



Djalil Salah-bey

Data Engineer | Analytics Engineer | Recherche CDI

Ingénieur Data orienté plateforme et architecture, je conçois et déploie des environnements analytiques modernes reposant sur Databricks, Airflow et ClickHouse.

Je mets un point d'honneur à industrialiser les pipelines, standardiser les modèles métiers et garantir la gouvernance dans une logique DataOps complète.

Mon ambition : faire de la donnée un actif maîtrisé, performant et durable.

Expériences professionnelles

Data Engineer -Freelance

De mars 2025 à septembre 2025 **PowerUp Technology** Paris

- Conception et industrialisation d'une plateforme analytique unifiée sur Azure Databricks et ClickHouse, dédiée à la surveillance des performances de batteries industrielles connectées.
- Développement d'un module Python de détection d'anomalies et de validation de conformité pour renforcer la fiabilité des flux d'ingestion.
- Mise en œuvre d'une modélisation OLAP hiérarchique garantissant granularité, navigation multi-niveaux et chargements incrémentaux optimisés.
- Élaboration et déploiement de datamarts métier (sécurité, performance, usage) alimentant les tableaux de bord et heatmaps analytiques.
- Intégration complète en CI/CD (Airflow orchestré sous Docker, GitHub Actions, gestion sécurisée via Azure Key Vault et DevOps).
- Supervision et observabilité avancée : suivi des performances, journalisation applicative, alertes en temps réel sous Grafana / InfluxDB.
- Gouvernance : gestion centralisée des accès, traçabilité et chiffrement des données sensibles.

Stack : Azure Databricks, Airflow, PySpark, Clickhouse, PostgreSQL, influxDB, MongoDB, Github Actions, Grafana, dbt, DuckDB, Docker, Kubernetes.

Data Engineer -Alternance

De septembre 2022 à septembre 2024 **Koacher** Lyon

- Conception du Data Warehouse central consolidant l'ensemble des données sportives issues de la plateforme.
- Développement et automatisation de pipelines d'ingestion et de transformation sous Airflow et dbt Cloud, avec supervision continue et contrôles qualité.
- Participation à la formalisation des métriques métier (rétention, engagement, performance athlétique) et à la documentation analytique partagée.
- Intégration d'une chaîne CI/CD pour fiabiliser les déploiements et contrôle des dépendances inter-modèles.
- Optimisation des performances SQL et rationalisation des modèles analytiques pour garantir des temps de réponse homogènes.
- Monitoring proactif de l'environnement via Grafana et n8n.

Stack : Airflow, Databricks, DBT Cloud, Grafana, MongoDB, Docker, Kubernetes, n8n, Grafana, Big Query, Prometheus, Github Actions.

Developpeur ETL & Big Data -Alternance

De 2021 à 2022 **Mutli By EasyPlanet** Paris

- Développement et maintenance de flux ETL en Talend, Python et SQL pour l'intégration de données RH, logistiques et financières.
- Conception d'un data warehouse centralisé dédié au reporting BI et au suivi des indicateurs de performance.
- Automatisation des traitements et planification des tâches récurrentes, assurant une disponibilité continue des données.
- Mise en place de mécanismes de contrôle qualité et de journalisation pour garantir la traçabilité des flux.
- Intégration de modules d'intelligence artificielle (OCR, reconnaissance faciale) au sein des chaînes de traitement.

Stack : Talend, Python, SQL Server, PostgreSQL, Power BI, Google Cloud Platform, Big Query, Airflow.

Diplômes et Formations

IA School — Master Ingénierie et Science des Données

De 2023 à 2025 **IA SCHOOL** Paris

IA School — Bachelor 3 Ingénierie et Science des Données

De 2022 à 2023 **IA SCHOOL** Paris

IAE Lille — Master Systèmes d'Information et Management de la Donnée

De 2021 à 2022 **IAE Lille** Lille

Université de Lille — Licence Banque Finance

De 2018 à 2021 **Université de Lille 2** Lille

✉ salahbeydjailil@gmail.com

🔗 <https://djo-project-portfolio.vercel.app/>

📍 France

☎ +33 6 11 27 91 53

Compétences

Plateforme / Cloud :

Azure (ADLS, DevOps, Key Vault), Databricks, ClickHouse, GCP, BigQuery

Orchestration / CI-CD :

Airflow, Docker, GitHub Actions

Transformation /

Modélisation : PySpark, dbt, SQL, OLAP, Kimball

Base de données :

PostgreSQL, MongoDB, InfluxDB

Monitoring /

Visualisation : Grafana, Prometheus, Power BI, Tableau

Langages : Python, SQL, PySpark

Langues

Français

Anglais

Arabe

Espagnol

Centres d'intérêt

Sport

Tennis

Basket Ball

Football

Voyages

Nouvelles cultures

Tour du monde

Tech

Conférences Data

Veille IA

Meetups