ALDORITH CHOUNA NGUEFA

Ingénieur MLOps / LLMOps

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

Cachan (94), France | +33 7 66 28 44 73 | chounanguefa005@gmail.com | portfolio.aldorith.com | LinkedIn | GitHub

Français natif · Anglais C1 · Allemand C1 · Disponible immédiatement · CDI / freelance

PROFIL

Ingénieur Data & IA (2 ans à **Société Générale**) orienté industrialisation : pipelines reproductibles, contrôle des fuites de données, conteneurs durcis, CI/CD, suivi d'expériences et **observabilité des applications LLM** (latences, coûts, traces). Formé aux MLOps sur Google Cloud.

EXPÉRIENCE PROFESSIONNELLE

Société Générale — Data & AI Scientist (alternance), Audit & Inspection Générale Oct. 2024 – août 2026

Automatiser la rédaction des notes de risque de modèle à partir de dossiers documentaires volumineux.

- Conçu et développé une **application d'IA générative** de bout en bout : ingestion de rapports PDF, Word et Excel, extraction et préparation documentaire, backend **Python/FastAPI** et interface web.
- Recherche hybride **BM25 + FAISS** fusionnée par **RRF**, génération **GPT** par section, consolidation et **attribution de la source** de chaque passage ; validation humaine conservée.
- **Résultat :** premier brouillon structuré et sourcé en **quelques minutes** pour un processus manuel de 2 à 3 mois ; itérations avec les managers ; participation aux réunions stratégiques du département Data Science.

Le Comptoir d'Afrique (Afro Shop) — Développeur Full-Stack & IA, freelance Sept. 2026 – présent

Digitaliser une épicerie : boutique en ligne, commandes, retrait, livraison et gestion du magasin.

- Recueilli les besoins avec la gérante et rédigé le dossier de conception : parcours, **modèle de données SQL**, **contrat OpenAPI**, architecture cible NestJS/PostgreSQL et assistant **RAG**.
- **22 tests** Vitest, **7 parcours E2E** Playwright, contrôle d'accessibilité axe-core ; formation de l'équipe à l'outil et aux bons réflexes de cybersécurité. Version en validation, backend et paiement à intégrer.

PROJETS TECHNIQUES

Phishing Detection — pipeline MLOps · code Oct. 2026

- 11 055 lignes, validation de schéma et de dérive (Kolmogorov-Smirnov), **découpage par groupes de doublons** (5 206 doublons, 0 fuite train/test), 5 modèles comparés en validation croisée : **Random Forest F1 test 0,959**.
- API **FastAPI** et prédiction par lots, prétraitement + modèle sérialisés ensemble, **Docker**, 8 tests pytest, CI **GitHub Actions** ; suivi **MLflow**/DagsHub, déploiement initial **AWS ECR/EC2**, artefacts S3.

Folio — recherche financière multi-agent · démo 2026

- Observabilité : métriques **Prometheus** (p50/p95/p99, coûts, 8 règles d'alerte), tableau de bord **Grafana** (18 panneaux), traçage **LangSmith** masqué, charge **k6** ; Docker durci (lecture seule, cap_drop) ; Oracle Cloud + Cloudflare.
- **4 agents LangGraph** (planificateur, chercheur, vérificateur, rédacteur) avec transferts signés HMAC ; données de marché via **MCP** Yahoo Finance derrière une passerelle en lecture seule.

MultiPix — recherche photo multimodale (projet de fin d'études, binôme) · code 2025 – 2026

- Mon périmètre : recherche texte → image et image → image, RAG multimodal hybride, clustering et orchestration des pipelines d'indexation : API FastAPI, PostgreSQL, file **Redis**, worker d'embeddings **TinyCLIP** (poids figés SHA-256), **Qdrant** ; fusion plein texte + vecteurs par **RRF**, reranking optionnel par VLM local.
- **34 tests** pytest, smoke test de bout en bout, CI GitHub Actions, Docker Compose ; benchmarks OpenCLIP / BLIP / HDBSCAN sur 5 000 images COCO.

COMPÉTENCES

MLOps : pipelines scikit-learn, validation de schéma, MLflow/DagsHub, registre et empreintes de modèles, prédiction par lots

Conteneurs & CI/CD : Docker, Docker Compose (durcissement), GitHub Actions, pytest, Ruff, k6

Observabilité : Prometheus, Grafana (tableaux de bord, alertes), LangSmith, p50/p95/p99, coûts tokens

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

Données & IA : Python, SQL, PostgreSQL, MongoDB, Redis, Qdrant, FastAPI, RAG, LangGraph

CERTIFICATIONS & FORMATIONS

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

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

Udemy — Deep Agent : Multi-Agent RAG avec Gemini et LangChain (2026)

FORMATION

ECE Paris — Diplôme d'ingénieur Data & IA (2024–2026)

EFREI Paris — Cycle ingénieur informatique (2023–2024)

RÉFÉRENCES

Pierre-Michel Becquet — manager d'alternance (2 ans), Société Générale · pierre-michel.becquet@socgen.com

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