# AHMED BAKKALI

Ingénieur en science des données, Big Data et intelligence artificielle
ahmedbakkali2003@gmail.com | +212 6 45 53 02 08 | linkedin.com/in/ahmed-bakkali-426076233 | github.com/Ahmed-BAKKALI

## PROFIL

Data & AI Engineer en fin de stage PFE, avec une expérience pratique dans le développement de solutions d'IA et d'automatisation, de la conception à la mise en production. Spécialisé en IA générative et LLMs, j'ai travaillé sur des pipelines RAG end-to-end, l'automatisation de processus, le scripting Python et la gestion d'APIs. Curieux et orienté impact, je recherche une nouvelle opportunité pour continuer à développer mon expertise technique et fonctionnelle.

## EXPÉRIENCE

**AI Engineer (Stage PFE)** Février 2026 – Août 2026 (6 mois)
*Siemens Digital Software Industries* *Rabat*

- Développement d'un pipeline RAG end-to-end pour automatiser la détection de crashes et de dégradations pilotée par LLM, en utilisant l'extraction de données via API, l'embedding, le chunking et la recherche sémantique & hybride indexée dans la base vectorielle Qdrant afin de prédire le build responsable parmi des centaines de suspects, avec plus de 90% de réduction du temps d'analyse et de l'intervention manuelle.
- Automatisation de processus QA manuels avec du scripting Python et l'orchestration de workflows via Prefect Framework afin d'économiser du temps et de réduire l'intervention manuelle.
- **Stack technique :** OpenAI API, LLMs (Qwen 3.6, GPT-OSS, GLM5), RAG, Prompt Engineering, Qdrant, Python, Prefect Framework, Jira, Git, GitLab, CI/CD.

**Data Analytics(Stage PFA)** Juillet 2025 – Octobre 2025 (4 mois)
*Africa Business School – UM6P* *Rabat*

- Conception et développement d'une plateforme data-driven pour l'écosystème des startups marocaines : **thepulse.ma**.
- Mise en place d'un pipeline ETL complet sur Azure (Data Lake Storage, Azure Functions, Azure SQL et Azure App Services), incluant la collecte de données par web scraping et APIs, la transformation et le nettoyage des données, ainsi que la visualisation des KPIs/tableaux de bord.
- **Stack technique :** Azure Data Lake Storage, Azure Functions, Azure SQL, Azure App Services, Python, Flask, Web Scraping, APIs, Beautiful Soup, Selenium, Playwright, Pandas, NumPy, Git, GitHub.

## PROJETS

**Projet Deep Learning : Détection du groupe sanguin à partir des empreintes digitales**
- Développement d'un système basé sur une architecture de réseaux de neurones CNN comprenant deux modèles : VGG16 avec une précision de 88% et ResNet avec une précision de 84%.
- Optimisation des prédictions grâce à des techniques avancées de prétraitement et d'augmentation de données.

**Projet Data Analysis & Web Scraping : analyse du marché immobilier en France**
- Développement d'un pipeline ETL de collecte de données avec web scraping et Beautiful Soup.
- Nettoyage et enrichissement des jeux de données avec Pandas et NumPy, création d'un système de recommandation avec des modèles ML et restitution des insights via un tableau de bord interactif Power BI.

## ÉDUCATION

**Élève ingénieur en science des données, Big Data et IA** 2021 – 2026
*ENSA Tétouan*

**Baccalauréat Sciences Mathématiques A** 2019 – 2021
*Lycée Imam El Ghazali*

## COMPÉTENCES TECHNIQUES

| | |
|---|---|
| **Data** | NumPy, Pandas, Polars, data preprocessing. |
| **Intelligence artificielle** | Machine Learning, Deep Learning, LLMs, OpenAI API, RAG, bases vectorielles, Qdrant, embedding, chunking, recherche sémantique, recherche hybride. |
| **Bases de données** | MySQL, Oracle Server, PostgreSQL, SQL Server. |
| **Visualisation** | Power BI, Excel, Matplotlib, Seaborn, Plotly. |
| **Data Engineering / Big Data** | Pipelines ETL et ELT, Apache Spark, Apache Kafka. |
| **Langages de programmation** | Python, C/C++ |
| **Outils et frameworks** | Jira, Docker, Git, GitHub, UML, Shell/Linux, Scrum Agile. |
| **Web Scraping / APIs** | Requests, Beautiful Soup, Selenium, Playwright. |
| **Cloud** | Azure Data Lake Storage, Azure Functions, Azure SQL et Azure App Services |

## CERTIFICATS

- 365 Data Science : Data Preprocessing with NumPy.
- 365 Data Science : Data Cleaning and Preprocessing with Pandas.
- 365 Data Science : SQL.

## ACTIVITÉS PARASCOLAIRES

**Chef de cellule Contact et Sponsoring** Avril – Octobre 2025
*7e édition du Forum d'entreprises ENSA Tétouan* *Association des étudiants de l'ENSA Tétouan*

- Management d'un comité de plus de 70 personnes réparties en 3 sous-cellules et pilotage de la prospection et de la négociation avec les entreprises partenaires et sponsors.
- Accueil et coordination de 25 entreprises ainsi que de plusieurs invités et intervenants.

**Responsable de mobilité et stages** Mars 2025 – Janvier 2026

## LANGUES ET LOISIRS

**Langues** Français : niveau C1 ; anglais : niveau C1 ; arabe : langue maternelle.
**Loisirs** Football, jeux d'échecs, voyage.