

MOHAMMED GHALIB

AI / RAG Engineer • Full-Stack Engineer (Python / Django • React / TypeScript)

Nairobi, Kenya • +254 758 299 089 • ghalibmohammad92@gmail.com

[Portfolio](#) • [LinkedIn](#) • [GitHub](#)

PROFESSIONAL SUMMARY

Full-stack software engineer with 3+ years shipping production systems end-to-end — Python / Django REST on the backend, React / TypeScript on the frontend, containerized with PostgreSQL, Redis, Celery, JWT auth, and CI/CD pipelines.

I specialize in production AI: retrieval-augmented (RAG) systems built to be trustworthy, not just impressive — similarity-thresholded retrieval, reranking, refusal on low confidence, and source-grounded citations. Built and shipped in Sheria AI (Kenyan-law assistant) and Ummly (Islamic knowledge platform).

Experience spans AI/RAG engineering at Datawise Africa, full-stack and AI product work with Digital Appex, and enterprise systems at a national bank (KCB) and across private clients — always with end-to-end ownership: architecture, backend, frontend, deployment.

CORE SKILLS

Languages & Frameworks: Python, TypeScript/JavaScript, Django, Django REST Framework, FastAPI, Celery, React 19, React Router 7 (SSR), React Native (Expo), Tailwind CSS, Redux Toolkit, TanStack Query

AI / RAG Engineering: RAG pipeline design, LangChain, vector databases (Pinecone, ChromaDB), Sentence-Transformers (all-MiniLM-L6-v2), reranking (bge-reranker-v2-m3), semantic retrieval with similarity thresholds, grounding & citation, hallucination reduction, prompt engineering, intent routing, eval/regression suites

LLMs: OpenAI API (GPT-4 family), LLaMA 3.1 (via Groq), Google Gemini 2.5, DeepSeek, multi-provider fallback orchestration

APIs & Integrations: REST APIs, JWT (RS256 / SimpleJWT), WebSockets/Channels, Paystack, M-Pesa STK Push, PayPal, Stripe, ZeptoMail, SMS/email gateways, KRA iTax

Data & Infra: PostgreSQL, Redis, SQLite, vector stores; Docker & docker-compose, Gunicorn, Nginx (nginx-proxy + Let's Encrypt), DigitalOcean, Vercel, Netlify, Render, Supabase

DevOps & Quality: GitHub Actions CI/CD, SonarQube quality gates, pytest + coverage, Vitest, semantic-version releases, staging/prod environments, Git, Agile/Scrum, Jira, Slack

Enterprise: Oracle Fusion HCM/Financials, RBAC, user provisioning, security policies

PROFESSIONAL EXPERIENCE

Software Engineer — Datawise Africa

2026 • Remote

In charge of Sheria AI end-to-end — an AI legal assistant for Kenyan law (RAG backend + SSR frontend).

- Built the **Sheria AI** (sheria.africa) **RAG backend** in Django 5.2 / DRF — 32 REST endpoints across 10 data models — with a two-step **triage → answer** pipeline: a structured classifier routes each message, then a grounded generator answers only in-scope legal questions. LLaMA-3.1 via Groq + LangChain; Sentence-Transformers (all-MiniLM-L6-v2) embeddings.
- Engineered the hallucination controls that are the product's core: **dual-corpus retrieval** over ChromaDB (Kenyan case law + the Constitution) with calibrated cosine-distance cutoffs (0.58 cases / 0.60 constitution) that drop off-topic hits so the model answers from general knowledge rather than forcing irrelevant context; exact-name case lookup; relevance-scored **source citations**; and a strict “never fabricate a case, citation, or section” system prompt with graceful refusal/redirect for out-of-scope and gibberish input.
- Replaced brittle threshold unit tests with a **prompt-driven evaluation suite** (triage scorecard + tagged regression cases) that locks routing and reply-language behavior against real production failures.
- Owned the Sheria AI **frontend** end-to-end — React Router 7 (SSR), React 19, TypeScript, Redux Toolkit, TanStack Query: streaming chat UI, markdown answers with relevance-ranked source cards, and a Vitest test suite.
- Set up the delivery pipeline on **GitHub Actions**: SonarQube quality gate, pytest + coverage, automated PR notifications to Slack, and semantic-version release automation; Gunicorn / PostgreSQL / Redis / Celery runtime.

Software Engineer — Digital Appex

2025 • Remote

Kenyan technology company — full-stack and AI product engineering.

- Built and shipped digitalappex.co.ke end-to-end: a Django 6.0 / DRF backend of **50 REST endpoints across 11 apps** (products, orders, quotes, payments, warranties, messaging, projects, contact, chatbot) and a React 19 + Vite + TanStack Query frontend.

- Implemented JWT auth (SimpleJWT), **Paystack** payment integration, role-based access, and Celery + Redis background tasks.
- Containerized with **Docker + docker-compose** (web / PostgreSQL / Redis); a Gunicorn entrypoint runs collectstatic → migrate → serve. GitHub Actions CI checks for migration drift and runs the test suite; deployed to a **DigitalOcean droplet behind nginx-proxy with automated Let's Encrypt SSL**.
- Define the engineering standards, documentation, and release process for the company.

Financial Systems Intern — Kenya Commercial Bank (KCB)

Jan 2025 – Jul 2025 • On-site, Nairobi

Production support for Oracle Fusion at one of Kenya's largest banks.

- Provided daily production support for Oracle Fusion HCM and Financials serving 1,000+ staff.
- Created and governed user roles, access restrictions, and security policies, enforcing internal controls and audit requirements (RBAC).
- Designed and exposed secure REST API endpoints integrating Oracle Fusion with internal services and external systems, including KRA iTax and payment gateways.
- Resolved tickets within SLA and coordinated fixes across HR, Finance, and IT stakeholders.

Full-Stack Developer Intern — Bee MultiScent Ltd

Feb 2023 – Dec 2023 • Hybrid

Company-wide Management Information System (MIS).

- Co-built an MIS used across operations, HR, and inventory — Django / DRF backend, React.js frontend with real-time dashboards.
- Delivered 30+ REST endpoints consumed by the company's React Native mobile app; added RBAC, task assignment, document management, and WebSocket notifications via Django Channels.
- Earned an official recommendation letter for performance and teamwork (available on request).

Independent Software Engineer (Contract) — Private Clients

Jan 2024 – Aug 2025 • Remote

End-to-end full-stack platforms designed, built, and deployed for clients.

- **Ummly (formerly Islamify) — Islamic knowledge platform** (ummlly.co.ke). Built the full product solo across **backend, web, and mobile**: a Django / DRF backend of 8 apps, 34 data models, and 93 migrations with a **Pinecone RAG** assistant grounded in the Qur'an and Hadith — server-side llama-text-embed-v2 embeddings, **bge-reranker-v2-m3 reranking**, a confidence floor that returns a low-confidence message instead of guessing when no source clears the threshold, intent classification, query rewriting, bilingual EN/AR answers, and per-excerpt citations. LLaMA-3.1 via Groq with Google Gemini fallback. Containerized (Gunicorn / PostgreSQL 16 / Redis 7 / Celery) and deployed via GitHub Actions to a DigitalOcean droplet (nginx-proxy + Let's Encrypt); ZeptoMail email and Paystack payments. Companion web app in React Router 7 (SSR) on Vercel and an Expo / React Native mobile app.
- **Collaborative Approach — referral & tour-booking SaaS** (collaborativeapproach.africa). React + DRF; integrated M-Pesa STK Push, PayPal, and Stripe; referral tracking with automated commission calculation, JWT auth, Celery background tasks, and role-based portals (Admin / Agent / Customer).
- **Sheikh Siraj Rahman College — school management system** (sheikhsirajrahman.ac.ke). Multi-portal (Admin / Teacher / Student / Parent) on React, Django, and PostgreSQL — online admissions, fee payments, results publishing, and SMS/email automation; deployed on HostPinnacle with a CI/CD pipeline.

EDUCATION

Bachelor of Business Information Technology — Mount Kenya University

2022 – 2025

Major in Information Technology, Software Development, and Artificial Intelligence.

CERTIFICATIONS

- Build AI-Powered Apps — Code with Mosh
- Ultimate Django Series — Code with Mosh
- Ultimate Docker Course — Code with Mosh
- Ultimate React + TypeScript Course — Code with Mosh

REFERENCES

Available on request.