsqu'elles sont altérées, viennent traduire un rapport au monde et à soi perturbé.
A l'aide d'un corpus unique, composé d'entretiens psychiatriques transcrits, notre approche consiste à utiliser des algorithmes d'apprentissage profond afin d'analyser les caractéristiques subjectives du langage utilisé (émotions notamment) et d'augmenter la valeur prédictive du statut de « patient à risque ».

Mission:
La personne recrutée aura pour mission principale de traiter les entretiens des patients avec les outils nécessaires pour l'analyse textuelles et audio. Cette mission comprendra la création d'un corpus de référence, le prétraitement des données, l'analyse de ces données avec des modèles d'apprentissage automatique, et l'évaluation des résultats obtenus.
Elle travaillera avec une équipe interdisciplinaire: linguistes, psychiatres et informaticiens. Il convient dans un premier temps de définir les différentes modalités de l'expression subjective dans ce type de discours: tristesse, euphorie, hallucination, perte de contrôle, insensibilité aux émotions, etc. Il faudra par la suite expérimenter plusieurs méthodes en apprentissage automatique et apprentissage profond pour le repérage de ces informations, en les croisant avec des métadonnées cliniques ou des analyses prosodiques existantes. Enfin, il faudra proposer un protocole d'évaluation à différents niveaux, permettant de faire des comparaisons entre la cohorte des patients et des personnes contrôle, dans l'objectif de créer un modèle prédictif de l'évolution des troubles psychotiques.

Activités principales :
- Concevoir et piloter un dispositif de recherche .
- Proposer une démarche adaptée à l'objet d'étude.
- Concevoir et organiser des bases de données ou des corpus.
- Coordonner la conduite des opérations de production et de recueil de données.
- Organiser le traitement des données et mettre en œuvre des méthodes pour leur analyse.
- Contribuer à la réflexion sur les techniques et les méthodes employées dans le projet.
- Concevoir et réaliser une veille scientifique et méthodologique.

Profil recherché:
Nous recherchons une personne titulaire d'un master ou d'un diplôme d'ingénieur en traitement automatique des langues, science des données, informatique ou domaine connexe, avec expérience ou intérêt pour les neurosciences, la santé ou les sciences du vivant.

Compétences:
- Maîtrise de la programmation en Python et familiarité avec les outils de traitement automatique des langues, d'apprentissage machine et d'apprentissage profond (PyTorch, TensorFlow, Keras, Transformers, Sklearn, Pandas, Numpy, Gensim, spaCy, NLTK, etc.).
- Expérience en analyse de données audio et phonétique (prosodie, intonation, rythme, pauses, etc.), cruciale pour comprendre les nuances émotionnelles et subjectives du langage dans les troubles psychotiques émergents.
- Intérêt pour les neurosciences, la santé ou les sciences du vivant, et volonté de contribuer à des projets interdisciplinaires à l'intersection de la technologie et de la recherche académique.

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**.**

Encadrement : Non.

Rémunération: selon barème de rémunération des contractuels BIATSS de Sorbonne Université

Horaires : travail en présentiel, avec possibilité d'une journée de télétravail

Prise de fonction : septembre 2024, contrat de 9 mois (IR) et 10 mois (IE).

*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 26/05/2024 à :

- Motasem Alrahabi, (motasem.alrahabi@sorbonne-universite.fr)
- Julien Desclés (julien.descles@inserm.fr)
- Christophe Lemey (christophe.lemey@chu-brest.fr)