

MOHAMAD ABU ABBAS – DATA ENGINEER

Lebanon, Beirut
+961 71 975 084
mohamadabuabbas@gmail.com | [GitHub](#) | [LinkedIn](#)

Results-driven Data Engineer with proven ability to design, build, and deploy both batch and real-time data pipelines. Expert in ETL/ELT workflows, orchestration, data warehousing, and business intelligence solutions. Proficient in Apache Airflow, Prefect, Kafka, Spark Structured Streaming, PostgreSQL, Power BI, and AWS Cloud services. Brings strong backend and DevOps expertise for production-grade, scalable, and reliable data solutions.

PROJECTS & EXPERIENCE

Global E-Commerce Sales Data Pipeline

- Designed and implemented an end-to-end batch data pipeline to process sales data from multiple CSV and Excel sources.
- Cleaned, transformed, and aggregated data using Pandas. Loaded data into a PostgreSQL data warehouse.
- Orchestrated workflows with Apache Airflow.
- Built Power BI dashboards for sales trends and regional analysis.

Real-Time Ride-Hailing Data Pipeline

- Built a real-time streaming data pipeline simulating ride-hailing trip data ingestion with Apache Kafka.
- Processed streaming data using Apache Spark Structured Streaming.
- Stored processed data in PostgreSQL.
- Visualized live analytics using Apache Superset.
- Managed orchestration and monitoring with Prefect.

WatchGuard AI – Full-Stack AI Threat Detection System

- Built a full-stack AI platform to detect misinformation, toxicity, and deepfakes using Hugging Face Transformers.
- Designed scalable architecture: Fast API backend, Redis queues, Celery workers, PostgreSQL database.
- Integrated multimodal ML tasks: image classification, object detection, ASR, and text classification.
- Developed a React + TypeScript dashboard for content moderation and threat visualization.
- Containerized using Docker & Docker Compose; implemented CI/CD pipelines with GitHub Actions.

Smart Audit Ops AI- Modular DevOps Automation Tool

- Developed a containerized platform to automate code quality checks, sentiment analysis, and UI compliance.
- Built a Fast API backend with real-time ML inference and document-based QA.
- Designed GitHub Actions pipelines for continuous integration and Docker image generation.
- Implemented observability stack: Prometheus, Grafana, Loki, and Redis.

EDUCATION

BACHELOR IN COMPUTER SCIENCE, LEBANESE INTERNATIONAL UNIVERSITY LEBANON, (2023)

CERTIFICATES / INTERNSHIP / COURSES

- AWS Certified Cloud Practitioner (2025)
- 42 Beirut – 4- week Assessment Program (2024)
- RASPBERRY PI- LIU Robotics Club (2023)
- IoT, Arduino Level 1 & 2- LIU Robotics Club (2022- 2023)
- L21 English Course, LINGUAPHONE Institute LTD., LONDON UK (2019)

SKILLS

- | | |
|---|---|
| • Programming & Data Processing: Python, Java, Pandas, Py Spark, NumPy, SQL | • Databases: PostgreSQL, SQLite, Firebase |
| • Data Engineering: ETL/ELT, Data Warehousing, Batch & Streaming Pipelines | • CI/CD & Infra: Git, GitHub Actions, Docker, Docker Compose, Prometheus, Grafana, Loki |
| • Orchestration: Apache Airflow, Prefect | • Visualization: Power BI, Apache Superset |
| • Streaming: Apache Kafka, Apache Spark Structured Streaming | • OS & Scripting: Bash, Linux |
| | • Backend & APIs: Fast API, Flask, REST APIs, Redis, Celery |

LANGUAGES

Arabic (Native)	English (Fluent)	Spanish (Beginner)
-----------------	------------------	--------------------