







HAMZA EL BELGHITI

RECHERCHE D'UNE ALTERNANCE À TOULOUSE (24 MOIS)

 Hamagistral.me  Hamagistral  Hamza El Belghiti
 hamza.lbelghiti@gmail.com  Toulouse  (+212) 660081086

FORMATION

MASTER INGÉNIERIE DES DONNÉES ET PROTECTION | 2024-2026

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

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

MASTER SCIENCES ET INGÉNIERIE DES DONNÉES | 2022-2024

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

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

LICENCE MATHÉMATIQUES APPLIQUÉES | 2019-2022

 **Faculté des Sciences et Techniques de Mohammedia (FSTM), Mohammedia**

EXPÉRIENCES PROFESSIONNELLES

INGÉNIEUR DATA | Février - Juillet 2024

 **Blent.ai, Paris (À distance)**

- 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 les données clients.
- **Consulting DGAC:** Automatisation de l'extraction d'informations d'aérodromes à partir de PDFs

Technologies: Airbyte, MongoDB, Mage, BigQuery, dbt, Vertex AI, 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

 **OMDNA 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 | 2023

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

ANALYSE D'OFFRES D'EMPLOI D'INGÉNIEURS DE DONNÉES | 2023

PROJET DATA SCIENCE & ENGINEERING |  **GITHUB**  **DEMO**

- Web scraping hebdomadaire des offres d'emploi d'ingénieurs de données sur Glassdoor.
- Conception d'une **pipeline ETL** avec **Mage** qui extrait et transforme les données depuis S3.
- Prédiction des salaires de data engineers avec **86% de précision** avec Random Forest.
- Création d'un dashboard pour analyser les salaires et compétences des ingénieurs de données.

Technologies: Python, Selenium, AWS S3, Boto3, Mage, Scikit-learn, Altair Viz, Streamlit



COMPÉTENCES

PROGRAMMATION

Python • Java • C

BASES DE DONNÉES

MongoDB • PostgreSQL • SQL Server • BigQuery • Snowflake

INGÉNIERIE DES DONNÉES

Terraform • Airbyte • Mage • Airflow • dbt • Databricks • Spark • Kafka • PowerBI • Looker Studio

DEVELOPEMENT WEB

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

CLOUD

Google Cloud • AWS • Azure • Firebase

OUTILS

Linux • Docker • Git • Clickup

LANGUES

Arabe • Anglais (C2) • Français (C2)

CERTIFICATS

GCP Professional Data Engineer

En cours de préparation
(Date examen: 8 Juillet)

DATA ENGINEER

Délivré par **DataCamp**
Septembre 2023

MACHINE LEARNING - PYTHON

Délivré par **FreeCodeCamp**
Avril 2022

CS50 WEB PROGRAMMING

Délivré par **Harvard**
Février 2022

PARASCOLAIRE

STREAMLIT FOR EDUCATION

Étudiant Ambassadeur Leader | 2023-2024

INNOVASUS - FSR CLUB


Membre | 2023-2024

CENTRES D'INTÉRÊT

 Blogging

 Hackathons

 Échecs

 E-sports