



✉ salahbeydjailil@gmail.com  
🔗 <https://djo-project-portfolio.vercel.app/>  
📍 France  
☎ +33 6 11 27 91 53

## 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

## Compétences

- **Data Platforms :** Databricks, Delta Lake, dbt, Unity Catalog
- **Data Engineering :** Spark (PySpark, SparkSQL), ETL/ELT, batch/streaming, Kafka
- **Databases :** PostgreSQL, MySQL, MongoDB, InfluxDB, ClickHouse
- **Cloud & DevOps :** Azure, GCP, Docker, Kubernetes (AKS/GKE), Terraform
- **Orchestration :** Airflow, Databricks Workflows, CI/CD (GitHub Actions, Azure DevOps)
- **Observabilité :** Prometheus, Grafana
- **BI & Apps :** Power BI, Streamlit, APIs Flask
- **ML/MLOps :** MLflow, scikit-learn, TensorFlow, PyTorch

Djalil Salah-bey

## Data Engineer | Analyst Engineer

Je suis Data Engineer avec 3 ans d'expérience, spécialisé dans les data platforms cloud, en particulier Databricks Lakehouse et les solutions cloud natives.

Au-delà de la technique, j'aime concevoir des pipelines end-to-end fiables et mesurables, qui transforment la donnée brute en indicateurs clairs, optimisent les coûts et fournissent aux équipes des produits data prêts à l'usage (dashboards, datasets ML, APIs).

## Expériences professionnelles

### Data Engineer

De mars 2023 à octobre 2025 **PowerUp Technology -EnergyTech** Paris

- **Conception et déploiement d'une Data Platform Lakehouse sur Azure (Databricks + Delta Lake) pour la gestion des données batteries IoT.**

- Développement et maintenance de **pipelines d'ingestion** (Azure Data Factory, Azure Blob, Kafka).
- Conception de pipelines de **transformation** sur **Databricks** (dbt, SparkSQL, PySpark) selon l'architecture Médaille (Bronze/Silver/Gold).
- Mise en place de **tests unitaires et règles qualité** (dbt tests, PySpark, Great Expectations).
- Orchestration des workflows avec **Databricks Workflows & Airflow** (SLA, retries, backfills).
- Participation à la conception de pipelines **CI/CD** (Azure DevOps, Git, Docker, Kubernetes AKS).
- **Monitoring & observabilité** avec **Grafana & Datadog** (alertes temps réel, métriques critiques).
- Gestion des accès et gouvernance via **Unity Catalog** (RBAC, conformité RGPD).
- Optimisation des performances : **partitionnement, Z-Ordering, caching** (requêtes <500 ms sur 40M+ points).
- Préparation de datasets pour le **Machine Learning** et suivi du cycle de vie avec **MLflow**.
- Collaboration avec les **Product Owners** pour livrer des produits data (dashboards Power BI/Streamlit, marts analytiques, datasets ML).

**Stack technique :** Azure (ADF, Databricks, Blob, AKS, DevOps), Delta Lake, Spark (PySpark, SparkSQL), dbt, Unity Catalog, MLflow, Airflow, Kafka, Docker, Kubernetes, Grafana, Datadog, Git, CI/CD.

### Data Engineer

De décembre 2022 à novembre 2023 **Koacher -SportTech** Lyon

- **Développement et maintenance d'une plateforme SaaS de coaching sportif (web et mobile).**

- Conception et optimisation des bases de données relationnelles (**PostgreSQL, MySQL**).
- Développement d'**APIs REST sécurisées** en Python/Flask pour l'intégration front-end/back-end.
- Mise en place de pipelines **ETL (Python, dbt)** pour centraliser et transformer les données utilisateurs (logs d'entraînement, sessions vidéo).
- Orchestration des workflows avec **Airflow** pour l'automatisation et la reprise sur incident.
- Intégration de solutions analytiques (**Power BI, dashboards internes**) pour le suivi des performances utilisateurs.
- Mise en place de **tests unitaires & d'intégration** (pytest, dbt tests) pour fiabiliser les pipelines.
- Conteneurisation avec **Docker** et déploiement sur **Kubernetes (GKE)** pour la scalabilité.
- Implémentation de l'authentification (OAuth2) et chiffrement pour assurer la **sécurité et conformité RGPD**.
- Optimisation des performances applicatives (profiling APIs, SQL tuning, caching).
- Suivi de production et **monitoring applicatif** avec **Prometheus & Grafana**.

**Stack technique :** GCP, Kubernetes (GKE), Python, Flask, PostgreSQL, MySQL, dbt, Airflow, Power BI, Docker, Git, CI/CD (GitHub Actions), Prometheus, Grafana.

## Diplômes et Formations

### BAC + 5 Ingénierie & Science des données

De 2023 à 2025 **IA School** Aix-en-Provence

### BAC +3 Ingénierie & Science des données

De 2022 à 2023 **IA SCHOOL** Paris

### BAC + 3 Licence Banque et finance

De 2019 à 2022 **Université de Lille 2** Lille