

Ingénieur en Big Data avec plus de 6 ans d'expérience dans la conception, le développement, et la gestion de plateformes de données à grande échelle. Expert en technologies Big Data, en particulier sur la stack Hortonworks (HDFS, Spark, Hive, Kafka, HBase) avec une solide maîtrise de plusieurs langages de programmation (Python, Java, Scala). Passionné par la résolution de problématiques complexes et l'amélioration des processus métiers via l'implémentation de solutions techniques innovantes. Proactif, avec un fort esprit d'équipe, une bonne adaptabilité, et un penchant pour les méthodologies agiles.

Expériences professionnelles

URSSAF Caisse Nationale - Data Engineer de 2022 à Aujourd'hui :

Projet Horus : Mise en place d'une plateforme Big Data pour la lutte contre la fraude en récupérant les transactions des plateformes d'économie collaborative. Ingestion de données reçu par la DGFIP dans le Datalake, puis rapprochement au SI Ucn avec la methode levenstein.

Datamining LCTI : Utilisation d'algorithme sur des données du datalake afin de scorer les entreprises fraudeuse.

IA GEN : projet POC sur la mise en place d'une IA generative pour aide au developpement de code

- **Technologies** : Hortonworks (HDFS, Spark, Hive, Kafka, HBase, Oracle SQL, PostgreSQL)
- **Langages** : Pyspark
- **Méthodologie** : Agile avec JIRA
- **Résultats** : Amélioration de la détection des fraudes grâce à une fiabilité accrue des données collectées.

THALES - Développeur Big Data de 2021 à 2022 :

Projet Horus : Participation à l'implémentation d'une solution Big Data pour l'Urssaf Caisse Nationale, avant internalisation dans le projet Horus.

- **Technologies**: Spark, HDFS, Hive, Kafka
- **Langages** : Python, Java, Scala
- **Méthodologie** : Agile avec JIRA

THALES - Développeur Big Data de 2018 – 2021 :

Projet MDS (Military Digital Stack) : Développement d'une plateforme Big Data pour les forces armées françaises.

Projet Defense

Technologies: Hortonworks CDP / HDP, HDFS, stack ELK, Kafka, MongoDB, Oracle SQL, Openshift, Docker, Ansible

- **Langages** : Scala, Java, Python
- **Méthodologie** : Agile avec JIRA
- **Résultats** : Mise en place d'une architecture de données robuste permettant de gérer des volumes massifs de données critiques.

Groupe Renault - Data Engineer (Apprentissage) 2015 à 2018

Intégré à l'équipe Data Stream du service après-vente (SAV Renault) pour développer des solutions d'analyse de données. Ingestion à l'aide d'un ETL de donnée dans le datalake.

Transcription d'algorithme pyspark en scala .

Requetage Hive.

- Technologies : Hortonworks (Zeppelin, HIVE, BASH)
- Langage : Scala
- Méthodologie : Agile avec JIRA

Formations

Ingénieur en Informatique

ESIEA - Paris/Laval - 2015 à 2018

DUT Informatique

Université Paris 8 - 2013 à 2015

Intérêts

- Technologies émergentes
- Cinéma
- Sport (football, judo, natation)

Compétences Techniques

- Langages : Python, Java, Scala
- CI/CD : Git, Jenkins, SonarQube, Nexus
- Big Data Platforms : Spark, HDFS, Hive, HBase, Oozie, Sqoop, NIFI
- Bases de Données : SQL, PostgreSQL, Lucene
- Virtualisation : Docker, VMware, Openshift (Kubernetes), Venv
- Outils de Monitoring : Ambari, Ranger, Grafana, ELK
- Notebooks : Zeppelin, Jupyter
- Méthodologie : Agile (Scrum, JIRA)