



# Aymane MAGHOUTI

## INGÉNIEUR DATA | BIG DATA | SOLUTIONS IA

✉ aymanemaghouti16@gmail.com

☎ +33 7 76 43 11 82

📅 23 ans

🔗 aymane-maghouti.github.io

### PORTFOLIO

🌐 @aymane-maghouti

🐙 @aymane-maghouti

### LANGUES

Arabe

Natif

Français

Courant professionnel

Anglais

Courant professionnel

### COMPÉTENCES TECHNIQUES

#### Langages & Frameworks

Python - Flask - FastAPI - SQL - Java

#### Bases de données

MySQL - PostgreSQL - SQL Server - Mongo DB

#### Big Data & BI

Hadoop (Hive, Hbase) - Spark - Kafka - Snowflake - DBT - Airflow - (SSIS/SSRS/Power BI)

#### Cloud

- **Azure** : ADLS, Data Factory, Databricks, Synapse Analytics, Azure ML
- **GCP** : BigQuery

#### Compétences Fonctionnelles

ETL/ELT - Data Modeling - Data Visualization - DB Optimisation

#### AI et Data Science

Machine Learning - Deep Learning - LLM - RAG - MLOps - Model Deployment

#### Outils & Méthodes

Git - Docker - Postman - Agile - Scrum

### COLLABORATION

Université Sol Plaatje (South Africa) | 07/30/2025

Conférencier invité en Data Engineering : Big Data, architectures et intégration IA.

Université Sol Plaatje (South Africa) | Jan. - Avril 2025

Collaboration sur une application de reconnaissance de la langue des signes en Cloud Hybrid.

Étudiant en Master 2 Big Data & Data Science, avec expérience en migration de plateformes et pipelines data intelligents. En recherche d'un stage PFE de 5 à 6 mois dès janvier 2026 en Data Engineering, Big Data ou IA.

### EXPÉRIENCES PROFESSIONNELLES

#### Consultant Data

De février 2025 à septembre 2025 INWI Casablanca, Maroc

- J'ai conçu et déployé des **pipelines temps réel** basés sur **Kafka** et **Python** pour capter les flux entrants, les stocker temporairement en CSV, puis générer automatiquement des fichiers consolidés intégrés au **Data Warehouse**. Ce processus traite plus de **500k événements / 15 min**, garantissant une ingestion fluide et continue des données.
- J'ai développé des workflows **SSIS** capables de traiter plus de **50 millions de lignes par jour**, réduisant le temps d'exécution de **2h à 4 min**.
- J'ai optimisé le **stockage** en mettant en place des **stratégies de purge partitionnées**, ramenant le traitement de **6h à moins d'une minute** et libérant plus de **350 Go** de données chaque mois.
- J'ai renforcé la **fiabilité du système** grâce à des **rapports SSRS** de supervision couvrant les jobs ETL, le stockage et la santé du système, assurant une **disponibilité de 99 %**.
- J'ai également créé des **tableaux de bord Power BI** pour le département **Qualité**, offrant un suivi en temps réel des services ADSL, Fibre et 4G, et améliorant la réactivité ainsi que la satisfaction client.
- Enfin, j'ai évolué dans un cadre **Agile (Scrum)**, participant activement aux sprints, daily meetings et reviews, tout en gérant le backlog via Jira pour garantir des livraisons incrémentales et alignées sur les besoins métier.

**Stack** : SQL Server, SSIS, Power BI, SSRS, Kafka, Python, Jira, Scrum.

#### Stagiaire Ingénieur Data & IA

De juin 2024 à septembre 2024 Shiftbricks Al Hoceima, Maroc

- J'ai automatisé la **collecte et l'ingestion** des documents depuis des plateformes juridiques marocaines, grâce à des scripts Python connectés à Azure Blob Storage pour un stockage sécurisé et centralisé des données.
- J'ai développé des **pipelines d'IA** capables de transformer les documents bruts en fichiers JSON, améliorant considérablement la qualité et l'exploitabilité des informations.
- J'ai intégré des **pipelines d'embeddings sémantiques** pour permettre la recherche intelligente et l'indexation des documents dans un cadre **RAG** (Retrieval-Augmented Generation), offrant des réponses contextuelles et précises.
- L'**automatisation** du processus a permis de réduire le temps de traitement manuel de 2h à 15 min, avec une fiabilité de 99 % sur plus de 5 DAGs orchestrés avec Airflow.
- J'ai également conçu une **application web de validation** facilitant la revue humaine, divisant le temps de vérification par 6 (2h → 20 min).
- Enfin, j'ai travaillé selon la méthodologie **Agile (Scrum)**, garantissant un delivery rapide et itératif.

**Stack** : Azure (Blob Storage, Document Intelligence, OpenAI), LLM, PostgreSQL, CosmosDB (MongoDB API), Python, Airflow, FastAPI, React.js, Docker, Git, Jira.

#### Stagiaire Ingénieur Data

De juillet 2024 à septembre 2024 ONCF Rabat, Maroc

- J'ai développé un **pipeline** Python orchestrant le chargement automatique de plus de 10 millions d'enregistrements par jour dans un **Data Warehouse SQL Server**, garantissant une ingestion fluide et performante.
- J'ai conçu un **modèle en étoile** (1 fait, 8 dimensions) optimisé pour réduire la latence des requêtes de 60 % et faciliter l'analyse multidimensionnelle.
- J'ai réalisé des **dashboards Power BI** interactifs intégrant des KPIs métiers grâce à DAX, offrant une vision claire et actionnable aux décideurs.
- L'automatisation quotidienne des mises à jour a assuré des rapports fiables et toujours à jour, tout en renforçant la collaboration **Agile (Scrum)** à travers les sprints et daily meetings.

**Stack** : SQL Server, Python, SQLAlchemy, Power BI (DAX), Jira.

### FORMATION

#### M2 en informatique

Depuis septembre 2025 École d'ingénieurs du Littoral Côte d'Opale Calais, France

- Spécialisations : Data Science, Big Data, Analyse & Visualisation, Machine Learning, Deep Learning, IA.

#### Diplôme d'ingénieur d'État en ingénierie des données

De septembre 2020 à juin 2025 École Nationale des Sciences Appliquées Al Hoceima, Maroc

- Spécialisation en **Data Engineering & Big Data** : bases de données, systèmes distribués, cloud, ML/DL/NLP.
- Major de promotion (1er de la classe).