

Ibrahim

DataOps
Spark, Airflow, Trino
7 ans d'expérience
Disponible immédiatement

FORMATIONS & CERTIFICATIONS

2018 Master en Génie Informatique et Interaction Humain Machine - Université de Lorraine

COMPÉTENCES

Techniques :

- Langages : Java, Scala, **Python**, TypeScript, R, SQL.
- Big Data : **Spark**, Kafka, Hadoop, **Trino(Starburst)**, **Kubernetes**, **Openshift**, ElasticSearch, HDF.
- Plateformes & Frameworks : JavaEE, Spring Boot, NestJS, **Django**, **NodeJS**, **Angular**.
- DevOps : **Azure DevOps**, Jira, **GitLab CI/CD**, **Jenkins**, **Ansible**, GitHub Actions.
- **GitOps** : Argo CD, Argo Workflow, Airflow.
- Bases de Données : MySQL, MinIO, SQL Server, **MongoDB**, Firebase, Oracle.
- Monitoring : **Dynatrace**, **Kibana**, Grafana.
- Systèmes d'Exploitation : Windows, Linux, macOS.
- Outils & Méthodologies : Maven, UML, Design Patterns, Scrum, Git, **Docker**, **SonarQube**, Jupyter.

Fonctionnelles :

- Cadrage du besoin.
- Analyse de l'existant et étude des impacts.

LANGUES

Anglais professionnel

EXPERIENCES PROFESSIONNELLES

Octobre 2023 à octobre 2025	Crédit Agricole CIB DataOps
-----------------------------	--------------------------------

Outil de migration de larges volumes Iceberg (S3) vers Oracle via Trino (connecteurs : Iceberg + Oracle).

Taches et responsabilités :

- **Optimisation des performances de migration (paramétrage Trino : parallélisme, split size, query memory, etc.).**
- **Orchestration des Workflows via Airflow.**
- Conception et développement de packages PL/SQL pour la gestion des tables Oracle.
- Génération dynamique des requêtes INSERT / DELETE et pilotage des procédures PL/SQL.
- **Mise en place de la chaîne CI/CD (build, tests, Sonar ...).**
- Rédaction de la documentation et transfert de connaissances.

Trino / Starburst Query :

- ✓ Mise en place et optimisation de flux de migration de données depuis Iceberg (S3) vers Oracle via Trino, avec **règlages de performance (parallélisme, split size, query memory...)**.
- ✓ **Packaging et déploiement via Helm et un pipeline GitOps avec ArgoCD.**

Airflow :

- ✓ Orchestration des workflows de migration et automatisation des traitements quotidiens.

Spark :

- ✓ **Développement et tuning de jobs Spark sur Kubernetes** pour le calcul et l'agrégation de gros volumes de données financières (projet Bâle IV).

CI/CD & Kubernetes :

- ✓ Déploiement GitOps (ArgoCD), intégration continue (GitLab CI), monitoring et support en production.

Environnement Technique : Java 21, **Airflow**, SQL, PL/SQL, Spring Boot, **Trino**(Starburst), Apache Iceberg, Minio S3, **Oracle**, Git, Gitlab CI/CD, Kubernetes.

Le projet Bâle IV répond aux obligations réglementaires imposées par la Banque Centrale Européenne concernant le calcul des risques de marché.

Ce système critique agrège, transforme et restitue de très grands volumes de données financières issues de multiples sources (marchés, référentiels, valorisations, PnL, etc.) grâce à des traitements distribués à haute performance.

Taches et responsabilités :

- **Implémentation et optimisation des jobs Spark sur Kubernetes (tuning, partitionnement, configuration, etc.).**
- Création et optimisation des workflows avec Argo Workflows.
- Support N2/N3 : analyse des incidents en production via les outils de monitoring (Dynatrace, Kibana, Grafana).
- Déploiement sur kubernetes en mode GitOps via ArgoCD.
- Creation des pipelines CI/CD sur Gitlab.
- Application de la stratégie de branching (GitFlow), gestion des versions (tags) et conduite des revues de code.
- Utilisation de Jupyter Notebook pour investigation, débogage, exploration des jeux de données et réalisation de POC techniques.

- Rédaction et mise à jour de la documentation technique et des guides d'exploitation.
- Implémentation des fonctionnalités et validation des résultats en collaboration avec les équipes métiers.

Environnement Technique : Kubernetes (K9s), Argo Workflows, ArgoCD, Oracle, Spark (Java & PySpark), Spring Boot, Dynatrace, Kibana, Grafana, Jupyter, GitLab, GitLab CI/CD, Jira, Agile (Scrum).

Mai 2020 - Oct 2023

EDF

Développeur Full-Stack / Big Data

Participation à la conception et la réalisation de Emoby, une solution au service des ressources humaines, destinée au sourcing et à la réalisation de cartographie des compétences.

Taches et responsabilités :

- Participation à la conception des nouvelles fonctionnalités.
- Implémentation d'un pipeline d'acquisition des CVs et des rapports d'activités.
- **Implémentation des batches spark de génération des cartographies de compétences et de métiers.**
- Participer au développement d'une application web de restitution (recherche des candidats pour un besoin spécifique et la génération des cartographies).

Environnement Technique : Docker, Ansible, Gitlab-CI, Spark, MinIO, Java, Angular, Python, Spring-Boot, Kafka, Postgres.

Avr 2018 - Avr 2020

Assystem

Développeur Full-Stack

Conception et le développement d'une solution Web conçue pour détecter automatiquement des données relatives à des équipements dans des documents.

Taches et responsabilités :

- Conception de l'architecture globale et sélection des technologies (Java, Angular, RabbitMQ, Docker, etc.).
- Développement et intégration des services backend et frontend, y compris la mise en place de bases de données et l'interface utilisateur.
- Dockerisation de l'ensemble des services pour faciliter le déploiement. • Réalisation de tests complets (unitaires, d'intégration).
- Configuration des environnements de développement.

Environnement Technique : Java / JHipster, TypeScript, Angular, Azure DevOps, Docker, Git, MinIO, Spring-Boot, Postgres, RabbitMQ.