

Aqib Ali Ghumro

AI Engineer | Backend Systems

aqibali280@gmail.com | +92 324 2920658 | Karachi, Pakistan | LinkedIn | GitHub

Professional Summary

AI Engineer with a software engineering background and 1+ year of professional software experience. Worked with local and international businesses to understand manual workflows and take them from requirements to implementation using software, automation, and AI. Built RAG assistants, voice agents connected to booking APIs, multi-agent workflows, and AI video pipelines using Python, FastAPI, OpenAI, LangChain/LangGraph, Vapi, PostgreSQL, and Docker, with a focus on reliable tool calling, validation, and backend integration.

Experience

Software Engineer

Jul 2025 – Present

Touchperia

Karachi, Pakistan

- Build and maintain backend APIs and product workflows across Makaanza, HairZone, and an internal HR platform, covering authentication, bookings, payments, notifications, and third-party integrations.
- Built core HR workflows for employee records, attendance, payroll, leave, documents, and onboarding; set up GitLab CI/CD and supported AWS deployment troubleshooting.
- Write unit, integration, API, and Playwright end-to-end tests; work with product stakeholders on requirements, estimates, implementation, verification, and releases.

Projects

AI Call Center | *Vapi, Node.js, REST APIs, OpenAI, Tool Calling*

- Built voice-agent workflows for availability checks, booking, rescheduling, cancellations, human handoff, transcripts, and contextual call summaries.
- Integrated Vapi webhooks and function/tool calls with validated REST APIs, separating conversational intent from backend execution and failure handling.

AI Document Assistant | *FastAPI, Next.js, LangChain, Chroma, PostgreSQL, Docker*

- Built a RAG application for PDF, DOCX, TXT, and Markdown ingestion with embeddings, metadata filtering, grounded answers, and citations tied to retrieved source passages.
- Added organization-scoped authentication, conversation history, streaming SSE responses, and grounding checks to reduce unsupported answers.

Multi-Agent Operations Platform | *FastAPI, LangGraph, LangChain, PostgreSQL, Next.js*

- Designed routed support, order-management, and analytics agents with shared state, structured tool calls, retries, and audit logging.

Channel-IQ | *Python, Django, Whisper, OpenAI, YOLO, OpenCV, FFmpeg*

GitHub

- Built a pipeline that downloads long-form video, transcribes audio, selects highlights, generates short clips, applies face-aware framing and captions, and prepares SEO metadata for YouTube publishing.
- Reduced roughly one-hour video turnaround from 3–4 hours of manual editing to under 25 minutes on a GPU VM.

Technical Skills

Languages: Python, TypeScript, JavaScript, SQL

AI/LLM: RAG, LangChain, LangGraph, AI Agents, OpenAI API, Vapi, Prompt Engineering, Tool/Function Calling

AI/CV: Chroma, Embeddings, Whisper, YOLO, OpenCV, FFmpeg

Backend: FastAPI, Node.js, Express.js, REST APIs, PostgreSQL, Prisma, JWT

Delivery/Testing: Docker, AWS, GitLab CI/CD, Unit, Integration, API, Playwright E2E

Education

NED University of Engineering & Technology

Bachelor of Software Engineering

2021 – 2025

Karachi, Pakistan