

# Idriss Khattabi

E-mail: drisskhattabi6@gmail.com | Phone: +212 6 35 24 57 97

Portfolio: drisskhattabi6.github.io/id-kh

LinkedIn: linkedin.com/in/idriss-khattabi-b3a266235/

GitHub: github.com/drisskhattabi6



## À propos de moi

Data Analyst / Data Engineer titulaire d'un Master en Intelligence Artificielle & Data Science, spécialisé en analyse de données, ingénierie des données et business intelligence. Maîtrise de Python, SQL et des outils de data engineering/cloud, avec une solide expérience dans la conception de pipelines ETL, le nettoyage, la transformation et l'intégration de données issues de sources multiples. Expérimenté dans le développement de tableaux de bord interactifs avec Power BI, et familier avec l'automatisation des workflows et l'analytique prédictive. Analytique, orienté détail et passionné par la transformation de données en insights exploitables pour la prise de décision.

## Éducation

**M.S.T en Intelligence Artificielle et Data Science**

Faculté des Sciences et Techniques de Tanger

Oct 2023 – Juil 2025

Tanger, Maroc

**Licence en Informatique et Mathématiques**

Faculté Polydisciplinaire de Larache

Oct 2019 – Juil 2023

Larache, Maroc

## Expérience

**Stagiaire AI / Data Engineer**, TNA Consulting – Casablanca

Fév 2025 – Juil 2025

- Conception et développement d'une plateforme SaaS d'IA pour aider les institutions financières à interpréter et exploiter les publications & documents officiels des banques centrales en Afrique.
- Mise en place d'un **pipeline de données** automatisé et évolutif pour le scraping & la collecte de documents depuis des sites web, leur nettoyage, transformation & découpage, puis leur chargement dans DB.
- Développement d'un **chatbot intelligent** utilisant la technique RAG, capable de répondre aux questions des utilisateurs sur les documents des banques centrales africaines.
- Outils utilisés : **Python, Marker, Streamlit, Apache Airflow, ETL Pipeline, Flask, Ollama, Faiss, RAG, LLM, Selenium, Docker**

**Assistant de recherche**, Faculté Polydisciplinaire – Larache

Mars 2023 – Juin 2023

- Conception et implémentation d'un système d'optimisation pour améliorer la cohérence des matrices de comparaison par paires dans l'AHP, en exploitant **GA** et **PSO**, avec le développement d'une application web Django facilitant la génération des matrices AHP.
- Co-auteur de **trois** articles de recherche avec le **Dr. Chakir Tajani**, exemples :
  - A Metaheuristic Approach to Improve Consistency of the Pairwise Matrix in AHP
  - Generating a Set of Consistent Pairwise Comparison Test Matrices in AHP using PSO

## Compétences comportementales (Soft Skills)

Esprit analytique et résolution de problèmes – Apprentissage rapide – Travail en équipe et collaboration – Agilité et adaptabilité – Curiosité pour les nouvelles technologies

## Compétences techniques

- **Programming & Scripting:** Python, SQL, Java, JavaScript, C, Bash, Solidity
- **Data Analysis & Visualization:** NumPy, Pandas, Matplotlib, Seaborn, Power BI, Power Query, Tableau, Excel
- **Big Data & Distributed Systems:** Apache Spark (PySpark), Apache Kafka, Hadoop, Apache Airflow
- **Machine & Deep Learning & NLP:** Scikit-learn, TensorFlow, Keras, PyTorch, XGBoost, MLflow, NLTK, SpaCy
- **Generative AI & LLMs:** RAG, Prompt Engineering, Fine-tuning, LangChain, LangGraph, LlamaIndex, n8n
- **Databases:** MySQL, SQL Server, PostgreSQL, MongoDB, Supabase, IPFS, FAISS, Chroma, Qdrant, Milvus
- **Web Development:** HTML, CSS, Tailwind CSS, Javascript, ReactJS/NextJs, Django, Flask, FastAPI, Streamlit
- **DevOps & Tools:** Linux, Git/GitHub, Docker, Microsoft Azure, Selenium, BeautifulSoup, PyTest

## Langues

**Arabic:** Native proficiency

| **English:** Professional (B2)

| **French:** Professional (B2)

## Certifications

Associate Data Analyst in SQL – DataCamp  
Microsoft Azure Fundamentals (AZ-900) – DataCamp

Data Analyst with Python – DataCamp  
Associate Data Engineer – DataCamp

## Projets

### Construction et Analyse d'un Data Warehouse E-Commerce

[GitHub Repo]

- Mise en place d'un pipeline ETL pour créer un Data Warehouse destiné à l'analyse des produits et fournisseurs E-commerce, avec conception de tableaux de bord et rapports.
- Outils utilisés : Python, Pandas, SQLAlchemy, MySQL, Power BI

### Analyse en temps réel du sentiment des posts Twitter

[GitHub Repo]

- Création d'un système temps réel pour prédire le sentiment des publications Twitter.
- Outils utilisés : PySpark, Kafka, Docker, Django, MongoDB, ChartJS

### Météo Data Mining

[GitHub Repo]

- Application de techniques de fouille de données sur des données météorologiques pour extraire des informations pertinentes et identifier des schémas pour la prévision météo.
- Outils utilisés : Python, Machine Learning

### Application Desktop d'Analyse de Données & Machine Learning

[GitHub Repo]

- Application desktop en Python construite avec CustomTkinter permettant aux utilisateurs d'analyser des jeux de données qualitatifs et quantitatifs et d'appliquer plusieurs algorithmes de machine learning via une interface intuitive.
- Outils utilisés : Python, Tkinter / CustomTkinter, Scikit-learn, Pandas, Matplotlib

### Chatbot de Recettes Marocaines

[GitHub Repo]

- Un chatbot IA utilisant la technique RAG pour fournir des recettes marocaines authentiques avec ingrédients, étapes, et images optionnelles, compatible avec des LLMs locaux et cloud.
- Outils utilisés : Python, Streamlit, LangChain, Ollama, Sambanova, Faiss DB

### Analyse et Prédiction du Marché Boursier

[GitHub Repo]

- Implémentation d'un modèle de Deep Learning pour analyser et prévoir les prix futurs des actions à partir de données historiques.
- Outils utilisés : Python, Pytorch, prédiction de séries temporelles

### Espace de Données pour Dossiers Médicaux Électroniques

[GitHub Repo]

- Projet Data Space médical offrant une plateforme complète pour la gestion, le traitement et l'exploration de données de santé.
- Outils utilisés : Python, MySQL, pipeline ETL, API RESTful

### Convertisseur Texte en SQL

[GitHub Repo]

- Outil interactif permettant de convertir du texte utilisateur en requêtes SQL via un agent IA, d'interroger une base SQL et de visualiser les résultats.
- Outils utilisés : Python, LangGraph, Streamlit, Pandas

### Détection de l'Âge, du Genre & des Émotions avec CNN

[GitHub Repo]

- Détection de l'âge, du genre et de l'état émotionnel d'une personne à partir d'images faciales.
- Outils utilisés : Python, OpenCV, Tkinter, TensorFlow/Keras, CNN, NumPy, modèles pré-entraînés

### Analyste de Données Exploratoires

Analyse des ventes - Analyse du comportement des consommateurs - le diabète - les employés