

# HelleboreTechnologies

Bringing Machine Learning and Data Mining to OTC Markets

## Stage Ingénieur développement C# Extraction d'information à partir d'un PDF

### Entreprise

Fondée en 2013, Hellebore Technologies est une FinTech basée à Paris reconnue pour son expertise et ses solutions SaaS pour la gestion des crédits dérivés. Ses clients sont des sociétés financières basées à Londres et aux Etats-Unis.

Hellebore Technologies innove continûment, en particulier sur l'analyse des séries temporelles de grande dimension ou l'analyse automatique de l'information textuelle. Ses innovations tant en Big Data qu'en Machine Learning sont mises en application quotidiennement dans les solutions utilisées par ses clients.

En 2022, la société travaillera sur l'amélioration de son outil d'extraction automatique d'information textuelle : ScriptMiner.

### Mission

Le challenge du stage est de définir une méthode permettant d'extraire des informations de PDF. Le stagiaire commencera par étudier les bibliothèques permettant d'accéder aux informations contenues dans un PDF. Ensuite, il s'intéressera à l'extraction d'information de PDF ne contenant que des images via des bibliothèques OCR (comme <https://github.com/tesseract-ocr/tesseract>). Enfin, le stagiaire analysera les possibilités d'amélioration de l'extraction d'information offertes par l'utilisation de bibliothèques d'analyse d'image tel que OpenCV (<https://opencv.org/>) – détection des lignes et colonnes des tableaux, analyses des zones de couleurs, ...

La durée du stage recommandée est de 6 mois, la rémunération est compétitive avec possibilité d'embauche à l'issue.

### Profil recherché

- => Tu es en dernière année d'école d'ingénieur ou de master
- => Tu connais le langage C#/Java
- => Tu aimes concevoir des algorithmes
- => Tu es rigoureux(se), tu as le souci des tests et de la documentation
- => Tu veux pouvoir voir rapidement les résultats de ton travail
- => Tu es curieux(se), enthousiaste, tu veux travailler dans une équipe jeune, talentueuse et soudée au cœur de Paris
- => Tu aimes l'esprit Startup, tout en étant dans une jeune entreprise ayant déjà des clients et un sérieux chiffre d'affaire

Alors envoie ta candidature à : [stage@helleboretech.com](mailto:stage@helleboretech.com) !