

**Amadou MAMANE**  
**Lead AI & ML Engineer**  
**CentraleSupélec Alumni**

AI Engineer spécialisé dans la conception et la mise en production de systèmes IA générative, RAG et multi-agents, avec une forte expertise en architecture backend, pipelines de données, observabilité, évaluation et déploiement sur GCP ou clusters privés

✉ [mamanesarki@yahoo.fr](mailto:mamanesarki@yahoo.fr)

📁 <https://amadoumamane.fr>

🌐 <https://github.com/AmadouMamane>

☎ +336.52.78.18.48

🏠 Vincennes, France

🌐 Français

## COMPÉTENCES

---

### ✓ Techniques

- Frameworks ML **Scikit-learn, Spark ML, TensorFlow**, Pytorch, **Vertex AI**, Statsmodels
- Visualisation de données **Matplotlib, Seaborn**, Plotly, Bokeh, **Tableau**, Grafana
- Langages de programmation **Python**, Numpy, **Pandas**, Scipy, Scala, Java, Shell, **SQL**
- Bases de données **BigQuery, AlloyDB**, Hive, **Pinecone**, Cassandra, Oracle, **Firestore**
- MLOps Git, **GitLab CI/CD, Kubeflow, Terraform**, Airflow, **Docker, Kubernetes**
- Plateformes big data **Hadoop, Spark**, Flink, **Kafka, Nifi**, Streamsets
- Cloud computing **Google Cloud Platform, AWS**, Openstack
- IA Générative GPT, Llama, Gemini, Claude, **ADK**, LangChain, **LangGraph**, LangSmith

### ✓ Fonctionnelles

- **Statistiques descriptives et inférentielles** (estimation, test d'hypothèse), **Séries temporelles**
- **NLP**, Théorie de graphes et analyse de réseaux, **Transformers, RAG, Prompt engineering**
- **Vision par ordinateur, Deep learning, Systèmes de recommandation, IA générative**
- Apprentissage non supervisé, Segmentation, **Méthodologie agile**
- Data engineering, MLOps, **Architecture Data/IA**

### ✓ Certifications

- C# developer, Septembre 2014
- Applied Data Science Program: Leveraging AI for Effective Decision Making - MIT, Juillet 2024
- Google Cloud Certified - Generative AI Leader, Septembre 2025
- Google Cloud Certified - Professional Machine Learning Engineer, Octobre 2025
- Google Cloud Certified - Professional Cloud Architect, Avril 2026

### ✓ Langues

- **Français** Courant
- **Anglais** Professionnel

## FORMATION

---

- 2024 - 2024** **Applied Data Science Program**: Leveraging AI for Effective Decision Making, **MIT**
- 2014 - 2015** **M.S. Ingénierie des Systèmes Informatiques Ouverts**, Ecole Centrale Paris
- 2011 - 2012** Master recherche Génie Industriel et Informatique, **Ecole des Mines de Saint-Etienne**

## EXPÉRIENCE PROFESSIONNELLE

---

Juin 2025 – Aujourd'hui

Devoteam GCloud/Thalia, Allemagne

### AI Engineer

Au sein d'une équipe de 8 personnes

- ✓ **Conception et industrialisation** d'une plateforme **IA multi-agents** dédiée au **support client** et à la **recommandation de produits** dans un contexte retail omnicanal
  - Définition et implémentation de l'**architecture data/IA** sur **GCP**, optimisée pour la **scalabilité**, la **performance** et les **coûts**
  - Développement de **composants IA réutilisables** et d'un pipeline de **vectorisation multimodale** texte + images avec Vertex AI Embeddings, BigQuery ML et AlloyDB
  - **Industrialisation et automatisation** des traitements via GitLab CI/CD et Terraform
  - Mise en place des briques **d'observabilité**, supervision et traçabilité des traitements IA avec OpenTelemetry, Cloud Trace, Cloud Logging et Cloud Monitoring
- ✓ Contribution aux **choix d'architecture** et aux standards d'ingénierie pour garantir la **maintenabilité**, la **réutilisabilité** et la qualité de **mise en production**
- ✓ **Accompagnement technique** de l'équipe sur les sujets IA générative, architectures multi-agents, pipelines de retrieval et **bonnes pratiques d'industrialisation**

#### Environment technique

LangGraph, Vertex AI Embeddings, Cloud Run Jobs, AlloyDB, BigQuery ML, GitLab CI/CD, Terraform, OpenTelemetry, Cloud Trace, Cloud Logging, Cloud Monitoring

#### Environment fonctionnel

Librairie, retail omnicanal (online + 950 magasins)

---

Juillet 2024 – Mai 2025

MIT/Projet de recherche appliquée, À distance

### AI Engineer

Au sein d'une équipe internationale de 4 personnes

- ✓ **Conception et développement** d'un système intelligent de **Retrieval Augmented Generation** (RAG) pour l'interrogation en langage naturel de **sources structurées** et **non structurées** (PDF, CSV, HTML, bases SQL)
  - Conception d'un **pipeline automatisé** d'ingestion multi-formats et de préparation des données
    - **Nettoyage documentaire** et stratégie de **chunking robuste** avec fallback par taille de tokens
    - **Enrichissement** des données par **métadonnées exploitables** pour la recherche hybride
  - Mise en place d'un **pipeline de vectorisation** et de recherche sémantique
    - **Embedding** des documents via modèles spécialisés (Hugging Face)
    - **Indexation** dans Qdrant avec support de la **recherche hybride** (similarité vectorielle + recherche lexicale)
    - Implémentation de stratégies de **reranking contextuel** pour améliorer la **pertinence** des documents récupérés
  - **Génération** de réponses fiables et explicables à partir de **LLM open-source**
    - **Génération conditionnée** via LLaMA et Gemma
    - Mise en place d'un **scoring de confiance** pour la **réduction des hallucinations**
    - **Citation** automatique des **sources** et formatage structuré des réponses
    - **+30 % de précision** perçue lors des **évaluations qualitatives**

#### Environment technique

LangChain, LangGraph, Qdrant, FAISS, Nltk, LLaMA, Hugging Face Transformers

#### Environment fonctionnel

Recherche appliquée

**Data scientist**

*Au sein d'une équipe de 20 personnes*

- ✓ Au sein du département de **lutte contre la fraude**, participation au développement d'une solution visant à **sécuriser les transactions bancaires** par la **détection proactive d'attaques** ciblant les e-commerçants
  - **EDA** approfondie pour la détection de **signaux faibles**
  - **Feature engineering** sur des **données transactionnelles massives**
  - Implémentation d'un modèle de **Gradient Boosting** (Spark ML) optimisé pour signaux rares
  - **Industrialisation** complète via PySpark & Oozie
- ✓ **Recommandations** de bonnes pratiques de développement **Apache Spark**

**Environment technique**

**PySpark, Spark ML, Python, Hive, Impala, Oozie, Git, Séries temporelles, Data leakage**

**Environment fonctionnel**

**Banque, Lutte contre la fraude**

**Data scientist**

*Au sein d'une équipe de 30 personnes*

- ✓ Participation à la création et à l'évolution de la **plateforme centrale de gestion des données clients** du groupe, ainsi qu'au **développement d'applications data et ML stratégiques**, dans un contexte exigeant en **qualité, confidentialité et traçabilité des données**. Projet accompagné d'une migration des traitements d'un cluster Hadoop on-premise vers Google Cloud Platform (GCP)
  - **Conception** et implémentation d'un **score de prédiction** de l'attrition des clients
  - Développement d'un **moteur de déduplication des données** clients basé sur la **similarité vectorielle et les graphes**
  - **Modèles ML** : scoring churn, segmentation, détection de fraude
  - **Industrialisation des traitements data/ML** avec des pipelines CI/CD, Airflow et Kubeflow
  - Pilotage de plusieurs **POC** pour évaluer et intégrer de nouvelles technologies

**Environment technique**

**Scikit-learn, Spark ML, Scala, Python, GitLab CI/CD, GraphFrames, GCP (Vertex AI, BigQuery, Pub/Sub, Cloud Run, Dataproc, Composer), Nifi, Jira**

**Environment fonctionnel**

**Banque, Risque et crédit, Connaissance client, Méthodologie agile**

**ML engineer**

*Au sein d'une équipe de 20 personnes :*

- ✓ Contribution à un projet de **valorisation en temps réel** des **données d'interaction client** sur le réseau mobile (2G, 3G, 4G), avec pour objectif d'**alimenter les applications métiers** et les **services transverses** du groupe
  - Développement d'applications de traitement en temps réel des données réseau, normalisation et diffusion dans un format pivot à l'ensemble des équipes métier
  - Participation au cycle complet de développement logiciel : prototypage, tests, sélection et intégration de composants applicatifs
  - Suivi de l'activité réseau à l'aide de KPIs métiers

- ✓ Mise en production, supervision et maintenance de plusieurs applications critiques

Environnement technique

Kafka, Flink, Spark, Hadoop, InfluxDB, Grafana, Streamsets, Scala, Java, Git, shell

Environnement fonctionnel

Télécommunications, Connaissance client

September 2016 – April 2017

Altran/SNCF, La Défense

### Data scientist

Au sein d'une équipe de 20 personnes :

- ✓ Appui à la stratégie de **yield management** (optimisation de la politique tarifaire) à travers le développement d'**outils d'aide à la décision tarifaire**: prédiction de la sensibilité au prix et prévision des ventes de billets
  - **Collecte des données** web de recherches sur les trains
  - Développement d'un **modèle** de prédiction de remplissage de la première classe
  - Maintenance d'un modèle de prédiction de la sensibilité des trains au prix
  - Collecte des données web de recherches sur les trains
- ✓ Contribution à l'**amélioration de la qualité de code et mise en place de bonnes pratiques de développement**

Environnement technique

Scikit-learn, PySpark, Python, Pandas, scala, Spark, Git, Bokeh, Jenkins, Luigi

Environnement fonctionnel

Yield management, Méthodologie agile

Novembre 2015 – Septembre 2016

Altran/Orange, Guyancourt

### Data scientist

Au sein d'une équipe de 30 personnes :

- ✓ Participation à l'implémentation de la solution **Customer Experience Management On Demand (CEMOD)** de Nokia, destinée à la **connaissance client** et à l'**amélioration de la performance du réseau mobile**
  - Préparation et transformation des données issues du réseau, de la facturation et du CRM
  - Intégration de ces données dans la plateforme CEMOD
  - Calcul et prédiction de scores permettant d'évaluer la satisfaction client et la qualité du réseau dans les zones de vie du client
  - Contribution à l'anonymisation des données sensibles conformément aux exigences de confidentialité et conformité

Environnement technique

Hive, Pig, Java, Oozie, shell, svn

Environnement fonctionnel

Télécommunications, Connaissance client

Avril 2015 – Septembre 2015

Société Générale, Val de Fontenay

### Datascientist/Stage

Au sein d'une équipe de 20 personnes :

- ✓ **Projet de cybersécurité** innovant visant à détecter les **attaques par déni de service distribué**

**(DDoS)** à partir des **logs applicatifs**, en rupture avec les approches traditionnelles qui se concentrent sur les indicateurs réseau (volumétrie, paquets, connexions)

- **Recherche sur l'état de l'art des méthodes de détection DDoS** basées sur les logs web
- Nettoyage des données brutes et reconstruction des sessions de navigation
- **Feature engineering et visualisation** pour mettre en évidence des **patterns anormaux**
- **Clustering** pour identifier les **groupes de sessions de navigation**
- **Labélisation** semi-automatique des données
- Développement, entraînement et optimisation de **modèles de classification supervisée** pour la détection d'attaques

**Environnement technique**

**Hadoop, Scikit-learn, Pig, Hive, Python, Java, Pandas, Numpy, Scipy, Matplotlib, Git**

**Environnement fonctionnel**

**Banque, Méthodologie agile**