

HAITHAM HATTAN

AI Engineer & Full Stack Software Engineer

Riyadh, Saudi Arabia — Saudi National | info@haithamhattan.sa | +966 53 743 0772
linkedin.com/in/HaithamHattan | github.com/Haitham8888 | x.com/HattanHaitham | haithamhattan.sa

PROFESSIONAL SUMMARY

Generative AI Engineer and Full Stack Software Engineer shipping client products, teaching AI, and building enterprise systems. Specializes in Retrieval-Augmented Generation (RAG) pipelines, fine-tuning and serving local Large Language Models (LLMs), agentic AI tools, and GPU-accelerated LLM training, engineering fully offline, privacy-first AI infrastructure and enterprise systems sustaining 99.9% availability, aligned with Saudi Vision 2030. A self-taught programmer since 2014 who built his first Android app, a Word-to-PDF converter in Java, he is now a two-time consecutive Employee of the Year at MDS CS, with 860+ volunteer hours in AI education and a tech YouTube channel with 156K+ views.

PROFESSIONAL EXPERIENCE

AI Engineer

MDS for Computer Systems (MDS CS) | Riyadh, Saudi Arabia | Apr 2026 – Present

- Architecting *Retrieval-Augmented Generation (RAG)* solutions with *Python*, *FastAPI*, and *local Large Language Models (LLMs)*, powering production-ready, fully offline, privacy-first AI deployments.
- Training, fine-tuning, and serving large language models on **16+ NVIDIA H100 and H200 GPUs** with *SGLang*, optimizing throughput, latency, and inference cost while monitoring model performance.
- Building security-critical AI components in *C++* and *Rust*, while shipping the majority of AI systems in *Python*.
- Engineering *Milvus* vector search and LLM automation tools that improved developer productivity by **40%**, and managing *Milvus* clusters containerized with *Kubernetes (k8s/k9s)* and *Helm*.
- Building *MCP (Model Context Protocol)* servers that give AI agents device control and *Playwright*-based browser automation for automated vulnerability discovery.

Full Stack Software Engineer

MDS for Computer Systems (MDS CS) | Riyadh, Saudi Arabia | Jul 2025 – Apr 2026

- Developing enterprise applications with *Java*, *Spring Boot*, and *DB2*, integrating *RESTful APIs* that improved data exchange efficiency by **30%**.
- Deploying enterprise systems on *IBM WebSphere (WAS)* with high-availability patterns sustaining **99.9% uptime**, *HCL Digital Experience (HCL DX)* portals, and *IBM FileNet* document workflows.
- Delivering the fully offline **National Address GIS Platform** with *Esri GIS* and *Bun.js*, building high-performance web services with *TypeScript* and the *Bun runtime*.
- Integrating *Active Directory* with login interfaces across *Java*, *Python*, *C++*, and *Rust* applications.

Software Development Consultant

Self-Employed | Remote, Saudi Arabia | Dec 2014 – Jun 2025

- Self-taught since 2014**; built his first Android app, a Word-to-PDF converter in *Java*, growing into **10+ Android mobile apps** for clients.
- Delivered **20+ web platforms** for private clients reaching **5,000+ users**, including the G-Care medical product (*SolidJS*, *Vite*) owned end-to-end from UI design to deployment.
- Built **Edugate Portal** for Jazan University — student academic services (registration, grades, schedules) in *PHP* and *MySQL* — and redesigned the university website with *Nafath* login.
- Built **platforms, services, automation bots, and internal tools** — including WhatsApp-based bots and computer-vision AI systems (*Node.js*, *Python*, *OpenCV*) — and delivered client work through *Fiverr*, *Khamsat*, and *Mostaqel*.

Software Engineer Intern

HackerRank | Remote, Saudi Arabia | Oct 2023 – Jan 2024

- Solved **100+ advanced algorithmic challenges** in *Python* and *C++* applying *data structures and algorithms* across multiple difficulty levels.
- Applied **complexity analysis** to design **efficient, optimized solutions**, sharpening debugging, code-quality, and problem-solving skills with each challenge.
- Built a consistent **daily problem-solving routine** on the HackerRank platform, completing the program with a **certificate of completion**.

Network Security Engineer

eHealth SA | Baish General Hospital, Saudi Arabia | Jun 2023 – Sep 2023

- Deployed **80+ computers and printers** at Baish General Hospital, configuring technical support services, building in-house help-desk tooling, and connecting *Cisco telephones* — Certificate of Appreciation.
- Managed device inventory, domain/hostname assignment, and system cleanup, hardening defenses with *firewalls*, *IDS/IPS*, *F5 BigIP*, *Kali Linux*, and *Python*-based automations.
- Administered *Active Directory* and shared folders, installing and configuring *switches and servers* across hospital departments.

EDUCATION

Bachelor's Degree in Information Technology — Machine Learning Track

Jazan University | Jazan, Saudi Arabia | GPA: 4.56 / 5 | Sep 2021 – May 2025

Studied deep learning and NLP with hands-on experiments building Arabic-language models for Saudi dialects.

Relevant Coursework: Machine Learning, Deep Learning, NLP, Data Structures & Algorithms, Databases, Software Engineering

PROJECTS

hse-digital (FERAS) — Digital HSE safety management platform for construction (PTW work permits, approval engine, Google Drive archiving, Google Apps Script) built for a real construction project, a coffee factory in Jazan.

MortakizAi — Local AI engine (Python, FastAPI, Qwen2.5-Coder, SGLang, DeepSeek V4 Flash, CUDA 13, 4xH100 94GB, OpenAI-compatible, adaptive GPU modes) for secure code assistance in Cline/Continue for VSCode and running Claude Code fully offline, with zero internet or login, no data leaving the machine.

Matcher — Multi-entity matching engine and RAG with vector database (Rust, Python, C++, 4 deep models, Milvus clustered vector database) matching faces, images, similar texts, and similar names across millions of records with millisecond search speed, fully local and offline.

SumAi — Offline Arabic summarization API (Python, FastAPI, ALLaM-7B, AraBert NER) with custom prompts and batch support for Arabic documents.

Maysour (Yusr) — Graduation Project — 1st Place, Jazan University; FinTech platform (Flutter, Firebase, Node.js) digitizing debt records and payment tracking for small businesses.

VOLUNTEERING & LEADERSHIP

Lead Coordinator, Science & Technology Community

Fada Tech | Saudi Arabia | 300 hours

- Organized events and technical support; wrote and published content on X (Twitter).
- Co-organized a Vehicle Cybersecurity course with the *Digital Attaa Initiative* as a technical organizer, helping participants install and run *Kali Linux*.

Founder & President

HRSD (HackerRank Student Developers) | Jazan Region, Saudi Arabia | 240 hours

- Founded and led the club; organized six-day AI camps on *Python, computer vision, and object detection*.
- Organized web development courses and HackerRank certification pathways for members.

Student Ambassador

Microsoft | Saudi Arabia | 200 hours

- Trained **3,000+ students per year** in AI-assisted development (*GitHub Copilot*).
- Delivered a *Metasploit* penetration testing course and an online LinkedIn profile-building workshop.

Student Developer

Google Developer Student Club (GDSC) | Jazan University | 120 hours

- Presented *Flutter* and AI courses.
- Organized club training sessions and workshops for students.

CERTIFICATIONS

Fundamentals of Artificial Intelligence (AI) — SDAIA

AWS Certified Cloud Practitioner — Amazon Web Services

Artificial Intelligence Fundamentals — IBM SkillsBuild

Software Engineering Essentials — IBM

Google Data Analytics Professional Certificate — Google

Junior Cybersecurity Analyst Career Path — Cisco

AWARDS & HONORS

Employee of the Year — MDS CS — Q4 2025 & Q1 2026 (consecutive)

1st Place — Graduation Project, Jazan University 2025

1st Place — Beban 24 (Riyadh), Best Project among Saudi Universities

1st Place — Monsha'at Hackathon 2024 & Jaddathon Fikrathon 2023

2nd Place — Bank Albilad Hackathon 2024

3rd Place — Mubtakiroon & Bank AlJazira FinTech 2026 & Accounting Hackathon, KSU 2024

SKILLS

Soft Skills: Team Leadership, Mentoring, Training Coordination, Public Speaking, Agile Environments, Communication, Collaboration, Analytical & Critical Thinking, Continuous Learning

Programming: Python, Java, C++, C#, Rust, TypeScript, JavaScript, Flutter

AI & ML: RAG, RAG Pipelines, Local LLMs, LLM Training, Fine-Tuning, MLOps, LLMOps, Model Deployment & Serving, Prompt Engineering, AI Agents, PyTorch, SGLang, Milvus, MCP

Backend & Data: Spring Boot, REST & SOAP APIs, Microservices, Vector Databases, PostgreSQL, DB2, MySQL, MongoDB, SQL / NoSQL

DevOps & Security: Docker, Kubernetes, Git, GitHub, Azure DevOps, CI/CD, Linux, CUDA, Playwright, WebSphere, Active Directory, Cybersecurity, Firewalls, IDS/IPS, Kali Linux

LANGUAGES

Arabic — Native | English — Fluent