



OFFRE D'EMPLOI

Administration handi-accueillante attachée à la mixité et à la diversité

LE HCERES RECRUTE POUR SON DEPARTEMENT OST UN(E) STAGIAIRE EN SCIENCE DES DONNEES

Stage à pourvoir entre le printemps et l'été 2022

L'OBSERVATOIRE DES SCIENCES ET TECHNIQUES

L'OST conçoit et produit des indicateurs et des études sur la recherche et l'innovation. L'ensemble constitué par les bases de données internes de l'OST sur les publications, brevets et projets européens est unique en France. Cette richesse de données lui permet notamment de développer des comparaisons institutionnelles et internationales solides. Le département travaille aussi sur des bases de données externes en fonction des besoins.

Les travaux de l'OST contribuent à différentes évaluations du Hcéres et plus largement à l'analyse des systèmes de recherche et d'innovation en France et à l'étranger. Selon les cas, ces analyses sont publiées sur le site du Hcéres, dans des rapports officiels ou dans des revues des domaines de la scientométrie et de l'évaluation des politiques publiques.

Base de données sur les publications et environnement technique

L'OST construit ses études bibliométriques à partir d'une version interne de la base *Web of Science* – OST-WoS. Le processus de mise à jour annuel des données de la base OST-WoS s'accompagne d'enrichissements des données sources sur les dimensions géographique, disciplinaire, institutionnelle et les supports de publication. Le processus de calcul des indicateurs est ensuite décomposé en 17 unités de traitement correspondant elles-mêmes à une série de procédures. Les indicateurs produits sont stockés dans des tables comportant 4,5 milliards d'enregistrements. En intégrant les tables intermédiaires (transformation et préparation des données), l'ensemble représente un espace occupé d'environ 3To.

Les procédures de calcul utilisées sont en PL-SQL, les programmes pour l'analyse sémantique sont en Python, la documentation est au format Markdown. La base de données est hébergée sur un serveur Oracle Enterprise 11g, les serveurs de développement sont sous Windows server Enterprise 2012. SQL Developer, R Studio et Visual Studio Code sont les principaux outils de développement utilisés.

ACTIVITÉS DU STAGE

Dans la suite d'un projet de refonte de la base OST-WoS en cours, l'OST modernise le calcul des indicateurs et se dote de nouveaux outils de visualisation et d'exploration des données. La mission principale du stage s'inscrit dans le cadre de la modernisation du calcul des indicateurs. Le/la stagiaire devra comprendre les étapes clés du calcul et les grands principes mis en œuvre dans la construction des indicateurs afin d'explorer certaines pistes d'évolutions envisagées.

L'OST mène des travaux en matière d'analyse sémantique des brevets. Il n'a pas encore conduit de travaux comparables sur les éléments textuels des notices des publications scientifiques (résumés). Le/la stagiaire sera chargé(e) de contribuer aux premiers travaux d'adaptation des méthodes développées sur les brevets au contexte des publications scientifiques.

CONNAISSANCES ET COMPÉTENCES REQUISES

- Connaissances avancées en SQL, Python et R
- Compétences en statistiques et en mathématiques
- Rigueur, méthodologie, capacité d'adaptation
- Capacité à communiquer, à rendre compte et à travailler en équipe
- Anglais technique



PROFIL

Master 1 ou Master 2 en science des données / modélisation statistique / informatique
2^e ou 3^e année d'école d'ingénieur en sciences des données / mathématiques appliquées / informatique

DURÉE ET RÉMUNÉRATION

Stage à inscrire entre le printemps et l'été 2022.
Gratification mensuelle égale au minimum légal.
Remboursement partiel des frais de transport.

CANDIDATURE

Adresser CV et lettre de motivation par courriel à :
david.sapinho@hceres.fr et ost.recrutements@hceres.fr
Tél : 01 55 55 71 02

LOCALISATION

2, rue Albert Einstein – 75013 PARIS
Métro ligne 14 et RER C : station Bibliothèque F. Mitterrand
Tramway T3a: arrêt Avenue de France
Bus 62 et 89 : arrêt : Porte de France (terminus)

Le Haut Conseil de l'Évaluation de la Recherche et de l'Enseignement Supérieur est en charge de l'évaluation des formations et des unités de recherche des établissements d'enseignement supérieur et des organismes de recherche. Le Hcéres assure aussi la production d'analyses et d'indicateurs bibliométriques à l'échelle nationale et internationale. Il porte une mission d'intégrité scientifique.
Site : [Accueil | Hcéres \(hceres.fr\)](https://www.hceres.fr)