

## OFFRE DE STAGE

### Améliorer la qualité des données grâce à la détection automatique d'anomalies

**Date de publication :** 01/12/2021

**Lieu :** Verneuil-en-Halatte (60) - accessible en transports en commun, à 40 mn au Nord de Paris

**Type de contrat :** stage

**Contact :** [morgane.salomon@ineris.fr](mailto:morgane.salomon@ineris.fr) - Tél. : 03 44 55 65 93 pour plus d'informations

Le poste proposé s'effectuera au sein de l'unité Données, Applications Techniques et Acquisition (DATA) de la direction Stratégie, politique scientifique et Communication (SCI). Cette unité réunit 8 personnes (ingénieurs et techniciens) travaillant sur des bases de données environnementales (gestion et exploitation) et sur de l'instrumentation.

Avec le nombre croissant de données en base, il est nécessaire de contrôler les données de façon automatique et de proposer des corrections.

Au sein de cette unité, le stagiaire participera au développement d'un outil d'amélioration de la qualité des données grâce à la détection des anomalies. Il travaillera notamment sur la base de données des émissions polluantes (BDREP). Cette base compile depuis 2007 les déclarations annuelles d'émissions (air, eau, sol) des exploitants.

Durant ce stage, le stagiaire aura pour missions de :

- Prendre en main la base de données,
- Construire un référentiel d'anomalies,
- Réaliser une recherche bibliographique sur les différentes méthodes de machine learning,
- Construire un algorithme de machine learning permettant de détecter les anomalies (au moment de la collecte et dans les données en base)
- Construire un algorithme permettant de proposer des correctifs

## PROFIL

Bac +4 et plus - Ce stage s'adresse un(e) étudiant(e) en master ou en école d'ingénieur avec une bonne connaissance théorique et pratique en sciences des données, particulièrement en machine learning/IA.

### **Expérience / Compétence**

Vous disposez de bonnes connaissances en langage de programmation (R/Python) et en SQL. Autonome, responsable, aimant le travail en équipe. Personne engagée et enthousiaste, ayant le sens des initiatives. Capacité de synthèse, aptitudes rédactionnelles.

## DIVERS

Durée : entre 4 et 6 mois

**Ce poste est ouvert aux personnes en situation de handicap.**