Aymane MAGHOUTI
DATA ENGINEER | BIG DATA | AI SOLUTIONS

- Calais (62100)
- Date of birth 05/07/2002
- aymanemaghouti16@gmail.com
- Open to remote work
- French
- +33 7 76 43 11 82
- aymane-maghouti.github.io

PORTFOLIO
- in @aymane-maghouti
- @aymane-maghouti

LANGUAGES

Arabic
Native

French
Courant

English
Courant

TECHNICAL SKILLS

Programming languages & Frameworks
Python - Flask - FastAPI - Java - Spring - Spring Boot - Spring DATA JPA - ReactJS - SQL

Databases
MySQL - Oracle - PostgreSQL - SQL Server - Mongo DB - Cassandra

Big data tech & BI tools
Hadoop - Spark - Kafka - Hive - HBase - Snowflake - DBT - Airflow - SSIS - SSRS - Power BI

Cloud Services
- **AWS** : S3, Athena, QuickSight, RDS
- **Azure** : Blob Storage, Data Factory, Databricks, Azure ML
- **GCP** : Bigquery, GCS, Looker

Functional Skills
ETL - ELT - Data modeling - Data warehousing - Data visualization - Database Optimisation

ML & DL & Data Science
Scikit Learn - Tensorflow (Keras) - Regression - Classification - Clustering.

INTERESTED

Learn new technologies

Sport (Football- basketball)

I am a Data & Big Data Engineer with experience in data analysis and data migration, real-time analytics, and BI solutions. I have worked on modernizing enterprise platforms (Oracle → SQL Server, SSIS, Kafka, Power BI) and developing AI-powered data pipelines. Currently pursuing a Master's (M2) in Complex Systems Engineering at EILCO, France, I focus on delivering scalable insights and intelligent analytics by combining Big Data, Cloud, and AI.

WORK EXPERIENCE

Freelance Data & AI Engineer (Part-time, Remote)
Since February 2025 Shiftbricks Al Hoceima, Morocco

- Designed and implemented AI-powered pipelines for processing large volumes of Arabic legal documents.
- Developed data ingestion, structuring, and enrichment workflows using **Azure Document Intelligence**, **MongoDB/Cosmos DB**, and **PostgreSQL**.
- Built embedding pipelines with **Azure OpenAI** and **Cohere** for semantic search and RAG applications.
- Automated data cleaning, metadata extraction, and hierarchical structuring of complex legal texts.
- Contributed to the architecture of a scalable knowledge management platform integrating **Big Data**, **AI**, and **Cloud** services.
  Tech Stack : Azure Cloud, PostgrSQL, MongoDB, Python, beautifulsoup

Data Platform Migration & Real-Time Analytics Implementation
From February 2025 to June 2025 SetGet Consulting - Client Inwi Casablanca, Morocco

- **Led the migration of INWI's legacy Oracle platform to Microsoft SQL Server**, modernizing the data ecosystem.
- **Developed automated ETL pipelines** using **SSIS**, significantly reducing processing time and operational effort.
- **Implemented data purge strategies with table partitioning**, reducing purge time from 6 hours to 1 minute and optimizing storage.
- **Enabled real-time Analytics** by integrating **Kafka** and **Python**, enhancing decision-making for the Quality team.
- **Delivered dynamic reporting solutions** through **Power BI dashboards** and **SSRS reports** for both operational and strategic monitoring.
- **Reduced total cost of ownership** by optimizing infrastructure and licensing expenditures.
  Tech Stack: SQL Server, SSIS, Power BI, SSRS, Kafka, Python.

Data Ingestion Pipeline for AI Application
From June 2024 to September 2024 Shiftbricks Al Hoceima, Morocco

- Design of an ingestion pipeline transforming unstructured data into structured data with the Medallion architecture.
- Ensuring data quality and consistency.
- Automation and orchestration of data workflows with Airflow.
- Development of a web application for data validation.
  Tech Stack: Python, FastAPI, ReactJs, MongoDB, PostgreSQL, Airflow, Prompt Engineering, Git-GitLab, Big Data Architecture, Docker

Data Collection and Analysis of Train Circulations and Activity Dashboard Design
From July 2024 to September 2024 ONCF Rabat, Morocco

- Developed an ETL pipeline using Python to collect data and load it into a SQL Server data warehouse.
- Designed a star schema for the data warehouse to facilitate efficient analysis.
- Created visualizations and dashboards using Power BI to present key performance indicators.
  Tech Stack: SQL Server, Python, SQLAlchemy, Power BI

Development of a modern e-commerce application with a microservices architecture
From July 2024 to October 2024 Marketing Confort Fès-Meknès, Morocco

- Develop web interfaces using React.js as well as REST APIs using the Spring ecosystem (Spring Boot, Spring Data JPA, etc.).
- Integrate the frontend with the backend through API consumption.
- Migrate the configuration of the microservices to Config Server (to centralize the configuration).
- Collaborate with the development team using Git-GitLab for version control.
- Follow Scrum methodology in an Agile environment using Jira for project management.
  Tech Stack: Java, Spring Boot, React.js, Typescript, Microservices Architecture, PostgreSQL, Git-GitLab, Postman

LATEST PROJECTS

End-to-End AI Solution Migration – From On-Premises to Azure Cloud

- Developed an AI-powered smartphone price prediction system integrating **FastAPI** (model API), **Spring Boot** (backend), **React Native** (mobile app), **PostgreSQL**, and **Power BI** dashboards.
- Implemented an **MLOps** pipeline with **MLFlow** and **Apache Airflow** to automate model training, versioning, and deployment.
- Migrated the entire architecture to **Azure Cloud**, leveraging Azure App Service, Azure Database for PostgreSQL, Azure Machine Learning, and Power BI Service.
- Delivered a scalable, cloud-native solution with automated retraining, real-time prediction APIs, and interactive dashboards for market insights.

Smartphone Price Prediction Using Lambda Architecture in a Big Data Environment

- Implemented a Lambda Architecture-based big data solution supporting real-time and batch processing.
- Integrated **Kafka, HBase,** and **PostgreSQL** for scalable ingestion, real-time views, and batch analytics.
- Built real-time prediction pipelines using **XGBoost** for low-latency inference.
- **Developed a Spring Boot service** to serve real-time **price predictions** via REST API.
- **Created Power BI dashboards** for batch analytics and historical trend analysis.
- **Orchestrated batch workflows** with **Apache Airflow** and managed storage using **HDFS**.

Event-Driven Architecture with Apache Kafka

- Engineered a real-time data streaming and processing pipeline leveraging **Apache Kafka**, **Apache Spark Streaming**, and **Debezium** for **Change Data Capture (CDC)**, enabling low-latency data ingestion and transformation.
- **Architected a comprehensive data platform** integrating **PostgreSQL**, **MySQL**, **Kafka**, **Spark Streaming**, **Spring Boot**, and **React.js**, facilitating scalable data flow and real-time analytics.
- Developed a real-time smartphone data analysis dashboard, utilizing **React.js** for the frontend and **Spring Boot** for RESTful API development on the backend.
- Implemented real-time data aggregation and persistence by streaming data into **MySQL** using **Spark Streaming**, supporting continuous analytics and operational reporting.

EDUCATION

Master Ingénierie des Systèmes Complexes

Since September 2025 Ecole d'ingénieurs du Littoral Côte d'Opale Calais, France

Data Engineer

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