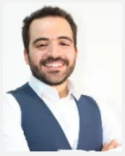
Hamza Ouhssaine



Senior Data Scientist · GenAI Engineer

BASED IN
Paris, France
Open to relocation

CONTACT
hamzaouhssaine@gmail.com
+33 7 63 73 77 89

ONLINE
github.com/hamzaca
LinkedIn

01 PROFILE

Senior data scientist and GenAI engineer with six years delivering production AI end to end (generative AI, computer vision, NLP, big data) in banking, heavy industry and travel. Today I design agentic and RAG systems for finance teams at BPCE & Natixis.

EXPERTISE **Generative AI & LLMs** / **RAG & AI Agents** / **Prompt Engineering** / **Machine Learning** / **Computer Vision** / **NLP** / **MLOps**

02 EXPERIENCE

Data Scientist, GenAI · BPCE & Natixis

Jan 2026 – Present · Paris

- **Steering ALM Metrics:** platform automating liquidity (LCR) monitoring and reporting across Groupe BPCE, built with Treasury, ALM and IT teams; an LLM-powered AI agent answers analysts' questions in plain language and drills into the data.
- **AAR (Accounting Analytical Review):** app streamlining quarterly statutory and consolidated accounting reviews for about 200 entities; drafts analyses of significant account movements and Audit Committee report sections for reviewers to edit and validate. Used by Permanent Controls and Consolidation teams to cut manual effort and standardize a process behind management decisions and audits.
- **Insights Generator:** financial analysis and BI platform for Natixis to import, pivot, chart and export reports; a generative AI assistant (LLM, RAG) extracts insights to improve business performance.
- **APAC T&E Policy Chatbot:** generative AI chatbot (LLM, RAG) answering employees' T&E policy questions with links to official sources; admins manage documents, rules and feedback from a dashboard.

SKILLS Machine Learning · Python · LLMs · Generative AI · RAG · LangGraph · LangChain · ChromaDB · Microsoft Graph API · AngularJS · JavaScript · Docker · Jenkins · OpenShift

Data Scientist · ArcelorMittal

Dec 2022 – Dec 2025 · Marseille

- **Product ID Recognition:** computer-vision pipelines reading steel product IDs at a 0.098% character error rate, reducing manual checks and improving traceability.
- **Predictive Models for Steel Production:** cooling-control models improving production reliability; computer-vision defect detection on casting rolls and products, integrated with cross-functional teams to cut manual inspection.

SKILLS Python · PyTorch · Deep Learning · Computer Vision · OpenCV · Image Processing · LLMs · RAG · Elasticsearch · Kibana · Dataiku · Decision Sciences

Big Data Engineer · Amadeus

Jan 2021 – Nov 2022 · Nice

- Built and maintained production ETL pipelines over large-scale airline datasets for Revenue Management, with automated testing and deployment.
- Delivered projects for Southwest, Polish Airlines, Singapore Airlines and Kuwait Airways; served as team leader and led hiring for the team.

Data Scientist · Meelo

Mar – Nov 2020 · Lille

OCR pipelines for administrative documents; credit-risk and fraud-detection APIs.

Deep Learning Researcher · CRIStAL Lab

Nov 2019 – May 2020 · Lille

Deep-learning steganalysis; designed and optimized a ResNet-12.

03 EDUCATION

Engineer's Degree, Data Science & AI · Centrale Lille

2016 – 2020

Master's-level engineering degree from one of France's leading Grandes Écoles, part of the Écoles Centrale group: two years of generalist, multidisciplinary engineering training, then a specialization in data science and AI taught by researchers and industry professionals.

COURSEWORK Machine learning · Big data · Statistical modeling · Data-driven decision-making · Applied mathematics · Optimization · Relational databases (SQL) · Project management · Innovation · Generalist core in physics, engineering sciences, computer science and economics · Data science integration project

SKILLS Data Science · Big Data · Deep Learning · Reinforcement Learning · LLMs · RAG · Dataiku

Master 2 (M2), Applied Mathematics · University of Lille

2019 – 2020

Master's degree in applied mathematics completed in one year in parallel with the engineering degree, through a partnership between the University of Lille and Centrale Lille.

COURSEWORK Stochastic processes: modeling random phenomena that evolve over time · Concentration of measure: high-dimensional probability, the theory behind generalization guarantees in machine learning

Bachelor's Degree, Mathematics · University of Lille

2017

Applied mathematics degree taken in parallel with the first year at Centrale Lille, through a partnership between the University of Lille and Centrale Lille.

COURSEWORK Topology · Probability · Differential equations · Differential calculus

Classes Préparatoires MPSI / MP* · Ibn Ghazi

2014 – 2016

Two-year intensive mathematics and physics program with weekly oral exams, preparing for the competitive entrance exams to France's top engineering schools. Admitted to Centrale Lille via the Centrale-Supélec exam.

COURSEWORK MPSI: analysis, algebra, finite probability, inner product spaces, linear differential equations · MP* (advanced "star" class): matrix reduction and the spectral theorem, normed spaces, power series, parameter integrals, differential calculus, discrete random variables and the law of large numbers · Also physics, chemistry, engineering sciences, computer science

(Python)

04 STACK

CODE
Python · Scala

ML & DL
PyTorch · TensorFlow · Keras · JAX · Haiku · scikit-learn · XGBoost

NLP & LLM
Hugging Face · spaCy · NLTK · LangChain · LangGraph

AI APPS
Chainlit · Streamlit

DATA
Pandas · Polars · Apache Spark · Hadoop

DATABASES
SQL / NoSQL · Impala · MongoDB · Cassandra · Chroma · Elasticsearch

PLATFORMS
Docker · Kubernetes · GCP · Dataiku

05 METHODS

Generative AI	LLM evaluation, hybrid retrieval (BM25 + vectors + RRF), reranking
Machine learning	OCR, object detection (YOLO), speech recognition (Whisper), time-series forecasting
Engineering	ETL pipelines, CI/CD, Git, Agile (SAFe, Scrum)

06 PERSONAL PROJECTS

Source code at github.com/hamzaca

01 Code Auditor

Autonomous LangGraph agent auditing and refactoring Git repos end to end: malware/secret preflight gates, static and security analysis, an LLM-generated risk-ranked roadmap, then sandboxed fixes with test-and-rollback. Treats the repo as hostile input: network-disabled containers, human approval gate. 421 backend tests.

LangGraph · FastAPI · network-disabled Docker sandboxes · ruff/mypy · bandit/semgrep/osv-scanner · ClamAV/YARA/TruffleHog · React

02 Sanad

Agentic RAG over private documents where every claim resolves to a page and bounding box. Structure-aware chunking, hybrid BM25 + vector retrieval with RRF and cross-encoder rerank; the agent decomposes multi-part questions and declines rather than guessing. Citation quotes verified against their source chunk; recall/MRR/groundedness eval harness.

LangGraph · FastAPI · hybrid BM25 + vector retrieval fused with RRF · cross-encoder rerank · vision-model OCR ingestion · pausable/resumable pipelines · Angular

03 Prompty

Prompt versioning, offline evaluation and A/B testing platform for LLM features: immutable versions with movable deploy labels (deploy and rollback are the same operation), LLM-judge grading against a golden set that blocks regressions, sticky traffic splitting, and a statistics layer: z-tests, Welch's t-test, power analysis, SRM detection, cost/latency guardrails.

FastAPI · Python SDK · LLM-judge eval harness · z-tests, Welch's t-test, power analysis, SRM detection · deterministic traffic splitting · React

04 Proving Ground

Agent evaluation benchmark scoring trajectories against world state, not the agent's own report: a deterministic event-sourced company simulator with 20 tools, 24 scenarios (a third unwinnable by acting), 4 scaffolds and 6 policy profiles with taint tracking and irreversible-action gates. Headline metric: hallucinated success, claiming "done" when the ledger disagrees.

FastAPI · WebSocket · event-sourced simulator · ReAct/Plan-Execute/Reflexion scaffolds · taint tracking · React

05 ML Engineer

Autonomous AutoML agent architecturally prevented from seeing its own holdout set: profiles a dataset, plans and codes scikit-learn experiments in a network-disabled sandbox with resource caps, runs six leakage checks it cannot overrule; holdout scored out-of-process. Reports the CV-minus-holdout gap against a naive baseline.

Python · scikit-learn · LLM code generation · network-disabled sandbox with CPU/memory/PID caps · out-of-process holdout scoring

06 Vantage / Insight Terminal

Full-stack market intelligence platform over ~194 instruments (equities, crypto, commodity futures): an ensemble of XGBoost, LightGBM, a PyTorch LSTM, ARIMA and Prophet against naive baselines with honest walk-forward backtesting, LLM-scored news sentiment, and an AI-written daily brief served from a local OpenAI-compatible model.

FastAPI · SQLModel · PyTorch LSTM · XGBoost/LightGBM · ARIMA/Prophet · walk-forward backtesting · Angular

07 Applied LLM Projects

Monorepo of five LLM systems behind one auto-discovering gateway with a unified cost tracker: a function-calling SQL agent with guardrails, a RAG eval harness (recall@k, LLM-judge groundedness, chunking ablations), a forecasting narrator with rolling-origin backtests, a drift sentinel with small-to-large model triage, and a distillation pipeline measuring an accuracy-vs-cost Pareto frontier.

FastAPI gateway · OpenAI-compatible local models & all major LLM providers · RAG evals · distillation · cost tracking · Angular

08 Tahanot ERP

Self-hosted Odoo-inspired mini-ERP (Sales, Purchase, Inventory, Dashboard) rebuilt from scratch: a generic model registry with domain filters, state-machine documents, quant-based stock with backorders, product variants, invoice payments and credit notes, plus a tool-calling AI assistant over the service layer. Bilingual EN/FR, with Postgres concurrency and locking tests.

FastAPI · SQLAlchemy · PostgreSQL · Alembic · Docker · tool-calling assistant · React

07 CERTIFICATIONS

- Build & Modernize Apps with GenAI, Google Cloud (2026)
- Advanced Machine Learning on Google Cloud
- Data Engineering, Big Data & ML on GCP
- Certified SAFe® 5 Practitioner

08 LANGUAGES

French	**English**	**Arabic**	**Tamazight**
Fluent	Fluent	Fluent	Native