OMAR EL HACHIMI

ARCHITECT / GENERATIVE AI ENGINEER

Paris, Île-de-France
oelhachi@gmail.com
French (native) · English (bilingual) · Arabic (native)

SUMMARY

- 8+ years architecting, building and deploying enterprise platforms across software development, DevOps, Security and applied AI; consultant in banking (LCL) and insurance (AXA).
- **Enterprise AI platforms:** build the layer that makes frontier models usable in regulated companies — self-service, identity, quotas and cost control, guardrails. LCL's GenAI platform IA4IT (500+ users, 90 teams); Claude Code on Amazon Bedrock at AXA (120+ users, 9 teams).
- **Evidence-grounded systems:** agentic platforms that turn documents into auditable decisions and reports under human validation (A-Pad, used by AXA project managers; KASTOR.AI); HIBOU, LCL's security command center across 700+ applications, adopted as the bank's enterprise security standard.
- **Adoption:** comparative tool testing, pilots, workshops and decision briefs for leadership (GitHub Copilot at LCL, Claude Code at AXA).

CERTIFICATIONS & EDUCATION

- AWS Certified **Generative AI Developer – Professional**
- AWS Certified **Machine Learning Engineer – Associate**
- AWS Certified **Solutions Architect – Associate**
- Microsoft **GitHub Copilot**
- **Certified Kubernetes Administrator** (CKA)
- **Certified Kubernetes Security Specialist** (CKS)
- IAPP **AI Governance Professional** (AIGP)
- **Diplôme d'Ingénieur (Master's degree), Information Systems** — Polytech Tours

PROFESSIONAL EXPERIENCE

AXA GROUP OPERATIONS (AXA GO) · Insurance — **Jan 2026 – Present**

Generative AI Engineering Lead · Consultant

Within Group Data & AI, AI Enablement initiative: help AXA's engineering teams evaluate, adopt and govern AI.

Claude Code on Amazon Bedrock – Enterprise Pilot & Cost Governance (Technical Lead)

- Designed and built the AWS layer of AXA's Claude Code pilot, with per-team access, budgets and usage dashboards.
- Built the cost-management layer: usage tracked per user, quotas per team and per user, alerts as consumption nears the limit, access blocked at the cap, chargeback to cost centres and a self-service spend view.
- Ran the pilot with users (support, power-user sessions, comparison of access models) and wrote the technical review and go/no-go decision brief for the program sponsors.
- **Impact:** 120+ users across 9 teams, each within its own budget, with spend reconciled to the AWS bill within 0.3%.

A-Pad – Configurable Agentic Evaluation & Generation Platform (Initiative & Full-Stack Build)

- Initiated A-Pad after analyzing AXA's project-gating process, and designed it as a fully configurable framework: each evaluation process or report is a template.
- Sole developer, front to back (React, FastAPI, LangGraph, PostgreSQL/pgvector): three reusable engines for evidence retrieval, agentic evaluation with self-correction and evidence-grounded confidence, and cited report generation, with human review on every result.
- Configured well beyond project gates: IT change management, climate & biodiversity reporting, registration documents, application risk and vendor due diligence.
- Built for production: fault-tolerant processing that resumes after failure, cluster-wide LLM cost control, fully auditable runs, 200+ automated tests.
- **Impact:** AXA project managers run gate assessments on it, replacing subjective reviews with evidence-cited reports.

AI Assets Marketplace – Governed Supply Chain for Agent Extensions (Architecture & Build)

- Architected and built the group marketplace for AI agent extensions (skills, agents, plugins, hooks and MCP configurations) serving both Claude Code and GitHub Copilot.
- Designed federated ownership: each AXA entity publishes its own catalogue on its own schedule, aggregated centrally.
- Built an automated security gate: every asset is scanned for secrets, unsafe code, prompt injection and data exfiltration before publication, with human review only when flagged.
- Built a census of AXA's GitHub Enterprise: 5,800+ distinct AI assets found in 40,000+ repositories, with their owners.

AI Tool Evaluation – Tech Radar & Compliance Assessment

- Built the market-intelligence view of AXA's AI Tech Radar, scoring and comparing 15 AI coding tools, with cost-at-scale and break-even analysis.
- Designed and prototyped an agent-based compliance assessment of AI coding assistants against CSA CCM v4 controls: agents gather public evidence, experts answer the rest, with human review.

LCL – CRÉDIT LYONNAIS · Groupe Crédit Agricole, Banking — **Jan 2023 – Present**

Generative AI Engineer · Consultant

Technical lead of two major initiatives: the bank's GenAI platform (IA4IT) and Security Command Center (HIBOU).

GenAI Platform – Internal "Custom GPTs" (Architecture & Build)

- Architected and industrialized the bank's central GenAI platform: self-service infrastructure through which 90 teams and 500+ users create, deploy and operate their own AI assistants.
- Developed the core capabilities: UI-configurable RAG, native connectors (Confluence, Jira, internal databases), advanced observability and debugging.
- Secured and scaled it: role-based access, quotas and cost tracking, rate limiting, load testing.
- Delivered use cases: multi-source documentation assistants, business content generation, Text-to-SQL, code analysis, K8s manifest review, CVE remediation assistance.

AI Strategy, Governance and Adoption (Lead)

- Drove GitHub Copilot adoption: comparative testing, collaborative workshops, ROI KPI definition.
- Ensured AI governance (EU AI Act compliance) and trained teams and leadership through workshops.

Security Command Center "HIBOU" – 360° Security View (Initiative & Lead)

- Proposed, designed and developed HIBOU: vulnerability management hub covering 700+ applications and 100+ teams.
- Built connectors to the bank's security scanners and ETL pipelines to consolidate and normalize scan data.
- Delivered real-time security posture: risk-based prioritization (vulnerability severity × application criticality), tracking by team and application, and team leaderboards on remediation speed.
- **Impact:** adopted as the bank's enterprise security standard; gives the bank auditable evidence of vulnerability remediation for its regulators under DORA; reference tool for CISOs, security managers and compliance.

Stack: Python, LangChain, LangGraph, AWS Bedrock (Claude), Open WebUI, Langfuse, LiteLLM, PostgreSQL, Elasticsearch, Kubernetes, Helm, GitLab CI/CD, Dynatrace, Prometheus, GitHub Copilot

KASTOR.AI — **Jan 2025 – Present**

Founder

- SaaS generating regulatory compliance reports (GDPR, EU AI Act, NIS2, ISO 27001) from client documents.
- Built solo from concept to production: RAG with hybrid search (HNSW + BM25), HyDE and re-ranking; multi-LLM orchestration (Vertex AI Gemini, Bedrock Claude) with LangGraph; FastAPI, Angular, Redis, MongoDB; GCP with Terraform.
- Product: dynamic report templates with AI or human control per section, an Angular 19 rich-text editor (TipTap) shipped as a PWA, Firebase authentication with RBAC.

EARLIER EXPERIENCE

SoLocal | DevOps Engineer · Groupe Pages Jaunes, Advertising — **2022 – 2023**

- Designed and maintained GitLab CI/CD pipelines with ArgoCD for multi-environment deployments; managed and optimized Kubernetes (right-sizing, performance tuning, Prometheus/Kibana observability).
- Built BigQuery ETL pipelines on click/impression data (100M+ daily events) and event-driven Cloud Functions.

Orange | Backend Developer · Notifications team, Telecom — **2020 – 2022**

- Backend for two critical event platforms: ZBus, an SQS-like REST messaging service (4M+ messages/month, 99.99% QoS), and Pushavoo, real-time push notifications (200M per day, 4,000 req/s).
- Drove the ZBus SaaS migration: containerization and Kubernetes deployment with ArgoCD, REST API over RabbitMQ/Kafka, Domain-Driven Design refactoring.
- Managed production releases, incident response, benchmarking and stress tests; rotating Scrum Master.

Worldline | Full Stack Developer · Worldline Private Cloud, Payments — **2018 – 2020**

- Designed the data architecture (database and data lake) with business teams for reporting modules.
- Built Talend and Spring Batch ETL jobs, REST APIs with automated PDF/Word/Excel reports, and Spring Boot/Angular dashboards, deployed on Kubernetes.

SKILLS

GenAI / LLM: LangChain, LangGraph, RAG (HyDE, Hybrid Search, Re-ranking), ReAct agents, MCP, Text-to-SQL, Prompt Engineering, Guardrails, DeepEval, Ragas

LLM Ops: AWS Bedrock (Claude), GCP Vertex AI (Gemini), Azure OpenAI, OpenAI API, Langfuse, LiteLLM, Open WebUI

Infrastructure: Kubernetes, Helm, Terraform, CloudFormation, Docker, ArgoCD, GitLab CI/CD, GitHub Actions

Full stack: Python (FastAPI), TypeScript (React, Angular), Java (Spring Boot), PostgreSQL/pgvector, Redis, RabbitMQ

Security: Vulnerability Management, Security Posture Management, SAST/DAST, SonarQube, Fortify, Container Security, HashiCorp Vault, CIS Benchmarks, Zero Trust Architecture