



Offre d'emploi

Assistant de recherche *Data Scientist* (6 mois)

CERDI, Université Clermont Auvergne-CNRS, France

Projet Big data & Environnement

Date limite de candidature : 20 juin 2021

Prise de poste : 1^{er} septembre 2021

Projet de recherche concerné

L'offre proposée s'inscrit dans la poursuite d'un travail de recherche sur le lien médias-pollution à partir de plusieurs grands ensembles de données : (1) la base GDELT (<https://www.gdeltproject.org/>) contenant des informations (émotions, etc.) sur les articles de presse publiés sur Internet ; (2) des données de surveillance en temps réels de la pollution de l'air en Chine.

Missions proposées

- **Traitement de données sur *Big Query* à partir de la base GDELT** : requêtes SQL afin d'extraire de nouvelles variables de la base GDELT et les fusionner avec les bases actuelles.
- **Création de nouveaux algorithmes d'analyse textuelle** : mettre en place de nouveaux algorithmes de *Machine Learning* et de *Text Mining* pour répondre à des besoins spécifiques d'analyse de documents textes associés à la base GDELT (par exemple : mieux identifier les articles sur le thème de la pollution, identifier la localisation des acteurs mentionnés dans les articles, cibler certains acteurs ou événements en particulier, etc.).
- **Missions secondaires** : contribution aux analyses économétriques et au travail de rédaction des articles ; participation possible à d'autres projets de recherche des membres du laboratoire.

Profil recherché

- Compétences en système de gestion de bases de données et Big Data (une première expérience avec BigQuery est un plus), en langage de programmation SQL et Python, et en algorithmes de *Machine Learning* et de *Text Mining/Natural Language Processing* (des expériences en *Deep Learning* appliqué à l'analyse de données textuelles est un plus)
- Maîtrise de logiciels de statistiques (Stata, R, ...)
- Maîtrise de l'anglais (lecture de documents et données en anglais)
- Masters 2 en statistiques appliquées, data science, économétrie



Processus de candidature

Les candidatures doivent être envoyées par email à Damien Cubizol (damien.cubizol@uca.fr) et Sébastien Marchand (sebastien.marchand@uca.fr) avec l'objet « Candidature Big data & Environnement ».

La candidature doit contenir :

- Un CV détaillé (copies des diplômes ou attestations de réussite)
- Une lettre de motivation (2 pages maximum) : description des motivations et des expériences passées (mettre en évidence la pertinence avec l'emploi proposé)
- Deux lettres de recommandation

Date limite de candidature : 20 juin 2021

Processus de sélection :

- Phase 1 : présélection des candidats
- Phase 2 : sélection finale (audition en visioconférence avant le 05 juillet)

Lieu de travail et rémunération

Un poste à temps plein est proposé pour une durée de 6 mois renouvelable 1 fois avec un salaire mensuel net d'environ 1 700 €.

Le candidat sélectionné travaillera avec une équipe d'enseignants-chercheurs du Centre d'Études et de Recherche sur le Développement International (<https://cerdi.uca.fr/>), une unité mixte de recherche Université Clermont Auvergne (UCA) et CNRS.

Le CERDI, basé à Clermont-Ferrand, offre un environnement de recherche passionnant, avec des groupes de travail et des séminaires. Clermont-Ferrand est une ville de province de taille moyenne qui offre un cadre de vie agréable à proximité du Parc Naturel des Volcans d'Auvergne (<https://www.auvergne-destination.com/>). L'Université de Clermont Auvergne rassemble 35 000 étudiants avec une forte ouverture internationale (<https://www.uca.fr/>).