



Hamza Ben Marzouk

Data & AI Engineer · Paris

contact@hamzabenmarzouk.com linkedin.com/in/bmhamza/

À PROPOS

Data & AI Engineer avec plus de 7 ans d'expérience, dont 5 en cabinet de conseil (OCTO Technology), dans des environnements exigeants : banque, assurance, énergie, médias, cloud. Je conçois et fiabilise des plateformes data (Snowflake, dbt, Databricks, Spark) et je mets l'IA en production : agents, évaluation de LLM, adoption par les équipes métier. Je fais le lien entre métier et technique, je fais monter les équipes en compétence et je laisse derrière moi des systèmes documentés, testés et maintenables.

CE QUE JE PROPOSE

Plateformes data modernes

Concevoir, reprendre ou moderniser une plateforme Snowflake ou Databricks : ingestion, modélisation dbt, orchestration, infrastructure as code et CI/CD.

Fiabilité, sécurité & gouvernance

Data quality, observabilité, data contracts, gestion des accès et conformité RGPD : rendre la plateforme digne de confiance et auditable.

IA en production

Agents IA, évaluation de LLM, chatbots et adoption de l'IA par les métiers : passer du POC à un usage en production, mesuré.

EXPÉRIENCE

Wakam

févr. 2024 – Aujourd'hui

Data Platform Engineer

avr. 2026 – Aujourd'hui

Retour sur la Data Platform pour industrialiser la sécurité, la gouvernance et la fiabilité de la plateforme.

- Conception d'un système GitOps d'accès temporaires à Snowflake : demande par pull request, validation, révocation automatique et journal d'audit, y compris pour les données sensibles (RGPD)
- Passage de la configuration Snowflake en infrastructure as code (Terraform) : politiques réseau, warehouses, comptes de service, tâches de sécurité ; traitement des recommandations du benchmark CIS
- Mise en place d'un miroir de données réglementaire (résidence des données UK) vers Azure Blob Storage, avec contrôle d'intégrité quotidien indépendant
- Fiabilisation de l'orchestration Prefect : isolation dev/prod, supervision des runs en retard, alerting Slack
- Migration de la CI/CD d'Azure DevOps vers GitHub Actions, environnements de recette éphémères par pull request, quality gate SonarCloud, migration Python 3.13 / uv
- Outillage agentique pour l'équipe : skills Claude Code automatisant PR, tickets, diagnostics et opérations courantes

Stack : Snowflake · Terraform · Prefect · dbt · Azure · GitHub Actions · Python · Claude Code

AI Engineer — équipe KamAI

août 2025 – mars 2026

Développement et déploiement de solutions d'IA, avec un double objectif : industrialiser l'évaluation des agents IA et maximiser l'adoption des outils IA par les collaborateurs.

- Conception et déploiement d'un système d'évaluation de bout en bout pour les agents IA à fort impact métier
- Pipeline de génération automatique de jeux d'évaluation synthétiques, validés par les experts métier dans une application Retool
- Pipelines d'évaluation automatique avec suivi centralisé des résultats sur Langfuse
- Mise en production d'un chatbot conversationnel sur wakam.com pour les assurés (workflows agentiques Dify)
- Pilotage de l'adoption de la plateforme Dust ; hackathons bimensuels (satisfaction moyenne $\geq 4/5$) et accompagnement des métiers sur des cas d'usage à forte valeur

Stack : Dust · Dify · Langfuse · Retool · LLMs · Python

Data Engineer — Data Platform

févr. 2024 – juil. 2025

Deux périmètres stratégiques, PDX (Partner Data eXchange) et DPF (Data Platform Foundation), pour garantir la fiabilité, la scalabilité et l'exploitabilité de la plateforme data.

- Référent technique data contracts : montée en autonomie de 3 équipes partenaires dans la définition et l'application des standards
- Mise en production d'une chaîne de bout en bout fiabilisée (échange partenaires → Snowflake), avec documentation, transfert de compétences et processus formalisés

- Optimisation du run mensuel : 50 % des demandes de support data traitées en self-service
- Observabilité complète via Datadog (alerting, dashboards) ; refonte du framework data quality et intégration d'Elementary
- Conception de pipelines ETL / reverse ETL et restructuration par domaine du code dbt ; réduction significative des incidents data

Stack : Snowflake · dbt · dlt · Databricks · Elementary · Datadog · Python

OCTO Technology

févr. 2019 – janv. 2024

Consultant — missions client

Scaleway — Consultant Data Ops

sept. 2023 – janv. 2024

Lancement du premier produit data de Scaleway : une offre de Spark managé (Spark as a Service).

- Validation des prérequis techniques : scalabilité, connectivité, résilience
- Implémentation et validation des premiers cas d'usage (POC)

Stack : Apache Spark · Spark as a Service · POC

Mobilize Financial Services (groupe Renault) — Développeur Backend

oct. 2022 – févr. 2023

Création de la néobanque Mobilize Pay, avec Accenture : développement et mise en production de l'application bancaire mobile.

RelevanC — Data Engineer

janv. 2022 – mai 2022

Refonte du data lake (optimisation et remise à niveau).

SACEM — Tech Lead

avr. 2021 – déc. 2021

Brique de réconciliation automatique des mises à jour de contrats de droits envoyées par les éditeurs musicaux.

Engie Digital — Développeur Backend

sept. 2019 – mars 2021

Plateforme smart city Livin' : qualité de l'air en temps réel, éclairage public, trafic.

BNP Paribas BDDF — Data Engineer

févr. 2019 – sept. 2019

Refonte des pipelines de traitement des dossiers clients : migration d'un mainframe DB2 vers une stack Big Data.

MISSIONS DE CONSEIL

Banque de France — Audit de code et d'architecture SI

AXA France — Modèle de données GED et cas d'usage data

Argos — Vendor due diligence

Monoprix — Audit des méthodes de déploiement

Rexel — Audit de l'organisation data

COMPÉTENCES

Data engineering Snowflake · dbt · dlt · Databricks · Apache Spark · Delta Lake · Kafka · BigQuery · Airflow · Prefect

Data quality & observabilité Elementary · Datadog · Data contracts · Alerting

IA & LLMs Agents IA · Évaluation de LLM · Datasets synthétiques · Dust · Dify · Langfuse · Retool · Claude Code

Cloud & infrastructure AWS · Azure · GCP · Terraform · GitHub Actions · GitLab CI · Docker

Langages Python · SQL · Java · Scala

Architecture & pratiques Data Mesh · Modélisation · Clean Architecture · Software Craftsmanship · Accelerate · Agile

Langues Français (bilingue) · Anglais (professionnel)

R&D & conférences Architectures frugales · Cache CPU · Plateforme Databricks

FORMATION

Master 2, Intelligence Artificielle, Systèmes, Données — Université Paris Dauphine · 2018
Diplôme d'ingénieur, Réseaux informatiques & télécommunications — INSAT, Tunis · 2017

CERTIFICATIONS

Databricks Certified Data Engineer Associate
 Databricks Certified Associate Developer for Apache Spark 3.0
 Databricks Partner Training — Solutions Architect Essentials
 AWS Certified Solutions Architect — Associate