

# Haroune Mohammedi Senior Data Engineer

 haroune.me  mohammedi.haroun@gmail.com  0626850080  GitHub  LinkedIn  Paris

Specialized in large-scale distributed data pipelines on Databricks/PySpark, currently at Engie, processing 50M+ rows/day of high-frequency data, optimizing workloads (up to 90%) and migrating legacy Oracle ETL workflows. Operates across the full project lifecycle, from functional analysis and architecture design to deployment and production support. Strong focus on understanding domain challenges and bridging the gap between technical implementation and business needs.

## EXPERIENCE

ENGIE | SENIOR DATA ENGINEER

Feb 2024 -- Present | Paris, France

- Designed and implemented large-scale data pipelines processing 50M+ rows/day of high-frequency energy consumption data (30-min intervals) on Databricks/PySpark.
- Migrated legacy ETL workflows from PL/SQL to PySpark on Databricks, reducing execution time from several weeks to a few hours and generating substantial compute cost savings.
- Optimized PySpark pipelines achieving 90%+ performance improvement (10 hours → under 1 hour) through Spark tuning, partitioning strategies, and scalable architecture design.
- Led projects end-to-end: functional analysis, architecture design, development, testing, deployment (MEP), and production support (RUN).
- Developed strong functional expertise in the energy domain, with a genuine interest in business challenges and operational processes.
- Ensured data quality and reliability of production pipelines in a critical context.

QUADRATIC | SENIOR DATA ENGINEER

May 2022 -- Jan 2024 | Paris, France

- Designed and implemented a real-time blockchain data integration pipeline, synchronizing on-chain events from Web3 smart contracts into MongoDB, making blockchain data accessible to Web2 applications.
- Designed the MongoDB schema and data models optimized for query performance and scalability, translating raw blockchain event structures into clean, queryable documents.
- Defined GraphQL schema and data access patterns to expose structured contract and transaction data to frontend consumers.

BIGMAMA TECHNOLOGY | SENIOR DATA ENGINEER

Jul 2019 -- Apr 2022 | Algiers, Algeria

- Core developer of a cloud-agnostic MLOps platform built on open standards (Docker, Kubernetes) supporting GCP, on-premise, and multi-cloud deployments.
- Deeply integrated MLflow for automated experiment tracking and reproducibility, and Seldon Core for scalable model serving on Kubernetes.
- Implemented one-click ML model packaging and deployment workflows.
- Evaluated, benchmarked, and integrated open-source MLOps tooling, developing deep expertise across the ML infrastructure ecosystem.
- Participated in architecture decisions and translated business requirements into scalable, maintainable platform features.

BIGMAMA TECHNOLOGY | DATA ENGINEER

Sep 2017 -- Jun 2019 | Algiers, Algeria

- Built and deployed a real-time health monitoring platform for elderly people living alone, using Akka, Spark, Kafka, Elasticsearch, and Firebase — with high-availability and fault tolerance as non-negotiable requirements.
- Designed complex distributed applications based on the actor model, contributing to architecture decisions across multiple systems.

## UNIVERSITY AND PERSONAL PROJECTS

HOMELAB | PROXMOX, TRUENAS, NGINX, VPN

2025 -- Present | Personal Project

Self-hosted personal infrastructure built on Proxmox and TrueNAS — serving as a private cloud, a 3-2-1 backup system, and a personal sandbox for exploring new technologies including local LLM deployment.

VIDEO PLAYER CONTROLLED BY HAND GESTURES | SCALA, AKKA, KERAS

2018 | University Project

Trained a hand gesture detection neural network with Keras, integrating the model into a fully asynchronous, reactive application built on the Actor model with Akka.

## SKILLS

Languages: Python • Scala/Akka • SQL • Java • Shell

Data Engineering: PySpark • Apache Spark • Databricks • Kafka • Elasticsearch

MLOps: MLflow • Seldon Core • Jupyter

Cloud, DevOps & Infrastructure: Docker • Kubernetes • Linux • Proxmox • Gitlab CI/CD • AWS • GCP

Databases & APIs: MongoDB • GraphQL • Flask • FastAPI

## EDUCATION

UNIVERSITÉ HOUARI BOUMEDIËNE (USTHB)

**PhD in Artificial Intelligence**

Sep 2019 -- Ongoing | Algeria Data services management in multi-cloud environments

**Master in Artificial Intelligence**

Sep 2017 -- Jun 2019 | Algeria Final project: Graph-based recommender system.

**Bachelor in Computer Science**

Sep 2014 -- Jun 2017 | Algeria Final project: Web services relationship prediction using Spark in collaboration with University of Michigan, USA.

## REFERENCES

**Pr. Youcef Aklouf**, Professor at USTHB | R&D Manager at Renault, France

 yaklouf@yahoo.fr

**Jugurtha Hadjar**, CTO, BigMama Technology, Algeria

 jugurtha.hadjar@gmail.com