

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

Diplômé de Master en science de l'IA et des données avec de fortes compétences en analyse des données, en génie des données et en développement complet. Compéoint dans Python, SQL et Django, avec une expérience dans la création de pipelines de données, de tableaux de bord et de solutions analytiques. Il est également familier avec l'apprentissage automatique et les applications d'IA. Apprenant rapide,

Éducation

MST en Intelligence Artificielle et Science des Données,

2023 oct - 2025 juillet

Faculté des Sciences et Techniques de Tanger

Tanger - Morocco

Licence en Mathématique et Informatique,

2019 oct - 2023 juillet

Faculté multidisciplinaire de Larache

Larache - Morocco

Expérience

Stagiaire Ingénieur en IA, TNA Consulting – Casablanca

Février 2025 – Juin 2025

- Conception et développement d'une plateforme SaaS IA full-stack pour aider les institutions financières à interpréter et exploiter les publications officielles des banques centrales africaines.
- Développement d'un **chatbot intelligent** capable de répondre aux questions des utilisateurs sur les documents des banques centrales africaines.
- Conception d'un **pipeline de données** automatisé et évolutif pour le crawling, le scraping, le nettoyage, la transformation et le chargement des documents financiers dans une base de données vectorielle.
- Outils utilisés : **Python, Marker, Streamlit, Apache Airflow, ETL, Flask, Ollama, Faiss, RAG, LLM**

Stagiaire Chercheur, Faculté Polydisciplinaire – Larache

Mars 2023 – Juin 2023

- Conception et mise en œuvre d'un système d'optimisation pour 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 sous Django facilitant la génération des matrices AHP.
- Co-auteur de **trois** articles de recherche avec **Dr. Chakir Tajani**, exemples d'articles :
 - 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 – Communication efficace – Curiosité pour les nouvelles technologies

Compétences Techniques

- **Programmation & Scripts** : Python, SQL, Java, JavaScript, C, Bash, Solidity, LaTeX
- **Analyse de données & Visualisation** : NumPy, Pandas, Matplotlib, Seaborn, Power BI, Tableau
- **Apprentissage automatique & profond** : Scikit-learn, TensorFlow, Keras, PyTorch, XGBoost, MLflow
- **Traitement du langage naturel (NLP)** : NLTK, spaCy, Hugging Face Transformers, GloVe, BERT, TesseractOCR
- **IA générative & LLMs** : RAG, Ingénierie de prompts, Fine-tuning (LoRA, QLoRA), LangChain, LangGraph, LlamaIndex, OpenAI, Ollama, Hugging Face, n8n
- **Big Data & Systèmes distribués** : Apache Spark (PySpark), Apache Kafka, Hadoop, Apache Airflow
- **Bases de données & Stockage** : MySQL, SQL Server, PostgreSQL, MongoDB, Supabase, IPFS, Bases vectorielles (FAISS, Chroma, Qdrant, Milvus, Weaviate)
- **Pipelines de données & Automatisation** : Processus ETL, APIs REST, Postman, Selenium, BeautifulSoup, PyTest
- **Développement Web** : HTML, CSS, Tailwind CSS, React.js, Next.js, Django, Flask, FastAPI, Streamlit
- **DevOps & Cloud** : Linux, Git/GitHub, Docker, Microsoft Azure

Langues

Arabe : Langue maternelle

| **Anglais** : Professionnel (B2)

| **Français** : Professionnel (B2)

Certifications

Associate Data Scientist in Python – DataCamp
Associate AI Engineer for Data Scientists – DataCamp
Microsoft Azure Fundamentals (AZ-900) – DataCamp

Associate Data Engineer – DataCamp
Data Analyst – DataCamp

Projects

Entrepôt de Données pour le Commerce et Analyse

[Dépôt GitHub]

- Pipeline ETL pour cree un data warehouse pour analyser les produits et fournisseurs avec un dashboard.
- Outils utilisés : Python, Pandas, SQLAlchemy, MySQL, Power BI

Analyse de Sentiment en Temps Réel de Twitter

[Dépôt GitHub]

- Système en temps réel pour prédire le sentiment des publications sur Twitter (X).
- Outils utilisés : PySpark, Kafka, Docker, Django, MongoDB, ChartJS

Systèmes de recherche 2D & 3D

[Dépôt GitHub]

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

Application Desktop d'Analyse de Données & ML

[Dépôt GitHub]

- Application Python avec interface graphique (GUI) utilisant CustomTkinter permettant d'analyser des jeux de données et d'appliquer des algorithmes de ML.
- Outils utilisés : Python, CustomTkinter, Scikit-learn, Pandas, Matplotlib

Medical Document Management Blockchain System

[GitHub Repo]

- A decentralized solution for secure medical document storage and access using Ethereum and IPFS. Empowers patients with document control, ensures transparency, and supports multi-platform access for doctors and admins. (Web & Desktop App)
- Tools Used: Solidity, Ethereum, IPFS, ReactJS, Python, Tkinter, Hardhat, MetaMask, Docker

ChatPDF – Discuter avec un PDF en local

[Dépôt GitHub]

- Application web locale RAG pour interroger des documents PDF via des modèles de langage, avec conversion PDF en Markdown, recherche vectorielle et chatbot intégré.
- Outils utilisés : Python, Streamlit, LangChain, Ollama, Sambanova, ChromaDB, Marker

AI Scraper

[Dépôt GitHub]

- Outil interactif pour extraire, traiter et structurer des données web en utilisant des LLMs.
- Outils utilisés : Python, Selenium, Crawl4AI, Streamlit, Ollama, Sambanova, Pandas

Détection de l'âge, du genre et des émotions à partir du visage via CNN

[Dépôt GitHub]

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

Bot de Résumé de Transcription YouTube

[Dépôt GitHub]

- Workflow d'automatisation basé sur l'IA qui extrait la transcription de vidéos YouTube envoyées via un bot Telegram, et la fournit avec un résumé sous forme de fichier .txt.
- Outils utilisés : n8n, API Telegram Bot, RapidAPI, JavaScript, Google Gemini

Développement Web avec Django

- Marketplace - Chat avec LLM en local - Réseau Social - Blog - Notes - Chat Room - E-commerce