



Djalil Salah-bey

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

Mon travail couvre tout le cycle de la donnée.
Je fais entrer les données régulièrement, je les range clairement et en sécurité, je les nettoie pour les rendre fiables, je les mets au format métier pour répondre aux questions clés, je l'automatise avec un planning et des alertes et je les partage pour toute l'équipe.
À chaque étape, il y a des tests et une chaîne CI/CD qui vérifie, déploie, puis contrôle que tout fonctionne avant de passer en production.

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

Passion

Sport
Programmation

Soft Skills

Autonome
Proactif
Pédagogue

Stack Technique

Cloud & Infrastructure
Azure Data Factory · Azure Databricks · Azure Synapse Analytics · Azure Data Lake Storage Gen2 · Azure DevOps · Azure Key Vault · Azure Purview · Entra ID · Terraform · Bicep · Docker · Kubernetes (AKS)

Data Engineering
Python (PySpark, Pandas) · SQL · Delta Lake · ClickHouse · dbt Core · Airflow · Data Modeling (Star / Snowflake / OLAP) · ETL / ELT · Great Expectations

Orchestration & CI/CD
Apache Airflow · Azure Data Factory · Azure DevOps Pipelines · Git · Bash · Linux

MLOps & Machine Learning
Azure Machine Learning · MLflow · AKS · ACI · Scikit-learn · Pandas · Numpy

Monitoring & Logging
Azure Monitor · Log Analytics · Prometheus · Grafana

Visualisation & BI
Power BI · Grafana · ClickHouse Dashboards

Expériences professionnelles

- PowerUp Technology

Freelance Data Engineer

Paris

De mars 2025 à octobre 2025

Projet : refonte d'une plateforme Data Lakehouse Azure & ClickHouse pour l'analyse thermique de systèmes IoT BESS

 - Conception d'une **infrastructure data unifiée** combinant Azure et ClickHouse pour l'analyse et le monitoring des données IoT industrielles.
 - Développement de **pipelines d'ingestion et d'orchestration** (Data Factory, Airflow AKS) pour capter, historiser et fiabiliser les flux de télémétrie en continu.
 - Structuration du **Data Lake ADLS Gen2** en zones *bronze / silver / gold* avec gestion incrémentale et historisation automatique.
 - Traitements distribués sous **Azure Databricks (PySpark / Delta Lake)** pour l'agrégation et la validation de mesures haute fréquence.
 - Conception d'un **modèle analytique OLAP** dans **ClickHouse**, avec hiérarchies spatiales et temporelles (cellule → pack → site → client).
 - Optimisation des performances via **vues matérialisées** et **agrégations 2D/3D** pour la visualisation des **heatmaps** dans Grafana
 - Mise en place du **monitoring et de la qualité** (Prometheus, Azure Monitor, Great Expectations) avec alerting et suivi SLA.
 - Industrialisation CI/CD** sous **Azure DevOps Pipelines** et adoption d'une démarche **DataOps** (Terraform, automatisation, observabilité).
 - Collaboration avec les équipes Data Science pour la **préparation et le déploiement de modèles prédictifs** (Azure ML, MLflow, AKS).

Environnement technique :
Azure Data Factory · Azure Databricks · ClickHouse · ADLS Gen2 · Airflow (AKS) · Azure DevOps · Prometheus · Grafana · Purview · Key Vault · Python · SQL
- Koacher

Alternant Data Engineer

Lyon

De décembre 2022 à décembre 2024

- Conception et mise en œuvre d'une **architecture Data Lakehouse Azure** couvrant ingestion, transformation et reporting analytique.
 - Développement de **pipelines automatisés** (Data Factory, Logic Apps) depuis APIs, bases SQL et Blob Storage pour centraliser les données produit et marketing.
 - Structuration du **Data Lake ADLS Gen2** en zones *bronze / silver / gold* avec contrôles de schéma et politiques de rétention automatisées.
 - Traitements et enrichissements sous **Azure Databricks (PySpark / Delta Lake)** pour fiabiliser et normaliser les jeux de données.
 - Conception d'un **modèle analytique Kimball** (dbt / Synapse Analytics) et création de marts dédiés aux KPIs métier : rétention, engagement, churn, activation.
 - Orchestration et industrialisation** via **Airflow (AKS)** et **Azure DevOps Pipelines**, avec automatisation des tests et des déploiements.
 - Mise en place du **monitoring et de la gouvernance** (Azure Monitor, Purview, Key Vault, Entra ID) pour garantir la traçabilité et la conformité.
 - Intégration d'une démarche **DataOps / MLOps** : infrastructure as code, suivi de la qualité des données, pipelines d'entraînement et scoring avec Azure ML & MLflow.
 - Conception de **dashboards Power BI** interactifs pour le suivi produit et la performance marketing.

Environnement technique :
Azure Data Factory · Azure Databricks (PySpark/Delta) · Azure Synapse Analytics · ADLS Gen2 · dbt Core · Airflow (AKS) · Azure DevOps · Azure Monitor/Log Analytics · Purview · Key Vault · Power BI · Python · SQL

Diplômes et formations

- IA School — Diplôme RNCP Niveau 7 Expert Data & Ingénierie de l'Intelligence Artificielle (2022-2024)
- IA School — Bachelor 3 Expert Data & Ingénierie de l'Intelligence Artificielle (2021-2022)
- Université de Lille — Licence Banque, Finance (2018-2021)