



Djalil Salah-bey

Data Engineer | Expert ETL/ELT (AirFlow, dbt, BigQuery, Databricks)

Ingénieur Data spécialisé dans la modélisation analytique, l'orchestration et la mise en production de pipelines de données en intégrant les bonnes pratiques de DataOps, CI/CD et observabilité.

Email
salahbeydjalil@gmail.com

Numéro de téléphone
+33 6 11 27 91 53

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

Langues

Français
Natif

Anglais
Courant

Soft Skills

Autonome
Curieux
Proactif
Pédagogue

Compétences Techniques

- Langages : Python, SQL, PySpark, Bash
- Intégration : Airflow, dbt Cloud, Talend
- DataOps : Azure, GCP, Terraform, Docker, GitHub Actions, Git
- Visualisation : Power BI, Looker Studio, Metabase
- Stockage Cloud : Databricks, GCS, ADLS, S3, BigQuery
- Bases de données & Moteurs analytiques : BigQuery, PostgreSQL, ClickHouse, DuckDB, MongoDB, InfluxDB

Diplômes et formations

- IA School — Diplôme RNCP Niveau 7 Expert Data & Ingénierie de l'Intelligence Artificielle (2023-2025)
- IA School — Bachelor 3 Expert Data & Ingénierie de l'Intelligence Artificielle (2022-2023)
- Université de Lille — Licence Banque, Finance & Analyse Quantitative (2019-2022)

Expériences professionnelles

- PowerUp Technology

Freelance Data/Analytics Engineer

Paris

De mars 2025 à octobre 2025

Refonte d'une plateforme BI vers **ClickHouse** pour des **heatmaps 2D/3D** thermiques performantes et une navigation hiérarchique fluide (*cellule* → *pack* → *site* → *client*).

 - Définition du grain d'analyse, des faits/dimensions et des hiérarchies spatiales et temporelles avec règles d'agrégation adaptées à ClickHouse.
 - Intégration des données depuis Azure Data Lake Storage, gestion des chargements incrémentaux, backfills contrôlés et idempotence des pipelines.
 - Conception d'un système de discrétisation spatiale pour agréger les températures et mesures dans des cartes 2D/3D homogènes, garantissant cohérence et précision sur de larges volumétries.
 - Génération automatisée de DAGs Airflow via templates Jinja, avec dépendances, SLA, alertes et exécutions multi-clients.
 - Création et optimisation de tables (partitionnement, compression, indexation) et de materialized views garantissant des latences <100 ms sur les requêtes analytiques.
 - Mise en place d'un modèle analytique dbt avec documentation automatisée, tests (unicité, fraîcheur, cohérence) et CI GitHub Actions pour validation continue.
 - Création d'un monorepo Databricks Repos pour standardiser le développement (Docker, conventions, packaging, linting) et uniformiser le déploiement.
 - Publication des heatmaps 2D/3D complètes avec toutes les fonctionnalités métiers demandées (zoom, filtres, agrégation dynamique, hiérarchies interactives).
 - Mise en place de data contracts, tests SQL et monitoring Grafana pour garantir la fiabilité opérationnelle.
 - Stack: ClickHouse, SQL, Python, dbt, Airflow, Databricks, Docker, ADLS, Grafana, Power BI, Git.

Résultat : Migration end-to-end livrée : modèle OLAP ClickHouse performant et gouverné, heatmaps 2D/3D interactives conformes aux exigences métier, pipelines automatisés et industrialisés (Airflow + CI).