

# 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

Analyste de données titulaire d'un Master en Intelligence Artificielle et Science des Données, spécialisé en analyse de données, ingénierie des données et intelligence d'affaires. Compétent en Python, SQL et Power BI, avec une expérience dans la création de pipelines ETL, le nettoyage et la transformation des données, ainsi que le développement de tableaux de bord interactifs pour l'aide à la décision. Esprit analytique et passionné par la transformation des données brutes en informations exploitables.

## Éducation

**Master 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 Pluridisciplinaire de Larache

Oct. 2019 – Juil. 2023

Larache, Maroc

## Expérience

**Stagiaire Ingénieur IA / Data**, TNA Consulting – Casablanca

Fév. 2025 – Juil. 2025

- Conception et développement d'une plateforme SaaS basée sur l'IA pour aider les institutions financières à interpréter et exploiter les publications et documents officiels des banques centrales à travers l'Afrique.
- Mise en place d'un **pipeline de données** automatisé et évolutif pour le scraping et la collecte de documents depuis des sites web, leur nettoyage, transformation, segmentation, puis intégration dans une base de données vectorielle.
- 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, Pipeline ETL, Flask, Ollama, Faiss, RAG, LLM, Selenium, Docker**

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

Mars 2023 – Juin 2023

- Conception et implémentation d'un système d'optimisation visant à améliorer la cohérence des matrices de comparaison par paires dans la méthode **AHP**, en utilisant les algorithmes **GA** et **PSO**, avec le développement d'une application web Django facilitant la génération de matrices AHP.
- Co-auteur de **trois** articles de recherche avec **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

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

## Compétences techniques

- **Programmation & Scripts** : Python, SQL, Java, JavaScript, C, Bash, Solidity
- **Analyse & Visualisation de données** : NumPy, Pandas, Matplotlib, Seaborn, Plotly, Power BI, Power Query, Excel
- **Big Data & Systèmes distribués** : Apache Spark (PySpark), Apache Kafka, Hadoop, Apache Airflow
- **Machine / Deep Learning & NLP** : Scikit-learn, TensorFlow, Keras, PyTorch, XGBoost, MLflow, NLTK, SpaCy
- **IA Générative & LLMs** : RAG, Prompt Engineering, Fine-tuning, LangChain, LangGraph, LlamaIndex, n8n
- **Bases de données** : MySQL, SQL Server, PostgreSQL, MongoDB, Supabase, IPFS, FAISS, Chroma, Qdrant, Milvus
- **Développement web** : HTML, CSS, Tailwind CSS, JavaScript, ReactJS/NextJs, Django, Flask, FastAPI, Streamlit
- **DevOps & Outils** : Linux, Git/GitHub, Docker, Microsoft Azure, Selenium

## Langues

**Arabe** : Langue maternelle

**Anglais** : Professionnel (B2)

**Français** : Professionnel (B2)

## Certifications

**Associate Data Analyst in SQL – DataCamp**  
**Data Manipulation – DataCamp**

**Data Analyst with Python – DataCamp**  
**Associate Data Engineer – DataCamp**

## Projets

### Construction et Analyse d'un E-Commerce Data Warehouse [GitHub Repo]

- Mise en œuvre d'un pipeline ETL pour créer un Data Warehouse destiné à l'analyse des produits et fournisseurs d'un site e-commerce, avec création de dashboards et de rapports analytiques.
- Outils utilisés : Python, Pandas, SQLAlchemy, MySQL, Power BI

### Real-time Twitter Post Sentiment Analysis [GitHub Repo]

- Développement d'un système en temps réel pour prédire le sentiment des tweets en streaming via Kafka, avec prédiction en temps réel dans PySpark et visualisation dynamique des résultats.
- Outils utilisés : PySpark, Kafka, Docker, Django, MongoDB, ChartJS

### Meteo Data Mining [GitHub Repo]

- Application de techniques de Data Mining sur des données météorologiques pour extraire des insights pertinents et identifier des patterns pour la prévision du climat.
- Outils utilisés : Python, Machine Learning

### Data Analysis & Machine Learning Desktop Application [GitHub Repo]

- Application de bureau en Python développée avec Tkinter permettant aux users d'analyser des datasets et d'appliquer plusieurs algorithmes de Machine Learning via une interface graphique intuitive.
- Outils utilisés : Python, Tkinter / CustomTkinter, Scikit-learn, Pandas, Matplotlib

### Data Space for Electronic Medical Records [GitHub Repo]

- Projet fournissant une plateforme complète pour la gestion, le traitement et l'exploration des données médicales électroniques.
- Outils utilisés : Python, MySQL, ETL Pipeline, RESTful API

### Stock Market Analysis & Prediction [GitHub Repo]

- Implémentation d'un modèle Deep Learning pour analyser et prédire les prix futurs des actions à partir de données historiques.
- Outils utilisés : Python, PyTorch, Time-Series Prediction

### 2D & 3D Retrieval Systems [GitHub Repo]

- Développement d'applications web pour la recherche d'images et de modèles 3D basées sur le contenu, en utilisant des descripteurs visuels et de forme.
- Outils utilisés : Python, Flask, MongoDB, (Angular / ExpressJS)

### Fake News Detection [GitHub Repo]

- Fine-tuning du modèle BERT pour détecter les fake news et déploiement du modèle entraîné via une application web Flask pour la prédiction en temps réel.
- Outils utilisés : Python, Flask, PyTorch, BERT

### Text To SQL Converter [GitHub Repo]

- Outil interactif permettant de convertir du texte utilisateur en requêtes SQL via un AI Agent, avec extraction et visualisation des données issues d'une base SQL.
- Outils utilisés : Python, LangGraph, Streamlit, Pandas

### Age, Gender & Emotion Detection using CNN [GitHub Repo]

- Détection de l'âge, du genre et de l'état émotionnel à partir d'images faciales en utilisant des réseaux de neurones convolutifs (CNN).
- Outils utilisés : Python, OpenCV, Tkinter, TensorFlow/Keras, CNN, NumPy, Pretrained Models