



# MAIA BAZERJI

Ingénierie IA & Data Science | Architectures RAG & Agents

Recherche d'alternance IA / Data — disponible immédiatement

Ingénieure IA spécialisée dans les architectures RAG, avec une orientation **R&D appliquée et industrialisation** (fine-tuning LLM, computer vision). Conception de pipelines ML en production (réduction de 80% du temps de traitement catalogue chez Kabeen).

## CONTACT

- 📞 06 12 39 74 50
- ✉ [bazerjimaia@gmail.com](mailto:bazerjimaia@gmail.com)
- ✉ [maia.bazerji@epitech.eu](mailto:maia.bazerji@epitech.eu)
- 📍 Meudon, France
- 🌐 [github.com/maiabazerji](https://github.com/maiabazerji)
- 🌐 [maia-bazerji](https://www.linkedin.com/in/maia-bazerji)
- 📖 [Medium Articles](#)

## FORMATION

### Master – Technologies de l'Information

Epitech Paris | 2024 – en cours  
Spécialisation IA & Data (BAC+5 / RNCP7)

### Master 1 & Licence – Informatique

Université Libanaise | 2020 – 2023

### Certifications

- Claude with Google Cloud's Vertex AI  
Anthropic Academy — 2026
- Evaluating and Debugging Generative AI  
DeepLearning.AI — 2026
- Data Science Coding Challenge  
Coursera – 2026
- Deep Learning Specialization  
Coursera – 2024

## LANGUES

FR / EN : bilingue

## COMPÉTENCES

### IA Générative & LLM

RAG, fine-tuning LLM/VLM, OpenAI API, Anthropic API, LlamaIndex, Hugging Face, embeddings, pgvector, FAISS, prompt engineering, chain-of-thought, NER

### Agents & Orchestration

Agentic workflows, multi-agent orchestration, AI agents, CrewAI, AutoGen, LangChain, function calling, tool use, mémoire d'agent, ReAct, orchestration de pipelines

### MLOps & Production

FastAPI, REST, Docker, Kubernetes, CI/CD, MLflow, PostgreSQL, GCP, model monitoring, évaluation de modèles

### Machine Learning

Scikit-learn, XGBoost, PyTorch, SVM, Naive Bayes, RL, modélisation prédictive, feature engineering, détection d'anomalies, séries temporelles, tests A/B

### Computer Vision

OpenCV, CNN, OCR, classification d'images

### Data Engineering & Visualisation

Python, pandas, NumPy, SQL, Git, web scraping, Spark, Databricks, Streamlit, Plotly, Matplotlib, Seaborn, EDA, statistiques descriptives et inférentielles

### Sécurité

Secure Coding Review, analyse de vulnérabilités

## ENGAGEMENT

- Ambassadrice Epitech (depuis déc. 2024)
- Hult Prize, UL (2023)

## EXPÉRIENCE PROFESSIONNELLE

### Kabeen

Saint-Cloud | Alternance | 2025 – 6 Avril 2026

#### Ingénieure Data & IA

- Conception d'un **pipeline RAG** pour la classification automatisée et l'extraction de métadonnées d'un catalogue applicatif (**LLM, NLP, OpenAI API**).
- Architected scalable data ingestion pipelines** via advanced web scraping for automated catalog enrichment.
- Mise en place d'un système de **Data Quality** et détection d'anomalies (Scikit-learn) garantissant l'indexation de données fiables.
- Full-stack integration** of AI modules into production environments using **React** and **FastAPI**, ensuring seamless industrialization; robust APIs and real-time sync (**PostgreSQL**).
- Résultat : réduction de 80% du temps de traitement** et amélioration de la scalabilité du catalogue.

### BazarSoft (freelance, temps partiel)

Dubai, UAE | Remote | 2025 – en cours

#### Ingénieure Logiciel / Data & IA

- App commerciale** et **dashboard** factures/KPIs vente ; **chatbot** et **reco IA** en production.
- Industrialisation et itérations sur ces briques ; monitoring and evaluating KPIs for automated billing efficiency and reduction in support ticket volume.

### Buy Any Insurance

Dubai, UAE | 2024

#### Software Engineer – Computer Vision / OCR

- Développement d'un moteur d'extraction (**OCR**) sur documents PDF avec **OpenCV** et **Tesseract** pour automatiser l'onboarding client.
- Amélioration de 15% de la précision** de lecture via des scripts de prétraitement d'images optimisés.

### Code Alpha

Remote | 2024

#### Cybersecurity Intern – Python Automation

Détection phishing.  
Application **Secure Coding Review** Python/Flask (qualité, sécurité du code).

## PROJETS & COMPÉTITIONS

### E-invoicing AI

Plateforme e-facturation (client web + app **Electron**) : saisie manuelle et assistée par IA, formats européens, intégration **PDP** ; API **Node.js**, service **Python** ; prototypage d'agents LLM (Claude).

### NER Travel Order

Extraction et structuration d'informations à partir d'ordres de mission pour automatiser la saisie et les contrôles (**CamemBERT, Whisper**).

### Taxi Driver – DRL

Comparaison de Q-Learning et SARSA, puis implémentation d'un agent **DQN** (**Gymnasium**).

### Sentiment Analysis Pipeline

Fine-tuning de modèles Transformers pour classifier et prioriser des messages ; API de scoring déployée en production (**Hugging Face, Transformers**).

### Computer Vision Quality Control

Classification d'images pour le contrôle qualité industriel, avec optimisation du chargement des données et data augmentation (**CNN, EfficientNet, SVM...**).

### KubeQuest

Déploiement d'un cluster Kubernetes : observabilité, GitOps et sécurisation (**Terraform, Helm, Prometheus**).