



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 (Delta Lake, dbt, Spark) 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).

✉ salahbeydjalil@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

Expériences professionnelles

Data Engineer

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

- Conception et déploiement d'une Data Platform Databricks Lakehouse pour la gestion des données batteries IoT (charge/décharge, télémétrie) en collaboration active avec 3 data engineers.
- Implémentation de l'architecture Médaille (Bronze / Silver / Gold) avec Delta Lake et automatisation des pipelines via dbt, gouvernance avec Unity Catalog.
- Développement de pipelines batch et streaming (Databricks Autoloader, Azure Blob Storage, Kafka) et transformations scalables en SparkSQL / PySpark.
- Optimisation des performances (partitionnement, ZORDER, caching) pour supporter l'analyse de milliards de points de mesure.
- Création de data marts analytiques (familles KPIs d'usage, de performance & de sécurité) exposés via Databricks SQL et intégrés à Power BI dashboards (cartes thermiques).
- Préparation de datasets pour le Machine Learning et suivi du cycle de vie avec MLflow.
- Orchestration des pipelines via Databricks Workflows et intégration CI/CD (Git, Azure DevOps, Airflow).
- Gestion des environnements de développement via Docker et déploiement via Azure Kubernetes Services (AKS).
- Intégration de bases SQL et NoSQL (PostgreSQL, MongoDB, InfluxDB) dans les pipelines de données.
- Observabilité & qualité : métriques/logs Grafana, alertings Airflow, tests dbt.

Stack technique : Azure, Databricks, Delta Lake, Spark (PySpark, SparkSQL), dbt, Unity Catalog, Grafana, MLflow, Docker, Kubernetes, PostgreSQL, MongoDB, InfluxDB, Azure Blob, Kafka, Git, CI/CD, Power BI.

Data Engineer

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

- Développement et maintenance d'une plateforme SaaS de coaching sportif (web et mobile) en collaboration avec une équipe de 4 dans le pôle informatique.
- Conception et implémentation des bases de données relationnelles (PostgreSQL, MySQL)
- Optimisation des requêtes SQL (schémas normalisés, indexation, analyse de plans d'exécution).
- Création d'APIs REST sécurisées en Python / Flask pour l'intégration entre front-end et back-end (authentification JWT, validation des schémas, rate limiting).
- Mise en place de pipelines ETL / ELT pour centraliser et transformer les données utilisateurs (logs d'entraînement, sessions vidéo) avec dbt (modèles, tests, documentation).
- Intégration de solutions analytiques pour suivre la performance des utilisateurs (dashboards internes Power BI : usage, rétention, engagement).
- Gestion du cycle de développement avec Git, Docker, CI/CD (GitHub Actions).
- Observabilité et supervision des services avec Prometheus (latence API, taux d'erreurs, des jobs ETL).

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

Diplômes et Formations

BAC + 5 Ingénierie & Science des données

De 2022 à 2025 **IA School** Aix-en-Provence

BAC + 3 Licence Banque et finance

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