







# HAMZA EL BELGHITI

## INGÉNIEUR ANALYTICS & IA

 hamagistral.dev  Hamagistral  Hamza El Belghiti  
 hamza.lbelghiti@gmail.com  Toulouse  (+33) 663045584

## FORMATION

### MASTER MIAGE INGÉNIERIE DES DONNÉES ET PROTECTION | 2024-2026

 **Faculté Sciences et Ingénierie (FSI), Université Toulouse III PAUL SABATIER**

**Cours:** SI Décisionnelles, Cloud et DevOps, Big Data, Projets IA, Projets Agiles, Gouvernance des SI

### MASTER SCIENCES ET INGÉNIERIE DES DONNÉES | 2022-2024

 **Faculté des Sciences de Rabat (FSR), Rabat (Maroc)**

**Cours:** Big Data, Business Intelligence, Datamining, Data Science, Administration Base Données

## EXPÉRIENCES PROFESSIONNELLES

### INGÉNIEUR ANALYTICS & IA | Septembre 2024 - En cours

 **Blent.ai, Paris (À distance) | ALTERNANCE**

- Migration de l'outil d'orchestration Mage vers Airflow et maintenance des pipelines.
- Développement d'un chatbot AI pour assister les étudiants sur la plateforme et automatisation de la correction des projets étudiants.
- Création de notebooks pédagogiques et d'exercices pratiques sur l'ingénierie analytique avec dbt, Airflow, et BigQuery.

**Technologies:** Airbyte, Airflow, BigQuery, dbt, Metabase, Langchain, HuggingFace (Mistral), Groq, OpenAI API, Pinecone

### INGÉNIEUR ANALYTICS & IA | Février - Juillet 2024

 **Blent.ai, Paris (À distance) | Stage PFE**

- Mise en place de **pipelines** ELT automatisés pour alimenter la Customer Data Platform.
- Développement de **Reverse ETL** pour les outils marketing tels que Hubspot et CustomerIO.
- Construction de dashboards sales et marketing pour analyser des KPIs des données clients.
- **Consulting DGAC:** Automatisation de l'extraction d'informations d'aérodromes à partir de PDFs

**Technologies:** Airbyte, MongoDB, Hubspot, Mage, BigQuery, dbt, Looker, Open Metadata Python, PDFplumber, Langchain, OpenAI API, ChromaDB, Streamlit

### DÉVELOPPEUR FULL STACK | Juillet - Septembre 2023

 **3D SMART FACTORY, Mohammedia |  GITHUB  DEMO**

- Déploiement d'un modèle de deep learning 3D à l'aide d'une API REST sur AWS Cloud.
- Réduction du coût de déploiement de **\$85 par mois fixe à \$1,3 pour 1000 requêtes.**
- Conception de l'interface utilisateur avec des librairies 3D et déploiement sur Vercel.

**Technologies:** Vite, React, Tailwindcss, Three, Vtkjs, Firebase, Pytorch, FastAPI, Docker, Vercel AWS ECR + Lambda + API Gateway

### DATA SCIENTIST JUNIOR | Mai - Juin 2023

 **OMDENA ÎLE-DE-FRANCE, À Distance |  GITHUB  DEMO**

- Collaboration avec une équipe internationale de 10 personnes pour développer un chat-bot qui facilite la recherche de transport public alternatif pendant les grèves en Île-de-France.
- Web scraping du site RATP pour l'état du trafic avec une **pipeline ETL sur AWS.**

**Technologies:** BeautifulSoup, AWS S3 + Lambda + Cloudwatch, Navitia API, Folium, Streamlit

## PROJETS

### ANALYSE DES DONNÉES D'UN COMMERCE DE DETAIL EN LIGNE

**PROJET FINAL DE ZOOCAMP |  GITHUB  DEMO**

- Conception et mise en œuvre d'une **pipeline de données ELT à l'aide d'Airflow sur GCP.**
- Ingestion des données dans GCP Storage Bucket et BigQuery et transformation avec dbt.
- Construction d'un dashboard pour **identifier et suivre le revenu total par pays et produit.**

**Technologies:** Python, Terraform, Airflow, dbt, Soda, GCP Storage + BigQuery, Docker, Looker



## COMPÉTENCES

### PROGRAMMATION

Python • Java • C

### BASES DE DONNÉES

BigQuery • PostgreSQL • MongoDB • Oracle • Snowflake

### INGÉNIERIE DES DONNÉES

Terraform • Airbyte • Mage • Airflow • dbt • databricks • Spark • Kafka • PowerBI • Looker Studio • Metabase

### LLM & IA

Langchain • HuggingFace • Pinecone

### DEVELOPEMENT WEB

HTML/CSS • Typescript • React • Nextjs • FastAPI • Django • Streamlit

### CLOUD

Google Cloud • AWS • Azure

### OUTILS

Linux • Docker • Git • ClickUp

### LANGUES

Arabe • Anglais (C2) • Français (C2)

## CERTIFICATS

### GCP Professional Data Engineer

Délivré par Google

**Juin 2024**

### DATA ENGINEER

Délivré par DataCamp

**Septembre 2023**

### MACHINE LEARNING

Délivré par FreeCodeCamp

**Avril 2022**

## PARASCOLAIRE

### STREAMLIT FOR EDUCATION

Étudiant Ambassadeur Leader | 2023-2024

### INNOVASUS - FSR CLUB

Membre | 2023-2024

## CENTRES D'INTÉRÊT

 Blogging

 Hackathons

 Échecs

 E-sports