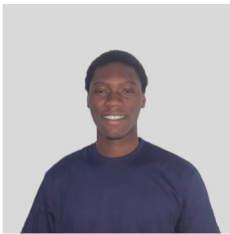


SANA YAVO

Lead Ingénieur IA, Data & Logiciel



RENSEIGNEMENTS PERSONNELS

Abidjan Abidjan
contact@sanayavo.com, +225 07 07 98 81 28
LinkedIn: linkedin.com/in/sanayasfp
Portfolio: sanayavo.com/about

PROFILE

Ingénieur IA & Logiciel avec plus de 4 ans d'expérience dans la conception d'architectures distribuées et l'industrialisation de modèles de Machine Learning (MLOps). Spécialiste en System Design, je conçois des infrastructures backend robustes (Rust, Python, TypeScript) et déploie des systèmes d'IA **auto-hébergés et souverains** alliant recherche sémantique (RAG), bases vectorielles (pgvector) et modèles de Computer Vision. Formateur et mentor dans l'âme, je possède une forte capacité à vulgariser des concepts techniques complexes pour accompagner la montée en compétences des équipes et aligner l'ingénierie avec les objectifs stratégiques.

EXPERIENCE PROFESSIONNELLE

Lead Ingénieur IA & Architecte Logiciel (Senior) de juin 2024 à Ce jour
Edane SA, Abidjan

- Direction Technique (Lead de 4 développeurs)** : Architecture et développement d'Edane AI, une plateforme multi-domaines d'agents IA conversationnels financiers interagissant via le standard Model Context Protocol (MCP).
- Gouvernance & Sécurité des données (Data Center / Banque)** : Conception d'un système d'anonymisation réversible des données PII (NER via Presidio/spaCy, substitution déterministe) et d'un Vault cryptographique (HMAC-SHA256/Fernet) garantissant le secret bancaire vis-à-vis des LLMs publics.
- Orchestration Multi-Agents & MLOps** : Implémentation d'un framework basé sur LangGraph avec gestion d'état distribuée, sous-agents spécialisés (analyse financière, scoring), et système de mémoire à long terme (MongoDB).
- Optimisation des coûts & Latence** : Création d'une architecture Fast-Path / Slow-Path utilisant des petits modèles locaux pour le routage des outils métier, réduisant drastiquement les coûts d'inférence et la latence.
- Auditabilité & Fiabilité** : Mise en place d'un système de journalisation inaltérable des requêtes IA et d'un LeakScanner asynchrone pour garantir la conformité aux audits bancaires.

Ingénieur Systèmes Cloud & MLOps (Distance) de avr. 2023 à mai 2024
Wearer Lab, Géorgie, États-Unis

- Architecture et déploiement d'une infrastructure Cloud pour l'ingestion et l'analyse temps réel de signaux audio médicaux.
- Optimisation matérielle et logicielle du pipeline de collecte de données pour l'inférence de modèles d'IA distants, assurant une latence minimale et une synchronisation réseau fluide.
- Mise en place de protocoles stricts de fiabilité et d'intégration d'API d'IA dans un environnement de santé exigeant.

Architecte IA & Logiciel (Temps partiel)

de avr. 2023 à janv. 2025

Chill Inc, Abidjan, Côte d'Ivoire

- Architecture de bout en bout et développement d'ELSA, un assistant conversationnel propulsé par des modèles d'IA générative.
- Pilotage du cycle de vie logiciel (CI/CD, Cloud) depuis la conception jusqu'à la mise en production pour absorber une demande utilisateur en forte croissance.
- Implémentation d'infrastructures scalables et gestion des cycles d'itérations rapides en collaboration avec les équipes produit.

Ingénieur Back end

de oct. 2022 à juil. 2023

Cool Lion Finance, Abidjan, Côte d'Ivoire

- Conception d'architectures backend distribuées focalisées sur la haute disponibilité et la robustesse pour soutenir la forte croissance de la startup.
- Optimisation de l'allocation des ressources de l'application et sécurisation des intégrations d'API.

Développeur Web & Data

de févr. 2022 à juil. 2022

Orange Digital Academy, Abidjan, Côte d'Ivoire

- Développement d'un portail de données ouvertes (Open Data) et d'une plateforme d'assistance B2B au sein d'une équipe pluridisciplinaire.
- Présentation et défense technique du projet lors de l'événement Orange Digital Meet'n Pitch (démonstration de capacités de communication et de vulgarisation).

COMPÉTENCES TECHNIQUES

- **Intelligence Artificielle & Data** : Entraînement de modèles (Computer Vision, PyTorch, YOLO, TensorFlow, scikit-learn, NLTK), RAG Hybride (Reciprocal Rank Fusion, pgvector), LLMs, Traitement du langage naturel (NLP), Pipelines MLOps, Data Augmentation.
- **Ingénierie Logicielle & Cloud** : Architecture de systèmes distribués, System Design, Développement d'outillages CLI (Rust), API REST, Haute disponibilité, Scalabilité.
- **Bases de données** : PostgreSQL avancé (PostGIS, pg_cron, pgmq, pgvector), Supabase, MySQL, Prisma, Redis.
- **Langages & Outils** : Rust, Python, TypeScript, PHP, SQL, Git.
- **Pédagogie & Leadership** : Mentorat de développeurs juniors, Vulgarisation scientifique, Rédaction de documentation technique, Direction technique de projets.
- **Langues** : Français (Natif), Anglais technique.

PROJETS OPEN SOURCE ET RECHERCHE APPLIQUÉE

- **Car Inspect AI** : Entraînement de A à Z d'un modèle de Computer Vision (YOLO11n sur dataset personnalisé) pour la détection automatisée de pièces automobiles. Développement d'un pipeline MLOps et d'un mini-ORM en Python.
- **Stack (Tooling en Rust)** : Conception d'un gestionnaire d'environnement de développement natif multi-projets (alternative à Docker sans virtualisation, interagissant avec vfox et Caddy).
- **Laplace Nearby** : Développement d'un moteur de recherche hybride (PostgreSQL, pgvector, PostGIS) pour IA conversationnelle, combinant recherche full-text, similarité vectorielle et proximité spatiale.

FORMATION

Executive Master of Engineering in Artificial Intelligence

de sept. 2025 à déc. 2026

ISM Paris/ISAE-Supmeca

Licence en électronique, énergie électrique et automatique

de janv. 2021 à déc. 2023

Université Polytechnique de Bingerville