Samir MAHAMAT SALEH

Applied Gen AI Engineer

Applications GenAI de bout en bout RAG, agents et workflows LLM multi-fournisseurs Python, React, APIs

Île-de-France · Permis B | mahamatsaleh.samiradam@gmail.com
github.com/MoSaleh47 | linkedin.com/in/samir-adam-mahamat-saleh

PROFIL

Diplômé Bac+5 en Intelligence Artificielle et Science des données (aivancity School of AI & Data For Business & Society, Programme Grande École), **Microsoft Certified: Azure AI Apps and Agents Developer Associate**, auteur d'une prépublication sur l'évaluation de LLM. Deux ans d'alternance à la DGAC/DSAC à **concevoir, architecturer et déployer de bout en bout des applications d'IA générative** réellement utilisées par des équipes métier : backend Python et APIs REST, interfaces web, intégration de LLM de plusieurs fournisseurs, workflows et agents, recherche vectorielle et pipelines CI/CD. **Disponible immédiatement**.

COMPÉTENCES TECHNIQUES

IA générative, LLM et agents: architectures RAG dense et hybride, embeddings, stratégies de chunking (dont parent-enfant), bases et index vectoriels (FAISS, Azure AI Search)
Fournisseurs de LLM et frameworks: Anthropic, OpenAI, Google Gemini, Mistral, Azure OpenAI, Azure AI Foundry, Groq, Hugging Face, Ollama, LangChain
Backend et frontend: Python, FastAPI, Flask, APIs REST, TypeScript, React 18 avec Vite, Tailwind CSS, React Flow, Recharts, Streamlit, Gradio
Cloud, DevOps et MLOps: Azure, Docker, Git / GitHub, CI/CD (GitHub Actions, Azure DevOps), pratiques MLOps et LLMOps, tests et logging
Évaluation et fiabilité: RAGAS, tracing Langfuse, jeux de questions de référence, taux de réponses non fondées, fidélité des citations, calibration
Données et ML: SQL, Pandas, NumPy, DuckDB, parsing XML / PDF / JSON, data quality, scikit-learn, TensorFlow, PyTorch

CERTIFICATION ET PUBLICATION

Microsoft Certified: Azure AI Apps and Agents Developer Associate - Microsoft, juillet 2026. Credential ID 3F9FED23D318BD9C. Applications et agents IA sur Azure AI Foundry et Azure OpenAI, RAG, orchestration, évaluation.
Prépublication - *VCBench* (SSRN, 2026), évaluation de LLM ouverts. ssrn.com/abstract=6832218

EXPÉRIENCE PROFESSIONNELLE

Chargé d'Études en Data Science - DGAC / DSAC Oct. 2024 - Sept. 2026
Direction Générale de l'Aviation Civile, Direction de la Sécurité de l'Aviation Civile *Paris, France*

- Conception d'une solution IA de traitement des déclarations d'événements de sécurité aérienne, transformant des comptes rendus en texte libre en **48 champs structurés au format JSON**, avec nettoyage, validation d'intégrité et contrôles de cohérence automatisés.
- Automatisation du pré-remplissage de formulaires réglementaires PDF à partir des données extraites, ramenant le temps de traitement d'une déclaration de **plus de 30 min à environ 5 min**, soit plus de 80 % de temps gagné sur la saisie et la recherche d'informations.
- Développement d'une chaîne applicative complète en Python comprenant extraction LLM, nettoyage et validation JSON, génération de formulaires, backend FastAPI avec plusieurs endpoints REST et interface web métier.
- Évaluation de la solution sur **une centaine d'occurrences** de sécurité aérienne, avec vérification des sorties structurées, des champs réglementaires et de la cohérence des données avant utilisation.
- Mise à disposition et tests auprès de 5 agents métier, avec recueil des retours utilisateurs, corrections, documentation technique et fonctionnelle et amélioration itérative de l'outil.

Stagiaire Développeur Python - UBA Tchad Juin 2023 - Août 2023
United Bank for Africa - traitement de données bancaires multi-agences *Tchad*

- Développement d'une application PyQt5 permettant aux équipes métier d'importer, nettoyer et consolider les fichiers Excel provenant de plusieurs agences bancaires.
- Automatisation du nettoyage des données avec Python et Pandas, puis génération de visualisations Matplotlib à l'appui de l'analyse et du reporting.

PROJETS TECHNIQUES ET RECHERCHE

AI Act Evidence Desk - agent RAG de conformité | *RAG hybride, agents, MCP, RAGAS, Langfuse* 2026

- Projet open source transformant un cas d'usage IA en cartographie de preuves sourcée au regard du règlement européen sur l'IA.

PFE - Benchmark de question-réponse en réglementation aéronautique | *Ingestion XML/PDF, FAISS, RAG, LLM* 2025 - 2026

- Projet de fin d'études : chaîne complète d'ingestion du règlement UE 376/2014 et des règles EASA, avec parsing XML et PDF, nettoyage et découpage en 133 chunks, puis construction d'un index vectoriel FAISS.

EuroMotion Model Factory - application web de simulation | *React 18, TypeScript, Vite, React Flow, Recharts* 2026

- Application web de dynamique des systèmes pour la planification stratégique d'une chaîne d'approvisionnement : moteur de simulation piloté par configuration JSON, diagrammes interactifs et restitutions graphiques. github.com/MoSaleh47/EuroMotion-Model-Factory

FORMATION

Programme Grande École - Diplôme d'études supérieures en Intelligence Artificielle et Science des données 2021 - 2026
aivancity School of AI & Data For Business & Society — Titre RNCP38584 Niveau 7 (Bac+5). Cursus hybride : intelligence artificielle, science des données, culture business, éthique de l'IA, RGPD et AI Act, gestion de projet. Projets en situation réelle au sein de la Clinique de l'IA *Paris, France*

LANGUES ET QUALITÉS

Langues: Français courant et professionnel. Anglais C1, cursus de master suivi en anglais.
Qualités: rigueur méthodologique, autonomie, curiosité technique, vulgarisation, documentation systématique, travail en équipe.