

ABOUBAKAR SALY

Élève ingénieur IA | NLP, RAG, agents LLM et Graph ML

Stage de fin d'études de 6 mois | disponible dès février 2027

Lyon, France | +33 7 62 57 51 69 | aboubakarsaly@aol.com

LinkedIn : aboubakar-saly | GitHub : abou-dev | Mobilité France et international

EXPERTISE TECHNIQUE

IA générative : RAG (Retrieval-Augmented Generation), agents LLM, embeddings, ChromaDB, évaluation RAGAS, OpenAI API, Claude Agent SDK, Model Context Protocol (MCP).

Machine learning : NLP, apprentissage sur graphes (Graph ML), PyTorch, Hugging Face Transformers, scikit-learn, PyKEEN, SPARQL.

Ingénierie : Python, SQL, FastAPI, Docker, CI/CD, Git, Linux.

EXPÉRIENCE PROFESSIONNELLE

Merck Group | AI Systems Architect Intern

Enterprise Architecture, Merck Life Science | Molsheim, hybride

sept. 2025 - janv. 2026

- Conçu et livré un assistant RAG multi-source (code, tickets, documentation) avec citations, puis étendu la solution à la revue automatisée de pull requests.
- Construit un protocole d'évaluation associant jeu de questions de référence, métriques RAGAS et retours utilisateurs. Comparé plusieurs stratégies de découpage documentaire.
- Prototypé des agents LLM avec Claude Agent SDK et un serveur MCP pour Bitbucket. Conteneurisé avec Docker et automatisé via n8n.

Mac Consulting Sénégal | Développeur web freelance

Cabinets comptables en Afrique de l'Ouest | Dakar, puis à distance

mai 2024 - août 2025

- Développé, hébergé sur OVH et sécurisé des sites Next.js. Assuré la maintenance et la formation des utilisateurs.

PROJETS IA SÉLECTIONNÉS

Détection multi-label de contenus toxiques - Jigsaw

Python | scikit-learn | PyTorch | Hugging Face Transformers

GitHub

- Comparé cinq approches (TF-IDF, SVM, BiLSTM, DistilBERT) sur six labels. Meilleur F1 micro : **0,821**. Optimisé les seuils par label et analysé les erreurs sur les classes rares.

Graphe de connaissances et collaboration scientifique - DBLP

RDF/OWL | SPARQL | NetworkX | Leiden | PyKEEN

contributeur principal

- Modélisé 786 auteurs et 2 659 co-publications. Évalué les communautés (**NMI 0,71**) et la prédiction rétrospective de liens (**ROC AUC 0,79**).

FORMATION

EPITA | Diplôme d'ingénieur en informatique, majeure SCIA-Graphes

Lyon | NLP, Graph ML, Knowledge Graphs, séries temporelles, Big Data, IA explicable

2024 - 2027

École Supérieure Polytechnique, UCAD | M1 Informatique, IA et Big Data

Dakar | projets Big Data avec Apache Spark et Scala

2023 - 2024

ESMT | Licence Télécommunications et Informatique

Dakar | 2020 - 2023

Langues : français courant, anglais professionnel (TOEIC 865/990). **Certifications** : CS50 AI (Harvard), Azure AI-900, GitHub Foundations.