

Haroune Mohammedi Senior Data Engineer

<http://haroune.me> mohammedi.haroun@gmail.com [0626850080](tel:0626850080) [GitHub](#) [LinkedIn](#) [Paris](#)

Spécialiste de pipelines de données distribués à grande échelle sur Databricks/PySpark, actuellement chez Engie : 50M+ lignes/jour de données haute fréquence, workloads optimisés jusqu'à 90%, migration d'ETL Oracle legacy. Intervient de bout en bout, de l'analyse fonctionnelle au support de production. Au cœur de ma démarche : faire le pont entre l'implémentation technique et les besoins business.

EXPÉRIENCE

ENGIE | SENIOR DATA ENGINEER

Févr. 2024 -- Présent | Paris, France

- Pipelines à grande échelle : 50M+ lignes/jour de consommation énergétique haute fréquence (30 min) sur Databricks/PySpark.
- Migration ETL legacy PL/SQL → PySpark/Databricks : exécutions ramenées de plusieurs semaines à quelques heures.
- Optimisation PySpark : 90%+ de gain (10 h → <1 h) via tuning, partitionnement et architecture scalable.
- Pilotage de bout en bout : analyse fonctionnelle, architecture, dev, tests, déploiement (MEP) et support production (RUN) sur infrastructure critique.

QUADRATIC | SENIOR DATA ENGINEER

Mai 2022 -- Janv. 2024 | Paris, France

- Pipeline temps réel d'intégration blockchain : événements on-chain des smart contracts synchronisés vers MongoDB, exposant la donnée Web3 aux applications Web2.
- Conception du schéma MongoDB et des modèles de données optimisés performance/scalabilité, à partir des structures brutes d'événements blockchain.
- Définition du schéma GraphQL et des patterns d'accès pour exposer contrats et transactions aux consommateurs frontend.

BIGMAMA TECHNOLOGY | SENIOR DATA ENGINEER

Juil. 2019 -- Avr. 2022 | Alger, Algérie

- Développeur principal d'une plateforme MLOps cloud-agnostique sur standards ouverts (Docker, Kubernetes), supportant GCP, on-premise et multi-cloud.
- Intégration approfondie de MLflow (tracking d'expériences) et Seldon Core (serving scalable sur Kubernetes).
- Workflows de packaging et de déploiement de modèles en un clic ; benchmarking de l'outillage MLOps open source.

BIGMAMA TECHNOLOGY | DATA ENGINEER

Sept. 2017 -- Juin 2019 | Alger, Algérie

- Plateforme temps réel de monitoring de la santé pour personnes âgées vivant seules (Akka, Spark, Kafka, Elasticsearch, Firebase) — haute disponibilité et tolérance aux pannes non négociables.
- Conception d'applications distribuées complexes basées sur le modèle acteur, contributions aux décisions d'architecture.

PROJETS UNIVERSITAIRES ET PERSONNELS

HOMELAB | PROXMOX, TRUENAS, NGINX, VPN

2025 -- Présent | Projet personnel

Infrastructure personnelle self-hosted sur Proxmox/TrueNAS — cloud privé, sauvegarde 3-2-1 et bac à sable pour explorer de nouvelles technos (déploiement de LLM en local, etc.).

LECTEUR VIDÉO CONTRÔLÉ PAR GESTES DE LA MAIN | SCALA, AKKA, KERAS 2018 | Projet universitaire

Entraînement d'un réseau de neurones de détection de gestes avec Keras, intégré dans une application réactive et entièrement asynchrone bâtie sur le modèle acteur avec Akka.

COMPÉTENCES

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

Bases de données & APIs : MongoDB • GraphQL • Flask • FastAPI

FORMATION

UNIVERSITÉ HOUARI BOUMEDIËNE (USTHB)

Doctorat en Intelligence Artificielle

Sept. 2019 -- en cours | Algérie
Gestion des services de données en environnements multi-cloud

Master en Intelligence Artificielle

Sept. 2017 -- Juin 2019 | Algérie
PFE : système de recommandation à base de graphes.

Licence en Informatique

Sept. 2014 -- Juin 2017 | Algérie
PFE : prédiction de relations entre services web avec Spark, en collaboration avec l'Université du Michigan, USA.

RÉFÉRENCES

Pr. Youcef Aklouf,
Professeur USTHB | R&D
Renault, France
[✉ yaklouf@yahoo.fr](mailto:yaklouf@yahoo.fr)

Jugurtha Hadjar, CTO,
BigMama Technology
[✉ jugurtha.hadjar@gmail.com](mailto:jugurtha.hadjar@gmail.com)