

# ISSAM SALMI

## Étudiant en ingénierie Big Data

@ issamsalmi55@gmail.com

+212 669081511

in issamsalmi

ISSAM-SALMI



Passionné par l'ingénierie des données et l'intelligence artificielle, je me spécialise dans la création de workflows ETL efficaces, la conception de pipelines de données robustes, et le travail avec les grands modèles de langage (LLMs). Dévoué à la résolution de problèmes complexes et au développement de solutions innovantes basées sur les données.

## SKILLS

- **Languages:** Python, Java, SQL
- **Pipelines:** Airflow, Dagster, Talend, Pentaho, ETL
- **Big Data:** Hadoop, Spark, Kafka, Hive
- **Databases:** PostgreSQL, MySQL, MongoDB, Hbase, Redis
- **Cloud:** AWS
- **OS:** Linux, Windows Server
- **Version Control:** Git, GitHub
- **Viz:** Power BI, Tableau, Matplotlib
- **LLM:** Fine-tuning, Embeddings, RAG
- **Containers:** Docker, Kubernetes

## SOFT SKILLS

- Collaboration et esprit critique
- Adaptabilité
- Travail d'équipe et communication

## CERTIFICATS

- **Oracle Cloud Infrastructure 2024 Generative AI Certified Professional**  
Oracle
- **AWS Cloud Quest: Cloud Practitioner**  
Amazon Web Services
- **Machine Learning Specialization**  
DeepLearning.AI
- **IBM Data Science**  
IBM
- **Associate Data Engineer in SQL**  
DataCamp

## BÉNÉVOLAT

- **Président & Responsable**  
Google Developer Groups  
Sur le campus  
2024 - En cours
- **Responsable IA**  
Google Developer Student Clubs  
2023

## LANGUES

- **Français :** Maîtrise professionnelle
- **Anglais :** Bonne maîtrise
- **Arabe :** Langue maternelle

## ÉDUCATION

### Cycle d'ingénierie en Systèmes d'Information et Big Data

#### École Nationale des Sciences Appliquées

Sept. 2021 – En cours

Berrechid, Maroc

### Baccalauréat

#### Lycée Ibno Yassine

Sept. 2019 – Juin 2020

Sidi Yahia, Maroc

## EXPÉRIENCE

### Stagiaire Data Engineer

#### Finea CDG

Août 2024 – Septembre 2024

Présentiel

- Développement de processus ETL avancés et de pipelines de données pour optimiser le flux et améliorer l'efficacité.
- Visualisation et présentation des insights dans Power BI pour soutenir la prise de décisions basée sur les données.
- **Technologies clés :** Talend Data Integration, Pentaho, Python, Power BI, Linux, PostgreSQL, SQL

### Développeur

#### Projet d'école privée

Juin 2022 – Juillet 2022

- Conception d'une application de bureau en Python pour la gestion des opérations scolaires.
- Intégration d'opérations OLAP pour permettre un suivi et un reporting avancés.
- **Technologies clés :** PyQt5, Python, SQL

## PROJETS

### Pipeline de Données – Architecture Cloud

- Conception d'un pipeline de données avec Airflow, Celery et Postgres pour gérer l'ingestion, la transformation et le stockage cloud de données Reddit.  
**Technologies utilisées :** Apache Airflow • AWS Glue • Amazon Athena • Redshift • PostgreSQL • Celery • Python

### Prédiction des Prix de Voitures d'Occasion – Marché Européen

- Développement d'un pipeline MLOps complet pour prédire les prix des voitures d'occasion, depuis le web scraping jusqu'au traitement automatisé et au réentraînement du modèle.  
**Technologies utilisées :** Airflow • Docker • Python • MongoDB • Web Scraping • Machine Learning • Linux

### Pipeline de Données – Analyse de Produits Affiliés

- Création d'un pipeline de données pour collecter automatiquement les données de produits affiliés et analyser leur performance avec Airflow, Selenium et PostgreSQL.  
**Technologies utilisées :** PostgreSQL • Power BI • Docker • Airflow • Selenium • Python