

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



Education

Master in Artificial Intelligence and Data Science

Faculty of Sciences and Technologies

Oct 2023 – July 2025

Tangier, Morocco

Bachelor in Computer Science and Mathematics

Faculty multi-disciplinary

Oct 2019 – July 2023

Larache, Morocco

Experience

Digitalization Engineer Intern, KOSTAL Group – Tangier

Jan 2026 - Present

- Analyzed production data to identify bottlenecks, defect patterns, and performance drifts
- Developed and implemented a digital platform that streamlined Project Launch Management, improving process tracking, reporting, and cross-team visibility.
- Acquired strong domain knowledge in industrial production systems, automation processes, and manufacturing performance metrics.

AI Developer Intern, Bid2Win – Remote - Part-Time

Dec 2025 - Present

- Developing an AI-powered platform to automate tender analysis, reducing manual document review time and improving filtering efficiency.
- Built automated data pipelines for scraping, processing, transforming, and indexing unstructured tender documents into a Vector DB.
- Implemented a RAG-based feature extraction pipeline to automate insight extraction from tender documents for scoring and recommendation.
- Engineered a microservices architecture using Docker and REST APIs, improving scalability and deployment flexibility on Azure.

AI Developer Intern, High Commission for Planning – Remote

Sept 2025 – Dec 2025

- Designed and developed an AI-powered platform to automate documentation and management of statistical processes based on the **GSBPM** framework.
- Engineered a robust data pipeline for document ingestion and processing from multiple sources.
- Implemented an advanced RAG engine using **HyDE** prompting to generate structured GSBPM-compliant documentation with evidence and audit trails.
- Reduced statistical documentation time from days to hours by automating information extraction and report generation workflows.

AI Developer Intern, TNA Consulting – Casablanca

Feb 2025 – July 2025

- Designed and developed an AI platform to help financial institutions interpret and leverage official African central bank publications and documents.
- Engineered an automated and scalable data pipeline for scraping, transforming, chunking, and indexing large volumes of financial documents into a Vector DB.
- Built an intelligent RAG-based chatbot capable of answering user questions on African central bank documents with contextual retrieval.
- Improved accessibility and searchability of financial regulatory information, reducing manual research effort for users.

Research Assistant, Faculty of Poly-disciplinary – Larache

March 2023 – June 2023

- Developed an optimization system to improve the consistency of AHP pairwise comparison matrices using Genetic Algorithms (GA) and Particle Swarm Optimization (PSO).
- Built a Django web application for automated AHP matrix generation and validation, simplifying the

decision-making workflow.

- Co-authored **three** research papers with **Dr. Chakir Tajani**: Paper 1 – Paper 2 – Paper 3

Languages

Arabic: Native proficiency | **English:** Professional (B2-C1) | **French:** Professional (B2)

Certifications

Associate Data Scientist in Python – DataCamp

Microsoft Azure Fundamentals – DataCamp

Associate AI Engineer for Data Scientist – DataCamp

Associate Data Engineer – DataCamp

Skills

- **Programming Languages:** Python, Java, JavaScript, SQL, C, Solidity, Bash
- **Machine Learning & Generative AI :** Scikit-learn, TensorFlow, Keras, PyTorch, XGBoost, MLFlow, LangChain, LangGraph, n8n, CrewAI
- **Web Development:** HTML, CSS, Tailwind, JavaScript, React/NextJS, Django, Flask, FastAPI
- **Libraries & Tools:** NumPy, Pandas, Matplotlib, Seaborn, Plotly, OpenCV, Selenium.
- **Big Data & Distributed Systems:** Apache Spark (PySpark), Apache Kafka, Apache Hadoop, Apache Airflow
- **Databases:** MySQL, SQL Server, PostgreSQL, MongoDB, Chroma, Qdrant, Weaviate, Snowflake, Redis
- **DevOps & Tools:** Linux, Git/GitHub, CI/CD, Azure, AWS, Docker/Docker-Compose, Power BI, Postman
- **Soft Skills:** Analytical & Problem-solving Mindset, Fast learning, Teamwork and collaboration

Projects

ChatPDF – Chat with PDF Locally

[GitHub Repo]

- A local RAG web app to query PDF documents using LLMs, with support for PDF-to-Markdown conversion, vector search, and chatbot-style Q&A.
- Tools Used: Python, Streamlit, LangChain, Ollama, Sambanova, ChromaDB, Marker

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

Real-time Twitter Sentiment Analysis

[GitHub Repo]

- Created a real-time system to predict sentiment from Twitter posts with visualizing the result in live.
- Tools Used: PySpark, Kafka, Docker, Django, MongoDB, ChartJS

Data Analysis & Machine Learning Desktop Application

[GitHub Repo]

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

Data Space for Electronic Medical Records

[GitHub Repo]

- Medical Data Space Project provides a comprehensive platform for managing, processing, and exploring healthcare data.
- Tools Used: Python, MySQL, ETL Pipeline, RESTful API

Web Development with Django

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