Ouamrhar Hamza

DATA ENGINEER

@ business.hamza2003@gmail.com   https://github.com/HamzaOuamrhar   ☎ +212640583630

in https://ma.linkedin.com/in/hamza-ouamrhar   ◐ https://medium.com/@houamrha

EDUCATION

Master in Data Engineering - ENSA, Khouribga, Morocco | Expected Graduation: June 2026

Software engineering - 1337, 42 Network, Khouribga, Morocco | Expected Graduation: June 2026

PROJECTS

- **Daily News Analytics Pipeline**

Batch pipeline for analyzing news and tweets to extract and predict trending entities.

• Designed Airflow DAG orchestrating tasks for ingestion, validation, Kafka publishing, Spark processing

• Implemented distributed NER extraction using a three-node Spark cluster

• Built next-day entity forecasting with XGBoost, storing predictions in PostgreSQL for trend analytics

• Containerized the full workflow with Docker Compose managing microservices and persistent volumes

Technologies: Python, Apache Airflow, Apache Kafka, Apache Spark, PostgreSQL, Docker, XGBoost, Grafana

- **Air Quality LLM Pipeline**

Cloud-native ETL and natural-language analytics platform for global air quality data.

• Built scalable Dataflow ETL processing 10,000+ compressed CSV records into partitioned and clustered BigQuery tables, reducing query cost by 60%

• Provisioned GCP infrastructure using Terraform modules across BigQuery, Cloud Storage, and Dataflow

• Developed a Text-to-SQL interface using Vertex AI Gemini with schema-aware prompts

• Created a Streamlit dashboard supporting real-time queries, visualizations, and templates for pollutant analysis across monitoring stations

Technologies: Python, GCP (BigQuery, Dataflow, Vertex AI, Cloud Storage), Terraform, Streamlit, Pandas, SQL

- **Monitoring System**

Provide end-to-end observability for a microservices-based full-stack application. (Individual contribution to a team project.)

• Deployed ELK stack with automated ILM policies, S3 snapshot backups, and data retention to optimize log storage

• Implemented Prometheus ecosystem with Node, cAdvisor, and Nginx exporters feeding Grafana dashboards for real-time insights

• Added Thanos for long-term metrics storage and integrated Alertmanager with Discord for proactive incident detection

• Configured automated certificate generation, TLS/SSL, and role-based access control across monitoring components

Technologies: Docker, Elasticsearch, Logstash, Kibana, Prometheus, Grafana, Thanos, Amazon S3, Amazon IAM

EXPERIENCE

Front-End Intern

OCP Group - Khouribga | Apr 2023 - Jun 2023

- Completed UI development tasks using React, contributing small reusable components.

TECHNICAL SKILLS

Python, PySpark, Apache Spark, SQL, Apache Airflow

Google Cloud Platform (GCP), AWS, Terraform, Docker, Git

PostgreSQL, MySQL, Redis, SQLite, MongoDB

PERSONAL SKILLS

Technical communication, team collaboration, and requirement understanding
Analytical problem solving and effective time management