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

About me

Data Analyst with a Master's in AI & Data Science, skilled in data analysis, data engineering, and business intelligence. Proficient in Python, SQL, and Power BI, with experience in building ETL pipelines, cleaning and transforming data, and developing interactive dashboards for decision-making. Familiar with machine learning applications and predictive analytics. Analytical, detail-oriented, and passionate about turning raw data into actionable insights.

Education

M.S.T in Artificial Intelligence and Data Science	Oct 2023 – July 2025
Faculty of Sciences and Technologies of Tangier	Tangier, Morocco
BS in Computer Science and Mathematics	Oct 2019 – July 2023
Faculty multidisciplinary of Larache	Larache, Morocco

Experience

AI / Data Engineer Intern, TNA Consulting – Casablanca   Feb 2025 – July 2025

- Designed and developed an AI SaaS platform to support financial institutions in interpreting and leveraging official central bank publications & documents across Africa.
- Engineered an automated and scalable **Data pipeline** for scraping & collecting documents from websites, cleaning, transforming & chunking documents, and loading into a Vector database.
- Built an intelligent **chatbot** using RAG Technique capable of answering user questions on African Central Bank documents.
- Used Tools : **Python, Marker, Streamlit, Apach Airflow, ETL Pipeline, Flask, Ollama, Faiss, RAG, LLM, Selenium, Docker**

Research Assistant, Faculty of Poly-disciplinary – Larache   March 2023 – June 2023

- Designed and implemented an optimization system to improve the consistency of pairwise comparison matrices in the **AHP**, leveraging **GA** and **PSO**, with Developing a Django web application that makes it easier to generate AHP matrices.
- Co-authored **three** research papers with **Dr. Chakir Tajani**, papers examples :
  - 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

Soft Skills

Analytical thinking and problem-solving – Fast learning – Teamwork and collaboration – Agility and Adaptability – Effective communication – Curious for new technologies

Technical Skills

- **Programming & Scripting:** Python, SQL, Java, JavaScript, C, Bash, Solidity
- **Data Analysis & Visualization:** NumPy, Pandas, Matplotlib, Seaborn, Plotly, Power BI, Power Query
- **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

Languages

Arabic: Native proficiency | **English:** Professional (B2) | **French:** Professional (B2)

Certifications

Associate Data Analyst in SQL – DataCamp   **Data Analyst with Python –** DataCamp
Data Manipulation – DataCamp   **Associate Data Engineer –** DataCamp

Projects

Build and Analyze E-Commerce Data Warehouse   [GitHub Repo]

- Implement an ETL pipeline to Create Data Warehouse for E-commerce product and supplier analytics with creating dashboards and Reports.
- Tools Used: Python, Pandas, SQLAlchemy, MySQL, Power BI

Real-time Twitter Post Sentiment Analysis   [GitHub Repo]

- Created a real-time system to predict sentiment from Twitter posts.
- Tools Used: PySpark, Kafka, Docker, Django, MongoDB, ChartJS

Meteo Data Mining   [GitHub Repo]

- Applied data mining techniques on meteorological data to extract meaningful insights and identify patterns for weather prediction.
- Tools Used: Python, Machine Learning

2D & 3D Retrieval Systems   [GitHub Repo]

- Built web applications for content-based image and 3D model retrieval using visual and shape descriptors.
- Tools used: Python, Flask, MongoDB, (Angular / ExpressJS)

Data Analysis & Machine Learning Desktop Application   [GitHub Repo]

- A Python desktop app built with CustomTkinter that enables users to analyze qualitative and quantitative datasets and apply multiple machine learning algorithms through an intuitive GUI.
- Tools Used: Python, Tkinter / CustomTkinter, Scikit-learn, Pandas, Matplotlib

Movie Recommendation System   [GitHub Repo]

- Built a movie recommendation system using two approaches: Content-Based Filtering and Collaborative Filtering.
- Tools Used: Python, Jupyter Notebook

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

AI Scraper   [GitHub Repo]

- An interactive tool to scrape, process, and structure web data using local or cloud-based LLMs.
- Tools Used: Python, Selenium, Crawl4AI, Streamlit, Ollama, Sambanova, Pandas

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

- Detecting person Age, Gender, and Emotional state from facial images using CNN.
- Tools Used: Python, OpenCV, Tkinter, TensorFlow/Keras, CNN, NumPy, Pretrained Models

Exploratory Data Analyst

Sales Analysis - Consumer Behavior Analytics - Diabetes Dataset - Employees Dataset

Web Development with Django

Marketplace Website - Chat With LLM - Blog Website - Note App - Room Chat App - E-commerce App