



# Aymane MAGHOUTI

## DATA ENGINEER | BIG DATA | IA SOLUTIONS

[aymanemaghouti16@gmail.com](mailto:aymanemaghouti16@gmail.com)

+33 7 76 43 11 82

23 years old

[aymane-maghouti.github.io](https://aymane-maghouti.github.io)

### PORTFOLIO

[@aymane-maghouti](#)

[@aymane-maghouti](#)

### LANGUAGES

#### English

Full professional proficiency

#### French

Full professional proficiency

#### Arabic

Native proficiency

### HARD SKILLS

#### Langages & Frameworks

Python - FastAPI - SQL - Java

#### Databases

MySQL - PostgreSQL - SQL Server  
- Mongo DB (NoSQL)

#### Big Data & BI

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

#### Cloud

- GCP: BigQuery, Cloud Storage, Google Cloud Dataflow
- Azure : ADLS, ADF, Databricks, Synapse Analytics

#### Functional Skills

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

#### AI & Data Science

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

#### Tools & Methods

Git- Docker - Kubernetes -  
Postman - Agile - Scrum

### COLLABORATION

#### Sol Plaatje University (South Africa) | 07/30/2025

Guest speaker in Data Engineering: Big Data, architectures, and AI integration.

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

Collaboration on a hybrid-cloud application for sign language recognition.

Master's student in Data & AI (M2), specializing in Data Engineering & Big Data, with hands-on experience in ETL/ELT pipelines, monitoring systems, and decision-making dashboards. Currently seeking a 5-6 month end-of-study internship starting January 2026 in Data Engineering, Big Data, or Artificial Intelligence.

### WORK EXPERIENCE

#### Data Consultant

From February 2025 to September 2025 INWI Casablanca, Morocco

- Design and deployment of real-time pipelines based on Kafka and Python for ingesting incoming data streams, performing temporary CSV storage, and automatically generating files integrated into the Data Warehouse (over 500k events every 15 minutes).
- Development of SSIS workflows (with unit testing and integration testing) capable of processing more than 50 million rows per day, reducing execution time from 2 hours to 4 minutes.
- Storage optimization through partition-based purge strategies, reducing processing duration from 6 hours to under one minute and freeing up 350 GB of disk space per month.
- Implementation of an SSRS monitoring system ensuring the traceability and 99% availability of ETL jobs, storage processes, and system performance indicators.
- Design of Power BI dashboards for the Quality department, delivering real-time visibility into ADSL, Fiber, and 4G services, ultimately improving operational responsiveness and customer satisfaction.
- Maintenance of technical documentation and application of quality standards to ensure maintainability and reliability.
- Active participation in the Agile (Scrum) framework: contributing to sprints, daily meetings, reviews, and backlog management via Jira, ensuring incremental deliveries aligned with business needs.  
Stack: SQL Server, SSIS, SSRS, Power BI, Kafka, Python, Jira, Scrum.

#### Data & AI Engineer Intern

From June 2024 to September 2024 Shiftbricks AI Hoceima, Morocco

- Automation of the collection and ingestion of Moroccan legal documents using Python scripts connected to Azure Blob Storage, ensuring secure and centralized data storage.
- Development of AI pipelines that transform raw documents into structured JSON files, significantly improving data quality and usability.
- Integration of semantic embedding pipelines to enable intelligent search and contextual text indexing within a RAG (Retrieval-Augmented Generation) framework.
- Automated orchestration of more than 5 Airflow DAGs, reducing manual processing time from 2 hours to 15 minutes while achieving 99% reliability.
- Design of a web-based validation application that streamlines human review, reducing verification time by a factor of 6 (2h → 20 min).
- Work within an Agile (Scrum) environment, with iterative deliveries and continuous collaboration across AI, data, and product teams.  
Stack: Azure (Blob Storage, Document Intelligence, OpenAI), LLM, PostgreSQL, CosmosDB (MongoDB API), Python, Airflow, FastAPI, React.js, Docker, Git, Jira.

#### Data Analytics Intern

From July 2024 to September 2024 ONCF Rabat, Morocco

- Development of an automated Python pipeline orchestrating the loading of more than 10 million records per day into a BigQuery Data Warehouse, ensuring scalable and high-performance cloud ingestion.
- Design of an optimized star schema, reducing query latency by 60% and enabling efficient multidimensional analysis through dbt.
- Creation of interactive dashboards connected to BigQuery, integrating business KPIs and providing clear, actionable insights for decision-makers.
- Implementation of automated Airflow orchestration to manage and schedule daily data flows.  
Stack: GCP, Python, SQL, BigQuery, dbt, Airflow, Jira.

### EDUCATION

#### Master 2 - Data et AI

Since September 2025 Engineering School of the Littoral Côte d'Opale (EILCO) Calais, France

- Specializations: Data Science, Big Data, Data Analysis & Visualization, Machine Learning, Deep Learning, Artificial Intelligence.

#### State Engineer's Degree in Data Engineering

From September 2020 to June 2025 National School of Applied Sciences Al Hoceima, Morocco

- Specialization in Data Engineering & Big Data: databases, distributed systems, cloud, ML/DL/NLP.
- Ranked first in the class (valedictorian).