



Djallil Salah-bey

Data Engineer | Analytics Engineer

Data Engineer passionné par la conception de plateformes data fiables et évolutives.

J'interviens sur l'ensemble du cycle de vie des données avec une attention particulière portée à la qualité, la performance et la traçabilité.

Habitué aux environnements cloud et à la production, j'aime construire des pipelines robustes, documentés et simples à maintenir, au service des équipes métier et data science.

✉ salahbeydjallil@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

Technologies Utilisées

- **Langages & Data** : Python, SQL, Jinja, YAML
- **Modélisation & Transformation** : dbt, Delta Lake, Kimball, SCD, ELT
- **Orchestration & Automation** : Airflow, Databricks Jobs, CI/CD (GitHub Actions, Azure DevOps)
- **Cloud & Stockage** : Azure (Blob, Data Lake), GCP (GCS, BigQuery), Docker, Kubernetes
- **Base de données & Moteurs** : ClickHouse, PostgreSQL, MySQL, DuckDB
- **Monitoring & Lineage** : Grafana, Prometheus, Marquez, Unity Catalog
- **BI & Analytics** : Power BI, Streamlit, Superset

Expériences professionnelles

Data Engineer Intern

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

- J'ai **structuré la plateforme de données** autour d'une architecture Lakehouse moderne, assurant la cohérence, la fiabilité et la traçabilité complète des informations.
- J'ai **stabilisé les flux de production** en automatisant les traitements quotidiens et en garantissant des mises à jour fiables et continues.
- J'ai **construit des jeux de données analytiques prêts à l'usage**, qui ont permis de fluidifier le travail des équipes Data Science et Business.
- J'ai **renforcé la qualité et la gouvernance des données** grâce à des contrôles systématiques et une documentation claire, facilement partageable entre équipes.
- J'ai **optimisé les performances et les coûts de calcul**, tout en améliorant la réactivité des analyses et la rapidité d'accès aux données.
- J'ai **mis en place un processus de déploiement maîtrisé**, avec validation, suivi et reprise rapide en cas d'incident.
- J'ai **développé une culture d'observabilité et d'anticipation**, permettant d'identifier rapidement les dérives et d'assurer la fiabilité du service en continu

Stack : Python, PySpark, Azure Databricks, Azure DevOps, Apache Spark, DBT, Unity Catalog, Grafana, MLflow, Docker, Kubernetes, PostgreSQL, MongoDB, InfluxDB, Kafka, Gitlens, CI/CD, Clickhouse.

Data Engineer en alternance

De décembre 2022 à novembre 2024 **Koacher -SportTech** Lyon

- J'ai **développé et fait évoluer une plateforme SaaS** de coaching sportif (web & mobile) utilisée par des centaines d'athlètes et coachs.
- J'ai **structuré les données** dès la base : conception des modèles relationnels, normalisation et optimisation des requêtes.
- J'ai **construit des flux d'ingestion et de transformation** pour centraliser les données d'entraînement, d'activité et de performance.
- J'ai **intégré des APIs métiers** permettant la communication fluide entre les services et la validation des échanges.
- J'ai **mis en place des indicateurs analytiques** pour suivre la performance des utilisateurs et la rétention des sessions.
- J'ai **instauré des processus de développement fiables** : versionnement, automatisation des tests et intégration continue.
- J'ai **assuré la supervision et la stabilité des services**, avec alertes et tableaux de suivi pour anticiper les anomalies.

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

Compétences Professionnelles

- **Architecture & Modélisation** : conception de plateformes data robustes et évolutives, structuration en couches analytiques (modèles logiques, historisation, schémas métiers).
- **Pipelines & Production** : développement et automatisation de flux d'ingestion, de transformation et de mise à disposition des données, avec supervision continue et reprise automatique.
- **Qualité & Gouvernance** : validation systématique des données, documentation standardisée, traçabilité et conformité de bout en bout.
- **Performance & Optimisation** : amélioration des temps de calcul, tuning des traitements et pilotage des coûts d'exécution en production.
- **Orchestration & Automatisation** : planification, suivi des exécutions, intégration et déploiement continus pour garantir la stabilité des environnements.
- **Observabilité & Supervision** : mise en place d'indicateurs, tableaux de bord et alertes pour assurer la fiabilité opérationnelle et anticiper les dérives.
- **Culture d'équipe & Méthodologie** : communication claire, documentation vivante et diffusion des bonnes pratiques data au sein des projets.

Diplômes et Formations

Master Ingénierie & Science des données

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

Licence Ingénierie & Science des données

De 2022 à 2023 **IA SCHOOL** Paris