



# ZAKARYA ALSALAH

AI & FULL-STACK ENGINEER

## / ALL ABOUT ME

AI & Full-Stack Engineer with 6+ years building and leading production AI systems — shipping LLM applications before ChatGPT existed. I architect the AI and ship the product around it: production document intelligence at 90%+ accuracy across 12+ document types, a live LangGraph RAG agent, and an MCP-based autonomous agent in build. An AI-native engineer who orchestrates coding agents and validates what ships. Led a team from 0 to 8. Native Arabic speaker based in Riyadh.

## / SOFT SKILLS

- > Leadership & Mentorship
- > Team Building & Hiring
- > Problem-Solving & Innovation
- > Communication & Initiative
- > Project Management

## / LANGUAGES

- > Arabic — Native
- > English — Fluent
- > Filipino — Fluent

## / CONTACT DETAILS

- 📍 AL-Malaz, Riyadh, Saudi Arabia
- ✉ zakaryaalsalahi@gmail.com
- 🌐 zalsalahi-portfolio.netlify.app
- ☎ +966 053-470-1120
- 🔄 github.com/hkzak · github.com/hkzak2
- 📄 linkedin.com/in/zakarya-alsalahi

## / WORK EXPERIENCE

### >> AI Engineering Lead

Confidential — *Singapore-based AI accounting company* | Jul 2023 – Present

- Own the company's **document-intelligence platform**: re-architected it from a legacy AWS Textract + rule-based pipeline to a **multi-LLM system**, lifting field-level extraction accuracy from **~50% to 90%+** across **12+ document types** (invoices, bank & credit statements, loans, rentals) in multiple languages and formats — PDF, Excel, images, typed & handwritten.
- **Led the development of a production RAG agent (LangGraph)** over hundreds of documents — mentoring the team through build, review, and deployment; OpenAI text-embedding-3-small + Supabase/pgvector, tuned chunking (512/20), top-5 semantic retrieval feeding a multi-step agent.
- Building an **MCP server** (Model Context Protocol) that exposes accounting endpoints as agent tools, plus an **autonomous accounting agent** on top — "Claude Code, for accounting."
- Architected a **multi-provider LLM layer** (OpenAI, Claude, Gemini) with provider benchmarking on accuracy/cost/latency; fine-tuned an open-source model (Qwen2.5, LLaMA-Factory) to evaluate self-hosting — a data-backed build-vs-buy decision.
- Built and scaled the AI & engineering team **from 0 to 8** (currently lead 6); **founded the internship program** that grew fresh graduates into engineers who now lead their own projects.
- Delivered, with the team: the document-intelligence platform (web + Android), a **tax-filing module**, a CRM, an HR app, and led the corporate-site rebuild (WordPress → enterprise web app); rescued **3 failing apps** to production. Stacks: Laravel, Node.js, React, React Native.

### >> Co-Founder

Raqmiyoon — *community tech & AI initiative* | May 2026 – Present

- Founded a community initiative (registering as a non-profit) helping Yemeni professionals and students adopt modern technology and AI.
- Built the website, formed the founding team, and run community events.

### >> Founding Engineer & Technical Lead (Part-Time)

Global Staffing — *staffing startup* | May 2025 – Mar 2026

- Co-founded the company; built the **CRM and landing page** and ran the tech & HR functions — client discovery, hiring pipeline, training, and an internship program.
- Delivered **AI applications for 3 clients**: AI **legal-document drafting** for non-legal professionals; **healthcare AI** that summarizes patient charts, auto-fills documents, and **flags critical conditions** for visiting nurses; and an AI-powered **"school OS"** (LMS modernization). Shipped one to production, one to staging.

### >> Back End Developer

Offshore Team | Jul 2022 – Jul 2023

- Built **Java / Spring Boot microservices** to enterprise standards; CI/CD with Docker, Jenkins, Linux; centralized configuration with config server + Vault.
- Created and deployed **mock-server containers** for cost-effective testing; built a PDF text-extraction chatbot POC (Haystack + open-source NLP). **Proposed and demoed a RAG system** ahead of client appetite.





## / TECHNICAL SKILLS

### AI / GENAI & AGENTS

- LLM APIs — OpenAI · Claude · Gemini
- RAG — embeddings, pgvector/Supabase
- Prompt Engineering
- LangChain / LangGraph
- AI Agents — CrewAI · Claude SDK · Google ADK
- Model Context Protocol (MCP)
- Fine-tuning — LLaMA-Factory, SFT
- Hugging Face · Haystack · FAISS
- AI-assisted coding — Claude Code, Copilot, Cursor

### LANGUAGES

- Python
- JavaScript / TypeScript
- Java
- PHP
- C++

### BACKEND & FRONTEND

- FastAPI · Django · Node/Express
- Spring Boot
- Laravel
- React
- React Native · Angular · Vue

### CLOUD & DATA

- Docker · CI/CD
- AWS · Azure · IBM Cloud
- PostgreSQL/pgvector · Supabase · MongoDB

## / CERTIFICATIONS

- Fihm.ai — AI Course (Completion + Attendance)
- DeepLearning.AI — Coursework

## / REFERENCES

Available upon request

## / WORK EXPERIENCE (CONTINUED)

### >> Full-Stack & AI Developer

DXC Technology | Jun 2019 – Jul 2022

- **Proposed and co-initiated the team's first generative-AI work.** Built 7+ LLM/NLP POCs with both **open-source and private LLMs — before ChatGPT existed** — with at least one reaching production: a **RAG chatbot** turned into a custom production system for **a major US airline**.
- Built a **generative document-extraction Q&A chatbot** using a **FAISS document store** to store and retrieve embeddings — answering in plain language while citing source paragraphs.
- Built an award-winning **OCR Kanban-board app** (photo → post-it detection → Google Cloud OCR → digital board); recognized on the company **Wall of Fame**.
- Built a **natural-language code interpreter** (→ pandas) letting non-technical managers query Excel in plain English; plus a sales-forecasting app.
- Modernized a major US airline's **mainframe applications to Java/Angular**; built APIs across Node.js, Java/Spring Boot, and FastAPI.

### >> College Lecturer (Part-Time)

Centro Escolar University (CEU), Manila | Aug 2023 – Aug 2025

- Taught **Object-Oriented Programming, Intelligent Systems, and Engineering Economics**.

## / EDUCATION

### >> Centro Escolar University (Manila) | Philippines

BS Computer Engineering | 2018

- > Graduated top of class — 1.82 weighted average (PH scale, 1.00 = highest)
- > Best Thesis: Intelligent Shopping Cart
- > PIFUA's 48th Most Outstanding International Student in the Philippines (2018)
- > CEU Ambassador / Mr. GoodWill (2017)

## / KEY PROJECTS

### >> Enterprise Document-Intelligence Platform

Confidential (current role) | 2023 – Present

Multi-LLM extraction pipeline: ~50% → 90%+ accuracy, 12+ document types, multi-language, multi-format. Production **LangGraph RAG agent** over hundreds of documents.

### >> AI Book Reader

Personal project | 2026

Upload a **PDF book** and read it in a dual reading + chat panel. Hit play and an **AI narrator reads it aloud page by page**, auto-advancing through the book. **Ask questions** about the text and take notes inline, with a **customizable reading persona** (tone, accent, and language).

### >> Medical Lab Billing System

Personal project | 2025

A **billing system for a medical laboratory** — managing patients, test orders, and invoicing through to payment.

### >> OCR Kanban Board

DXC Technology | 2021

**Award-winning app** that turns a **photo of a physical board into a digital one**: post-it detection → **Google Cloud OCR** → live Kanban board. Recognized on the company **Wall of Fame**.

### >> Healthcare Charting AI

Freelance / Staffing client | 2025

AI that **summarizes patient charts**, auto-fills documents, and **flags critical conditions** for visiting nurses to review.

### >> NL Code Interpreter for Excel

DXC Technology | 2022

App letting non-technical users query large Excel files in **plain English** (→ pandas), returning charts, tables, and answers. Streamlit + FastAPI.

