

# Yann Salako

Rennes, France | [yannsalako27@gmail.com](mailto:yannsalako27@gmail.com) | +33 6 46 33 38 44  
[linkedin.com/in/yann-salako](https://linkedin.com/in/yann-salako) | [github.com/Yann2808](https://github.com/Yann2808)

Étudiant en Master 2 MIAGE – Data & BI | Recherche d'un stage Data Engineering / Big Data – Mars 2026

## Profil

Étudiant en Master 2 MIAGE – Data Engineering, certifié Databricks en modélisation. Expert en pipelines et architectures Medallion / Star Schema, je maîtrise Spark/Databricks et les outils ETL (Talend). Orienté performance & problem solving, je recherche un stage de 6 mois pour industrialiser des flux de données critiques et complexes.

## Formation

- Master MIAGE – Data & Business Intelligence, Université de Rennes 2024 – 2026
- Licence Informatique – Parcours MIAGE, Université de Rennes 2023 – 2024

## Expériences Professionnelles

### AS PHARM GROUP – Paris

Avril – Juillet 2025

Stage – Développeur Data & Outils Décisionnels

- Automatiser l'ingestion de flux RH via des procédures SQL pour un dashboard en temps réel.
- Modéliser des couches de données consolidées pour fiabiliser le reporting sous Power BI.
- Cadrer les besoins métiers pour transformer des données brutes en indicateurs de performance (KPI).

### AVRIL SERVICES – Saint-Gérard-Croixanvec

Avril – Juillet 2024

Stage – Analyste BI & Data Engineering

- Optimiser les processus ETL du Data Warehouse pour garantir la haute disponibilité des données.
- Déployer des solutions de reporting complexes sur SAP BO BI4 et Power BI.
- Réaliser des PoC techniques sur l'intégration de nouveaux flux de données hétérogènes.

## Projets Data Engineering

### Analyse Big Data – Mobilité Électrique (Databricks & PySpark)

2025

- Exploiter Spark SQL pour traiter des données de télémétrie massives (IoT) de véhicules électriques.
- Optimiser le stockage via le format Delta Lake, réduisant le temps de traitement de 60%.
- Résultat : Identification automatisée des anomalies de charge via des Window Functions complexes.

### ShopNow – Plateforme Analytics E-commerce (Docker & Prefect)

2025

- Concevoir un pipeline ETL complet sous Python pour l'ingestion de données transactionnelles et clients.
- Orchestrer les workflows avec Prefect pour gérer les dépendances entre tâches et le monitoring des erreurs.
- Modéliser un entrepôt de données sous PostgreSQL en respectant un schéma en étoile (Star Schema).
- Industrialiser la solution via Docker, isolant les services (DB, ETL, BI) pour assurer la portabilité.
- Résultat : Mise en place d'un dashboard Metabase offrant une vision 360° des ventes en temps réel.

### Pilotage des Dépenses & Performance Financière (Talend)

2025

- Utiliser Talend Open Studio pour extraire et transformer des données comptables fragmentées.
- Modéliser le socle de données financier permettant l'automatisation du P&L et du suivi budgétaire.
- Résultat : Réduction drastique des erreurs de saisie et fiabilisation des reportings sous Power BI.

## Compétences Techniques

Data Engineering : Pipelines ETL/ELT, orchestration, automatisation, Data Quality — Python (Pandas), SQL, Talend (TOS & Cloud), Prefect, Snowflake.

Écosystème Big Data : Spark (RDD, DataFrames), Hadoop, Hive, HDFS ; architectures distribuées.

Bases de données : PostgreSQL, MySQL, SQL Server — modélisation, structuration, optimisation.

Versioning & Déploiement : Git, Docker, notions CI/CD, bonnes pratiques DevOps.

Visualisation & Reporting : Power BI, Tableau, Metabase — création d'indicateurs, dashboards.

Méthodes & Langues : Agile/Scrum, Français (Natif), Anglais (B1 – Technique).

## Certifications & Centres d'intérêt

- Certification Databricks : Data Modeling Strategies Obtenue en 2026
- Veille technologique : Suivi des blogs ingénierie et architectures de Data Modeling.
- Loisirs : Sport collectif, Gym, contribution à des projets personnels sur GitHub.