

Elmahdi Elaazmi

Senior Full-Stack AI Engineer
& Technical Lead

elaazmielmahdi@gmail.com
+212 675874185

Profiles

Linkedin
elmahdi-elaazmi

Github
Alcheemiist

Portfolio
elmahdi-elaazmi.com

Skills

Leadership & Delivery
Technical leadership & team management, Mentoring & onboarding, Cross-functional collaboration, Architecture & technical roadmap, Agile delivery, Stakeholder & client communication

Programming Languages
Python (Advanced), TypeScript, JavaScript, C, C++, Rust, SQL, Bash/Shell, HTML/CSS, Assembly

AI and Machine Learning
LLMs & Fine-tuning, Generative AI, AI Agents (LangGraph | LangChain), RAG, NLP & Transformers, Neural networks & predictive modeling, Computer Vision (YOLO | Detectron | MediaPipe), OCR

ML Engineering & MLOps
PyTorch, TensorFlow, Keras, Hugging Face, scikit-learn, OpenAI API, Anthropic API, ML CI/CD, Edge AI (NVIDIA Jetson), IoT Pipelines

Data & Databases
PostgreSQL, MySQL, Redis, Neo4j (graph), Vector databases, ETL, Feature Engineering, Statistical analysis, Pandas, NumPy, Seaborn

Backend & Frontend
FastAPI, Django, Flask, Node.js, Express.js, RESTful APIs, API integrations, React.js, Next.js, Streamlit

Cloud & DevOps
Docker & Docker Compose, Kubernetes, ArgoCD, CI/CD, AWS, DigitalOcean, Linux / Unix, Git / GitHub / GitLab, Vagrant, Shell Scripting

Summary

Strategic Senior Full-Stack AI Engineer & Technical Lead with 5+ years bridging advanced ML research and production-grade software. I lead cross-functional teams that turn AI prototypes into scalable systems, owning the full arc from architecture to delivery. Fluent across the full stack, AI/ML, agentic AI, RAG, robust backends, and cloud-native infrastructure. I lead by owning the whole problem, setting standards and mentoring engineers, and I bring a founder's ownership to everything I build.

Experience

Senior R&D Engineer - Technical Lead Hybrid, Rabat - MAR
UM6P - SCI October 2025 - Present

- Led a 5-person cross-functional team to rebuild Deliberatorium, a legacy platform that uses AI/LLMs to structure large-scale deliberation into a modern, production-grade system (Aideator).
- Re-architected the full stack (React/TypeScript, FastAPI, PostgreSQL, Redis), containerised with Docker Compose and deployed on DigitalOcean Linux servers, scaling capacity from ~50 to 1,000+ concurrent users (20x) with high uptime and real-time incident detection.
- Owned the end-to-end migration from the legacy system, preserving and ingesting historical data while translating research-paper concepts into production AI/LLM features.
- Designed the infrastructure and ML lifecycle and set the team's tooling and engineering standards (Docker-based deploys, code review, CI practices), adopted as the default across the project.

AI Engineer Consultant Remote, Dubai - UAE
SURETREE June 2025 - October 2025

- Served as a core contributor on a fintech's AI/Data team, partnering with product and data managers to design and ship bespoke GenAI/RAG solutions for both internal teams and external clients.
- Built agentic systems (LangGraph, LangChain, FastAPI) for real-time insight gathering powering document processing, internal automation, and decision support, with outputs delivered as automated reports, newsletters, and alerts that cut manual reporting time by 60%.
- Integrated a Neo4j knowledge graph, vector databases, and HubSpot to ground agents in company data and wire outputs into existing client tooling, sustaining 90% answer accuracy.
- Advised on system architecture to balance accuracy and cost, reducing LLM/inference spend by 30% and driving 5 proposed use cases into production, delivered solo or alongside the engineering team.

AI Engineer Hybrid, Marrakech - MAR
PROD'AIR November 2024 - April 2025

- Architect and lead developer of a real-time computer-vision system on NVIDIA Jetson, running YOLO inference to detect and classify urban objects for a digital-twin platform supporting urban services and incident prevention.
- Cut cloud costs by 30% and deployment time by 50% by moving inference from DigitalOcean cloud to edge processing on Jetson, improving detection accuracy by 60% over the prior cloud-based setup.
- Built an energy-monitoring pipeline across 50+ IoT sensors, streaming area-level consumption data at 95% fidelity to identify energy leaks and anomalies through automated ingestion.
- Delivered Llama-2 RAG pipelines with OCR to retrieve and surface insights from internal documents and archives, powering employee self-service.
- Owned all architecture and core development end-to-end, directing collaborating engineers on implementation and testing.

Community Volunteering

President of LeetMakers Robotics Club 1337 / UM6P

February 2021 - June 2022

- Led a 50+ member university engineering club, driving a 40% increase in member engagement.
- Organised 10+ technical workshops and 3 hackathons, building a hands-peer-to-peer learning community.
- Secured sponsorship from UM6P ecosystem tech companies to fund events and equipment.

Languages

English
Advanced

French
Upper-intermediate

Arabic
Native

Experience

R&D Engineer

DECENTURE

Remote, London - UK

May 2023 - June 2024

- Built 5+ production NLP models (GPT/BERT) reaching 95% validation accuracy, analysing student scores, remarks, and profiles to power personalised insights and services within the platform.
- Engineered reproducible MLOps workflows (Git, CI/CD, PostgreSQL/Django) enabling reliable retraining, faster model iteration, and automated deployment.
- Delivered full-stack features across a web/mobile product, internal tools and client dashboards using React, Node.js, and Express.js.

Co-Founder & Technical Lead

PHOENIX SERVICES

Benguerir - MAR

Jun 2021 – May 2023

- Co-founded and served as Technical Lead of a campus-digitalisation startup, building a campus app, student portal, and services platform from concept to production.
- Scaled to ~5,000 active users (students, staff, and PhD researchers) with campus-wide administration adoption. The platform sustained 5 years of operation and was ultimately acquired by UM6P.
- Architected the full system end-to-end (POC → production), owning all technology-stack and infrastructure decisions while leading a team of engineers.
- Ran the venture as a founder, fundraising, university partnerships, hiring, and product strategy and contributed to pitching and client meetings.

Full-stack Developer

GREEN AND SMART BUILDING PARK (IRESEN)

Part-time, Benguerir - MAR

August 2021 – March 2022

- Delivered a custom Django CMS-driven website for a green-energy research park, driving a 60% increase in digital engagement.
- Built admin dashboards with analytics to track site-management efficiency and energy-data reporting.
- Ran iterative client consultations to keep technical delivery aligned with green-energy operational needs.

Full-Stack Developer

UPWORK

Freelance

January 2021 – June 2021

- Delivered custom AI features, automation, and full-stack web applications for international clients.

Education

Certificate of RNCP7

1337 · 42 Network, Benguerir - MAR

AI & Data

Expected October 2026

- State-recognised with the French national qualifications framework (France Compétences) and EU-recognised as equivalent to a Master's degree (Bac+5).

Certificate of Digital Technologies Architect

1337 · UM6P · 42 Network, Benguerir - MAR

Computer Science - AI & ML

September 2019 - April 2026

- Computer Science with an AI & Machine Learning specialisation. Completed the advanced pathway, reaching Level 22 (Computer Architect).

Diploma of Specialized Technician

IFMIAC, Casablanca - MAR

Industrial Automated Systems

September 2017 - July 2019