

MARIUS AYRAULT

AI Engineer · Systèmes LLM de production

Français (langue maternelle)
Anglais (professionnel, TOEIC 895)
Full remote · France & international

✉ contact@mariusayrault.com mariusayrault.com linkedin.com/in/marius-ayrault github.com/mriusero

AI Engineer spécialisé dans les systèmes LLM de production : IA agentique, RAG, évaluation et observabilité, LLM Ops. Ancien ingénieur industriel dans l'aéronautique (Groupe Safran), devenu founding engineer d'une startup IA en pre-seed. J'apporte une résolution structurée des problèmes, une vision système et une culture du résultat mesurable à l'architecture IA.

EXPÉRIENCE

Founding AI Engineer & Architect

Août 2025 - Avril 2026

Arkim AI, startup en pre-seed · Mission freelance, remote (Los Angeles)

- Pipeline d'ingestion multimodal en production sur Databricks : OCR, image-to-text, chunking sémantique et embeddings à grande échelle.
- Système multi-agents à deux niveaux avec LangGraph : orchestration Plan-and-Execute de sous-agents ReAct, points de contrôle human-in-the-loop, streaming SSE via FastAPI et prompt caching sur AWS Bedrock.
- Framework d'évaluation RAG (Ragas + LangSmith) sur 5 pipelines candidats : le pipeline interne surpasse la baseline commerciale en Precision@k (0,82 contre 0,68) et en Recall@k (0,70 contre 0,56).
- Aux côtés des founders, de la vision produit à la production : observabilité et boucles de feedback en place pour une itération pilotée par les données.

Ingénierie IA indépendante

Octobre 2023 - Août 2025

Indépendant · Paris, France

- Transition délibérée de l'ingénierie industrielle vers l'ingénierie IA, menée comme un programme de R&D, avec une formation Data Analytics à la Sorbonne.
- Des systèmes livrés de l'exploration à la production : scoring crédit industrialisé (MLOps), pipeline de streaming temps réel sur Kafka, deep learning appliqué (vision, NLP, séries temporelles), solutions RAG et déploiement de modèles open source. Chaque projet est documenté sur mariusayrault.com.

Production Lead, documentation de maintenance

Octobre 2022 - Septembre 2023

Docaret, pour Dassault Aviation · Saint-Cloud, France

- Pilotage en binôme de la production externalisée de la documentation de maintenance du Rafale, sur trois équipes en France et à l'international : standards de rédaction technique, conformité XML et contrôle de versions.
- Workflows restructurés : backlog absorbé, puis capacité de production des sous-traitants augmentée jusqu'à 50 % ; outils de BI conçus pour un suivi de la production en temps réel.

Ingénieur méthodes & industrialisation

Septembre 2017 - Septembre 2022

Safran Electrical & Power · Apprentissage · Niort, France & Pitstone, Royaume-Uni

- Projet Lean Six Sigma Green Belt sur une cellule automatisée : TRS de 40 % à 87 %, rebuts de 8 % à 2 %, capacité Cp/Cpk de moins de 1 à plus de 2, 19 419 € d'économies annuelles.
- Industrialisation de l'assemblage des moteurs électriques ENGINEUS : conception d'une ligne semi-automatique dé-risquée par POC, PFMEA process animé à Pitstone (176 risques cotés avec 8 métiers), relocalisation post-Covid de l'assemblage de la power unit en France.
- Études R&D : busbars allégés jusqu'à 30 %, validés en essais thermiques ; automatisation par cobotique dimensionnée à 79 530 € de gains annuels, retour sur investissement en 1,94 an.

STACK TECHNIQUE

Agentic AI & LLM

LangGraph, LangChain, MCP, AWS Bedrock, vLLM, streaming SSE, prompt caching, structured outputs

RAG & document intelligence

OCR & VLM, chunking sémantique, embeddings, recherche hybride, reranking, ChromaDB, knowledge graphs

Évaluation & observabilité

Ragas, LangSmith, LLM-as-Judge, golden datasets, évaluation de trajectoires d'agents, suivi des coûts

ML & deep learning

PyTorch, TensorFlow, Scikit-learn, XGBoost, MLflow, Optuna

Backend & infra

Python, FastAPI, Docker, CI/CD, Kafka, Spark, Databricks, AWS, PostgreSQL

CERTIFICATIONS

Claude with AWS Bedrock

Anthropic · 2026

Agents Course

Hugging Face · 2025

Deep Reinforcement Learning

Hugging Face · 2025

Green Belt Lean Six Sigma

CESI Certifications · 2022

FORMATION

Data Analytics, Diplôme Universitaire

Université Paris 1 Panthéon-Sorbonne · 2024 - 2025

Diplôme d'Ingénieur généraliste

CESI École d'Ingénieurs · 2019 - 2022

DUT Génie Mécanique et Productique

IUT Poitiers · 2017 - 2019