



# Hamza Ben Marzouk

Data & AI Engineer · Paris

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

## ABOUT

Data & AI Engineer with 7+ years of experience, including 5 in consulting (OCTO Technology), across demanding industries: banking, insurance, energy, media and cloud. I design and harden data platforms (Snowflake, dbt, Databricks, Spark) and put AI into production: agents, LLM evaluation, adoption by business teams. I bridge business and engineering, upskill the teams I work with, and leave behind systems that are documented, tested and maintainable.

## WHAT I DO

### Modern data platforms

Design, take over or modernise a Snowflake or Databricks platform: ingestion, dbt modelling, orchestration, infrastructure as code and CI/CD.

### Reliability, security & governance

Data quality, observability, data contracts, access management and GDPR compliance: make the platform trustworthy and auditable.

### AI in production

AI agents, LLM evaluation, chatbots and AI adoption by business teams: move from proof of concept to measured production use.

## EXPERIENCE

### Wakam

Feb 2024 – Present

#### Data Platform Engineer

Apr 2026 – Present

Back on the Data Platform to industrialise the platform's security, governance and reliability.

- Designed a GitOps system for temporary Snowflake access: request by pull request, approval, automatic revocation and audit log, including for sensitive (GDPR) data
- Moved Snowflake configuration to infrastructure as code (Terraform): network policies, warehouses, service accounts, security tasks; addressed CIS benchmark recommendations
- Built a regulatory data mirror (UK data residency) to Azure Blob Storage, with an independent daily integrity check
- Hardened Prefect orchestration: dev/prod isolation, late-run monitoring, Slack alerting
- Migrated CI/CD from Azure DevOps to GitHub Actions, with ephemeral review environments per pull request, a SonarCloud quality gate and a Python 3.13 / uv migration
- Agentic tooling for the team: Claude Code skills automating PRs, tickets, diagnostics and routine operations

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

#### AI Engineer — KamAI team

Aug 2025 – Mar 2026

Built and deployed AI solutions with a dual goal: industrialise AI agent evaluation and maximise adoption of AI tools across the company.

- Designed and deployed an end-to-end evaluation system for business-critical AI agents
- Automated generation of synthetic evaluation datasets, validated by business experts in a Retool app
- Automated evaluation pipelines with centralised result tracking in Langfuse
- Shipped a conversational chatbot on wakam.com for policyholders (agentic workflows on Dify)
- Drove adoption of the Dust platform; ran bi-monthly hackathons (average satisfaction  $\geq 4/5$ ) and coached business teams on high-value use cases

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

#### Data Engineer — Data Platform

Feb 2024 – Jul 2025

Two strategic workstreams, PDX (Partner Data eXchange) and DPF (Data Platform Foundation), ensuring the data platform's reliability, scalability and operability.

- Technical lead on data contracts: brought 3 partner teams to autonomy in defining and applying the standards
- Delivered a hardened end-to-end pipeline (partner exchange → Snowflake), with documentation, knowledge transfer and formalised processes
- Optimised the monthly run: 50% of data support requests handled self-service
- Full observability with Datadog (alerting, dashboards); rebuilt the data quality framework and introduced Elementary

- Designed ETL / reverse ETL pipelines and restructured the dbt codebase by domain; significantly reduced data incidents

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

## OCTO Technology

Feb 2019 – Jan 2024

Consultant — client engagements

### Scaleway — Data Ops Consultant

Sept 2023 – Jan 2024

Launch of Scaleway's first data product: a managed Spark offering (Spark as a Service).

- Validated technical prerequisites: scalability, connectivity, resilience
- Implemented and validated the first use cases (proof of concept)

Stack : Apache Spark · Spark as a Service · POC

### Mobilize Financial Services (Renault Group) — Backend Developer

Oct 2022 – Feb 2023

Creation of the Mobilize Pay neobank, alongside Accenture: built and shipped the mobile banking app.

### RelevanC — Data Engineer

Jan 2022 – May 2022

Data lake overhaul (optimisation and remediation).

### SACEM — Tech Lead

Apr 2021 – Dec 2021

Automated reconciliation of music rights contract updates submitted by publishers.

### Engie Digital — Backend Developer

Sept 2019 – Mar 2021

Living smart city platform: real-time air quality, street lighting, traffic.

### BNP Paribas BDDF — Data Engineer

Feb 2019 – Sept 2019

Rebuilt client file processing pipelines: migration from a DB2 mainframe to a Big Data stack.

## ADVISORY ENGAGEMENTS

**Banque de France** — Code & IT architecture audit

**AXA France** — Document management data model & data use cases

**Argos** — Vendor due diligence

**Monoprix** — Deployment practices audit

**Rexel** — Data organisation audit

## SKILLS

|                                         |                                                                                                       |
|-----------------------------------------|-------------------------------------------------------------------------------------------------------|
| <b>Data engineering</b>                 | Snowflake · dbt · dlt · Databricks · Apache Spark · Delta Lake · Kafka · BigQuery · Airflow · Prefect |
| <b>Data quality &amp; observability</b> | Elementary · Datadog · Data contracts · Alerting                                                      |
| <b>AI &amp; LLMs</b>                    | AI agents · LLM evaluation · Synthetic datasets · Dust · Dify · Langfuse · Retool · Claude Code       |
| <b>Cloud &amp; infrastructure</b>       | AWS · Azure · GCP · Terraform · GitHub Actions · GitLab CI · Docker                                   |
| <b>Programming</b>                      | Python · SQL · Java · Scala                                                                           |
| <b>Architecture &amp; practices</b>     | Data Mesh · Data modelling · Clean Architecture · Software Craftsmanship · Accelerate · Agile         |
| <b>Languages</b>                        | French (bilingual) · English (professional)                                                           |
| <b>R&amp;D &amp; talks</b>              | Frugal architectures · CPU cache · Databricks platform                                                |

## EDUCATION

**Master's degree, Artificial Intelligence, Systems & Data** — Université Paris Dauphine · 2018

**Engineering degree, Computer networks & telecommunications** — 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