ALDORITH CHOUNA NGUEFA

MLOps / LLMOps Engineer

ML pipelines · Docker · CI/CD · MLflow · AWS · Prometheus · Grafana · LangSmith

Paris area (Cachan), France | +33 7 66 28 44 73 | chounanguefa005@gmail.com | portfolio.aldorith.com | LinkedIn | GitHub

French (native) · English C1 · German C1 · Available immediately · Permanent / freelance

PROFILE

Data & AI engineer (2 years at **Société Générale**) focused on productionisation: reproducible pipelines, leakage control, hardened containers, CI/CD, experiment tracking and **LLM application observability** (latency, cost, traces). MLOps training on Google Cloud.

PROFESSIONAL EXPERIENCE

Société Générale — Data & AI Scientist (apprenticeship), Audit & Inspection Générale Oct 2024 – Aug 2026

Automate the drafting of model-risk audit notes from large document sets.

- Designed and built an end-to-end **generative AI application**: ingestion of PDF, Word and Excel reports, document extraction and preparation, **Python/FastAPI** backend and web interface.
- Hybrid **BM25 + FAISS** retrieval fused with **RRF**, section-level **GPT** generation, consolidation and **source attribution** for every passage, with human review kept in the loop.
- **Outcome:** structured, sourced first draft in **minutes** for a process that took 2–3 months manually; iterated with managers; took part in strategic meetings with Data Science department leadership.

Le Comptoir d'Afrique (Afro Shop) — Freelance Full-Stack & AI Developer Sep 2026 – present

Digitising a grocery store: online shop, orders, pickup, delivery and store management.

- Gathered requirements with the owner and wrote the design file: user journeys, **SQL data model**, **OpenAPI contract**, target NestJS/PostgreSQL architecture and **RAG** assistant.
- **22 Vitest tests**, **7 Playwright E2E journeys**, axe-core accessibility checks; trained staff on the tool and cybersecurity basics. Version in acceptance testing; backend and payment to be integrated.

TECHNICAL PROJECTS

Phishing Detection — MLOps pipeline · code Oct 2026

- 11,055 rows, schema and drift validation (Kolmogorov-Smirnov), **duplicate-group splitting** (5,206 duplicates, zero train/test leakage), 5 models compared with cross-validation: **Random Forest test F1 0.959**.
- **FastAPI** and batch prediction, preprocessing + model serialised together, **Docker**, 8 pytest tests, **GitHub Actions** CI; **MLflow**/DagsHub tracking, initial **AWS ECR/EC2** deployment, S3 artefacts.

Folio — multi-agent financial research · demo 2026

- Observability: **Prometheus** metrics (p50/p95/p99, cost, 8 alert rules), **Grafana** dashboard (18 panels), masked **LangSmith** tracing, **k6** load tests; hardened Docker (read-only, cap_drop); Oracle Cloud + Cloudflare.
- **4 LangGraph agents** (planner, researcher, verifier, writer) with HMAC-signed hand-offs; market data via **MCP** Yahoo Finance behind a read-only gateway.

MultiPix — multimodal photo search (final-year project, team of two) · code 2025 – 2026

- My scope: text → image and image → image search, hybrid multimodal RAG, clustering and indexing pipeline orchestration: FastAPI API, PostgreSQL, **Redis** queue, **TinyCLIP** embedding worker (SHA-256-pinned weights), **Qdrant**; full-text + vector **RRF** fusion, optional local VLM reranking.
- **34 pytest tests**, end-to-end smoke test, GitHub Actions CI, Docker Compose; OpenCLIP / BLIP / HDBSCAN benchmarks on 5,000 COCO images.

SKILLS

MLOps: scikit-learn pipelines, schema validation, MLflow/DagsHub, model registry and hashes, batch prediction

Containers & CI/CD: Docker, Docker Compose (hardening), GitHub Actions, pytest, Ruff, k6

Observability: Prometheus, Grafana (dashboards, alerting), LangSmith, p50/p95/p99, token cost

Cloud: AWS (EC2, ECR, S3), Oracle Cloud, Cloudflare, Render

Data & AI: Python, SQL, PostgreSQL, MongoDB, Redis, Qdrant, FastAPI, RAG, LangGraph

CERTIFICATIONS & TRAINING

Google Cloud Skills Boost — Machine Learning Operations (MLOps) Fundamentals · badge earned

Google Cloud Skills Boost — Develop Gen AI Apps with Gemini and Streamlit; Agent Governance and Security

Udemy — Deep Agent: Multi-Agent RAG with Gemini and LangChain (2026)

EDUCATION

ECE Paris — Engineering degree (MSc), Data & AI (2024–2026)

EFREI Paris — Computer science engineering track (2023–2024)

REFERENCES

Pierre-Michel Becquet — apprenticeship manager (2 years), Société Générale · pierre-michel.becquet@socgen.com

Jean Djeto — Senior Auditor, Audit & Inspection Générale, Société Générale · jean.djeto@socgen.com