



Djalil Salah-bey

## Data Engineer | Analytics Engineer

- Data Engineer autonome et passionné par la construction de plateformes data modernes
- J'assure la collecte, le traitement et la fiabilité des données au sein d'environnements cloud.
- Mon travail vise à rendre la donnée accessible, gouvernée et exploitable pour la prise de décision.
- A la recherche d'une nouvelle opportunité , -Portfolio Projects : <https://djo-project-portfolio.vercel.app/>

### Expériences professionnelles

#### ● Data Engineer Databricks (Freelance 8 mois)

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

**Projet de déploiement d'une Data Platform Lakehouse pour la supervision et la prédiction des performances de batteries industrielles connectées (BESS – Battery Energy Storage Systems).**

- Conception et maintenance de **pipelines Databricks** pour la supervision et l'analyse des performances des **batteries**.
- Mise en place d'une architecture **Delta Lake (Bronze → Silver → Gold)** robuste et scalable pour centraliser les **données de télémétrie**.
- Développement de traitements **batch et streaming** connectés à **Kafka** et **Azure Blob Storage** pour l'ingestion des flux IoT.
- Exploitation de **Spark SQL, MLflow** et des **notebooks collaboratifs Databricks** pour la transformation et l'analyse.
- Intégration et normalisation de données **multi-sources** : capteurs IoT, API, bases relationnelles.
- **Exploitation des fonctionnalités avancées de Delta Lake** : *time travel, merge, versioning, vacuum*.
- Collaboration avec les Data Scientists pour **industrialiser les modèles ML** (prédiction de performance et vieillissement) et gérer les **feature tables**.
- Automatisation et supervision des workflows via **Airflow, Docker, Azure DevOps**.
- **Mise en place d'un monitoring applicatif** pour le suivi des jobs et des coûts Databricks.
- Application des pratiques **DataOps & CI/CD** : versioning, tests, déploiements automatisés.
- Mise en œuvre du **contrôle qualité, de la sécurité et de la traçabilité** (dbt, Unity Catalog, Key Vault).
- Optimisation de la **scalabilité et des performances Spark** sur clusters Databricks.
- **Gestion des accès, lineage et gouvernance centralisée** via Unity Catalog et Key Vault.
- Participation à la définition des **stratégies de scaling et de gouvernance** de la plateforme Lakehouse.

- **Résultats** : flux unifiés, jobs industrialisés, supervision temps réel, – 40 % de latence de traitement.
- **Stack** : **Azure, Databricks, Delta Lake, Spark/PySpark, Airflow, dbt, Kafka, ADLS, MLflow, Unity Catalog, Azure DevOps, Docker, Grafana, Prometheus, MongoDB, Clickhouse.**

#### ● Data Engineer DBT (Alternance 2 ans)

De décembre 2022 à décembre 2024 **Koacher** Lyon

**Projet de développement analytique au sein d'une plateforme web & mobile dédiée au suivi des performances sportives et à la mise en relation coaches / athlètes.**

- Conception et déploiement du **Data Warehouse analytique central** de la plateforme Koacher.
- Modélisation des données **staging → intermediate → marts sous dbt Cloud**.
- Développement de modèles **SQL testés, documentés et versionnés**.
- Intégration et orchestration des pipelines via **Airflow**.
- Mise en place d'une **chaîne CI/CD** avec **GitHub Actions** pour tests et déploiements automatisés.
- Suivi de la **fraîcheur des sources** et du **lineage automatique** via dbt Cloud.
- Utilisation de **macros Jinja** et de **packages dbt-utils / audit-helper** pour standardiser les modèles.
- Création d'un **catalogue de métriques business** : rétention, engagement, performance sportive.
- Gestion de la **qualité des données** : tests dbt, validation, fraîcheur et cohérence des sources.
- Intégration des **tests dbt** dans la chaîne CI/CD pour validation avant déploiement.
- Supervision des workflows et alertes via **Grafana** et **Prometheus**.
- Contribution à la **gouvernance analytique** et à la **documentation automatisée** (linéage, docs)
- Collaboration avec les équipes produit et data science pour le **reporting et les datasets ML**.
- **Optimisation des modèles SQL et suivi** des performances analytiques en temps réel.

- **Résultats** : standardisation des modèles analytiques, réduction de 60 % du temps de requête et adoption totale du framework dbt en production.
- **Stack** : **dbt Cloud, Jinja, Airflow, PostgreSQL, BigQuery, Python, Docker, Grafana, Prometheus, GitHub Actions, Great Expectations, Marquez, OpenLineage, Power BI.**

#### ● Data Engineer Talend (Stage 6 mois)

De janvier 2022 à juillet 2022 **Mutli By EasyPlanet** Paris

### Diplômes

#### ● IA School — Master Expert Data & Ingénierie de l'Intelligence Artificielle

De 2022 à avril 2025 **IA SCHOOL** Paris

#### ● IAE Lille — Master Systèmes d'Information & Management de la Donnée

De 2021 à 2022 **IAE Lille** Lille

#### ● Université de Lille — Licence Spécialisé en Banque et Finance

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

✉ salahbeydjalil@gmail.com

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

📍 France

☎ +33 6 11 27 91 53

### Compétences Techniques

#### Data Platform & Processing

Databricks · Spark · Delta Lake · ClickHouse · Kafka · Airflow · dbt Cloud · Talend

#### Data Storage & Databases

Azure Blob Storage · PostgreSQL · SQL Server · BigQuery · InfluxDB · MongoDB

#### DataOps & CI/CD

Git · GitHub Actions · Azure DevOps · Docker · Kubernetes

#### Quality & Governance

Unity Catalog · Key Vault · Great Expectations

#### Monitoring & Observability

Grafana · Prometheus

#### Visualization & Analytics

Power BI · Streamlit · Metabase

#### Languages

Python · SQL · YAML · Bash

#### Cloud & Environment

Azure · GCP · Linux (WSL2)

### Langues

Français

Anglais

Arabe

Espagnol

### Centres d'intérêt

#### Sport

Tennis  
Basket Ball  
Football

#### Tech

Conférences Data  
Veille IA