Mohamed El Korchi

Data Engineer · Data Platforms · AI

Agadir, Morocco · +212 652 923 251
elkorchimohameddev@gmail.com · mohamedelkorchi.com
linkedin.com/in/mohamedelkorchi · github.com/Med-Dev-99

PROFILE

Data Engineer with 3+ years building scalable data platforms, ELT pipelines and real-time systems in production with Apache NiFi, Azure Databricks (Spark), ADLS Gen2, Airflow, dbt and PostgreSQL/PostGIS. Recently built the data platform and grounded AI agent of FIREBRIDGE for the NASA Space Apps Challenge 2026: 666.9M satellite records acquired, audited and served through versioned releases. Strong foundation in machine learning, NLP and LLM agents.

EXPERIENCE

AI & Data Engineer · Freshtrack

Dec 2023 – Present

Agadir · On-site · Apache NiFi, Azure Databricks, ADLS Gen2, Spark, PostgreSQL, n8n, Python

- Designed the data-platform architecture and Apache NiFi pipelines, with failure handling, backpressure control and automated alerting.
- Built ELT on Azure ADLS Gen2 and Databricks (Spark) with Parquet storage, improving analytics performance and reducing cloud costs.
- Real-time maritime tracking that merges carrier feeds (Maersk, CMA CGM, MSC) into a single PostgreSQL source of truth with live alerts.
- Road-transport ETA prediction from historical GPS trajectory segmentation and real-time tracking.
- OCR document service for order workflows: classification, extraction, validation and mismatch alerts.
- Self-hosted n8n booking workflows, an AI analytics assistant, a vessel-tracking service and carrier-performance dashboards.

Data Engineering Intern · Brocoli Data

Aug 2023 – Dec 2023

Remote · Delta Live Tables, BigQuery, AWS Lambda, GCP

- Built scalable extraction and loading pipelines with Delta Live Tables.
- Serverless data workflows on Google Cloud Functions and AWS Lambda across S3, GCS and BigQuery.

ML & Python Developer Intern · AGILISYS

Jul 2023 – Nov 2023

Remote · PaddleOCR, OpenCV, NER, Flask

- Document extraction with PaddleOCR, zonal extraction, NER and table parsing; OpenCV preprocessing and a custom-trained NER model.
- REST API in Flask to automate extraction and integration workflows.

Master's Research, Big Data & AI · Ibn Zohr University

Feb 2023 – Jul 2023

Agadir · PyTorch, GNN

- Session-based recommendation with hybrid GAT, GCN and attention, 4% better than state-of-the-art models.

PROJECTS

FIREBRIDGE: NASA Space Apps Challenge 2026

Data platform & AI agent · team of 4 · ogote.com

Airflow 3 · dbt · Parquet · DuckDB · PostgreSQL/PostGIS · FastAPI · Fastify · React · Docker · GitHub Actions

- Built resumable acquisition and audit of NASA FIRMS fire records for 5 satellites (MODIS, VIIRS) and Landsat: 666.9M rows in a 23.3 GiB Parquet lake, 4,202 download windows audited, orchestrated by 10 Airflow DAGs.
- Published immutable, versioned releases to PostgreSQL/PostGIS with dbt models and data tests; regional spatial indexes cut analysis time-to-ready from 25.5 s to 4.7 s.
- Built a grounded AI agent (FastAPI, Groq LLMs): evidence engine and number guard that verify every figure, rate-limit-aware fallback across 4 models, short/long-term memory, prompt-injection screening in 7 languages; 138 tests.
- Main contributor to the React/MapLibre web app and Fastify/BullMQ API; CI/CD on GitHub Actions (tests, secret and image scans, GHCR) deploying hardened Docker Compose stacks.

Docker Inspector: open-source Linux desktop app

Sole author · github.com/Med-Dev-99/Docker-inspector

Electron · React 19 · TypeScript · Node.js · SQLite · WebSocket · Docker Engine API

- Records container metrics and events locally; diagnoses crashes in plain English with evidence, last logs and pre-crash metrics; 7-day time-machine history.
- Configuration auditor flags UFW-bypassing ports, secrets in environment variables and missing limits, with one-click fixes; hardened image (read-only, no capabilities, token and Host/Origin checks).

IoT Sensor Streaming Platform

Kafka · Spark · MongoDB Atlas · Grafana

- Ingested simulated sensor streams with Kafka, stored them in MongoDB Atlas, analysed at scale with Spark and built Grafana dashboards.

Twitter ETL & Analytics

Python · Pentaho PDI · PostgreSQL

- ETL pipeline from the Twitter API into PostgreSQL with analytical reporting.

SKILLS

Data engineering	Apache Spark, Kafka, NiFi, Airflow, dbt, Delta Live Tables, Parquet/Arrow, DuckDB, ELT, streaming
Cloud	Azure Databricks, ADLS Gen2, BigQuery, AWS S3 & Lambda, Google Cloud Storage & Functions
Databases	PostgreSQL, PostGIS, MongoDB, Cassandra, SQLite, Redis
AI / ML	PyTorch, TensorFlow, scikit-learn, HuggingFace, LangChain/LangGraph, LLM agents, OCR, NLP
Backend & DevOps	FastAPI, Flask, Node.js/Fastify, Docker, GitHub Actions, Linux, n8n, REST APIs
Languages	Python, SQL, TypeScript, Java, Go, C/C#, Dart

EDUCATION

Master's, Big Data & Artificial Intelligence

Ibn Zohr University, FPT · 2021 – 2023

Licence, Computer Engineering

Ibn Zohr University, FPT · 2020 – 2021

IT Development Technician Diploma

ISGI Laayoune · 2017 – 2019

AWARDS & CERTIFICATIONS

NASA Space Apps Challenge 2026: Galactic Problem Solver

NASA · Nov 2026

Machine Learning Specialization

DeepLearning.AI & Stanford Online

Deep Learning Specialization

DeepLearning.AI

Natural Language Processing Specialization

DeepLearning.AI