

# Consultant **Data Engineer/OPS**

REF ZKH

Tél : (+33)0756943180

## Compétences Techniques

Langages	Python, Scala, SQL, R
Analytique	Analyse de logs, Algorithmes supervisés (réseau de neurones), Python
BDD	MySQL, NoSQL, Big Data, Hadoop, Hive, Hbase, Dataiku, Oracle
Big Data	Spark, Kafka, Databricks, Kubernetes, Microsoft Azure
Cloud	Azure, appétence GCP & AWS
Anglais	Courant (TOEIC : 850)

## Compétences Fonctionnelles

Gouvernance de données  
Compétences analytiques  
Connaissance des systèmes de traitements Big Data  
Machine Learning

10/2023 à aujourd'hui Lead Data Engineer @ SNCF

### DSI SNCF :

- Traitement de la donnée (Spark, Scala, Python, SQL)
- Participer à la définition de la solution informatique applicative des applications confiées,
- Réaliser les conceptions techniques et détaillées,
- Concevoir, analyser, coder et tester les évolutions applicatives du parc en responsabilité conformément aux besoins et dans le respect des normes et procédures,
- Garantir de la cohérence et de la pérennité des développements,
- Participer à l'amélioration de la qualité et la performance,
- Intervenir dans la gestion des problèmes et dans la résolution d'incidents,
- Contribuer à l'élaboration des standards techniques,
- Œuvrer en contexte agile, utilisation de Jira et participation à la comitologie agile en place)
- Participer à la définition de la solution informatique applicative des applications confiées.

Stack **Spark, Scala, Python, Databricks, Aks, CI CD, YAML, Oracle, MySQL**

## 10/2022 – 10/2023 Data Engineer @ La Banque Postale

### DSI Groupe La Poste:

- Traitement de la donnée (Spark, Scala, Python, sur Databricks)
- Participer à la définition de la solution informatique applicative des applications confiées,
- Réaliser les conceptions techniques et détaillées,
- Concevoir, analyser, coder et tester les évolutions applicatives du parc en responsabilité conformément aux besoins et dans le respect des normes et procédures,
- Garantir de la cohérence et de la pérennité des développements,
- Participer à l'amélioration de la qualité et la performance,
- Intervenir dans la gestion des problèmes et dans la résolution d'incidents,
- Contribuer à l'élaboration des standards techniques,
- Œuvrer en contexte agile, utilisation de Jira et participation à la comitologie agile en place)
- Participer à la définition de la solution informatique applicative des applications confiées.

Stack **Spark, Scala, Python, Hadoop, Hive, Hbase, Dataiku, Oracle, MySQL**

## 8/2021 – 10/2022 Data Scientist / Engineer @ Stime

### DSI Groupement Les Mousquetaires :

- Traitement de la donnée (Spark, Scala, Python, sur Databricks)
- Modélisation de base de données
- Création de pipeline
- Calcul automatique distribué sous Spark (Scala, Pyspark)
- Orchestration de la donnée sur Microsoft Azure via datafactories
- Industrialisation de modèle de Machine Learning (IA, MLOps)
- Monitoring de data Lake
- Développement de pipelines avec Datafactories
- Procédures stockées en SQL
- Injection de données dans des bases de données NoSQL (Cosmos dB, redis, ...)

Stack **Spark, Scala, Python, Pyspark, Databricks, Azure Datafactory, SQL, Azure Cosmos DB**

## 3/2021 – 8/2021 Consultant Data @ Volvo

- Concevoir des solutions de données permettant de collecter, de nettoyer, d'organiser, de synthétiser et de modéliser efficacement les données ;
- Déployer et optimiser des flux de données complexes avec Azure Data Factory pour l'ingestion des données, Databricks pour le chargement, transformation et stockage des données en format parquet/delta tables et Spark/Python/Scala pour le développement des librairies de chargement et de transformation ;
- Mettre en place les pipelines de CI/CD permettant de valider la qualité de code et d'assurer des déploiements simplifiés ;
- Assurer le suivi de la qualité des données et de la bonne mise en œuvre des règles de gouvernance ;
- Travailler en étroite collaboration avec les Data Scientists et les lignes d'affaires ;

Stack **Python, Azure, Databricks, Spark, Scala, ...**

## 2/2020 – 2/2021 Data Scientist / Engineer @ Saur

- Création d'un algorithme NLP dans un projet IOT
- Déploiement de modèle de Machine Learning dans le cloud Microsoft Azure
- Création d'une plateforme IOT sous azure
- Création d'algorithme et de pipeline dans un environnement Big Data
- Déploiement d'une plateforme sous Azure Kubernetes services

Stack **Azure, Python, R, Spark, IOT, Kubernetes, Centreon**

## 10/2019 – 12/2020 Consultant Machine Learning @ Paul Wurth

- Reconnaissance d'outils dans un contexte de production
- Classification de photos en fonction des situations et du temps
- Réalisation d'une plateforme web permettant d'uploader une image et de le reconnaître via la vision par ordinateur
- Base de données avec information sur l'état de traitement des images

Stack **Python Flask, SQL, HTML/CSS**

## 10/2018 – 10/2019 Consultant Machine Learning @ Umanis

- Identifier les besoins et la problématique des directions métiers
- Définir une modélisation statistique qui permette de répondre à la problématique
- Construire des outils d'analyse pour collecter les données de l'entreprise.
- Sourcer et rassembler l'ensemble des sources de données structurées ou non structurées nécessaires à l'analyse et pertinences
- Organiser, étudier et synthétiser ces sources de données sous forme de résultats exploitables
- Modéliser les comportements et en extraire de nouveaux usages utilisateurs

### **Reconnaissance d'image et Application web**

- Classification de photos d'écriture cursive de chiffres (keras)  
Réalisation d'une plateforme web permettant d'uploader une image et de le reconnaître via la vision par ordinateur (Python Flask, HTML/CSS)
- Base de données avec information sur l'état de traitement des images (SQL)

### **Collecte de données et prédiction de prix immobiliers**

- Collecte des données (scrapping)
- Nettoyage, analyse et visualisation des données (pandas, numpy, matplotlib, seaborn)
- Features engineering et entraînement d'un modèle d'apprentissage supervisé de régression (sklearn)

Stack **Python, R, SQL, GIT, MongoDB, Elastic search, Hadoop, Spark, Docker**

2007 – 2018 Expert @ Supply Chain

**3 ans, Directeur d'exploitation @ StageCoach**

**5 ans, Directeur d'exploitation @ Confort Transport Service**

**4 ans, Responsable Grands Comptes @ Hilti Group**

## **Formations continues**

2019 Certification Développeur Data IA @ Ecole IA Microsoft

2019 Certification Dataiku DSS

2019 Certification Microsoft DP-100

## **Formations diplômantes**

2019 Master 2 Big data @ Centrale Supélec

2017 Master 2 Transport et Logistique @ Ecole supérieure des Transports

2007 Master 2 Ingénierie d'Affaires @ Sup de Vente