

Muhammad Arslan Tabish

Full Stack AI Engineer · Claude API · Agentic Pipelines · Python / FastAPI · React / TypeScript
+92 300 0835300 · arslantabish2001@gmail.com · muhammad-arslan-tabish · github.com/Arslan9272
Portfolio · arslantabish.com

Nationality Pakistani · **Location** Lahore, Pakistan · **Relocation** Open to relocation and remote · **Notice** 30 days

Profile

Full Stack AI Engineer with 2+ years of professional experience building LLM-powered products end to end — Python/FastAPI services and data models through to the React 19 / TypeScript interfaces I design and ship myself. Focused on **AI systems, not AI demos**: resumable multi-stage agentic pipelines, schema-constrained output with validation and retry/repair, grounded generation and RAG, guardrails and automated QA gates, model routing and per-call cost accounting on the **Anthropic Claude API**. Owns features from PostgreSQL schema to browser and into production on Docker, Vercel and Railway.

Technical Skills

Generative AI & LLMs: Anthropic Claude API, OpenAI API, LLM integration, prompt engineering & versioning, structured outputs, context engineering, model routing, prompt caching, token & cost optimisation

Agentic AI: agentic workflows, multi-stage AI pipelines, tool / function calling, LangGraph, LangChain, MCP, agent memory, multi-agent orchestration, human-in-the-loop

RAG & Retrieval: RAG, embeddings, vector search, vector databases, semantic search, hybrid retrieval, reranking, chunking & indexing, grounded generation

AI Reliability & Evaluation: schema validation & retry, output guardrails, automated QA gates, eval harnesses, retrieval evaluation, LLM observability, prompt / model evaluation, failure analysis

Computer Vision: OpenCV, YOLO, object detection, object tracking, real-time video processing

AI Integrations & Automation: Apify, Aircscale, MillionVerifier, Instantly, third-party APIs, webhooks, lead-enrichment pipelines, automated outreach workflows, data extraction and processing

Backend: Python 3.12+, FastAPI, async/await, asyncio, Pydantic v2, SQLAlchemy 2.0 async, PostgreSQL, WebSockets, background jobs, REST APIs, API authentication

Frontend: React 18/19, TypeScript, Next.js, Vite, TanStack Start / Router / Query, Tailwind CSS, Radix UI / shadcn, Framer Motion, React Three Fiber, React Hook Form, Zod

DevOps & Production: Docker, Vercel, Railway, AWS S3, Cloudflare, GitHub / GitLab, CI/CD, production monitoring, pytest

Professional Experience

Intagleo Systems

Jul 2024 – Present

Software Engineer (Full Stack / Python)

Lahore, Pakistan

- Build **LLM-powered product features on the Anthropic Claude API** — agentic workflows, RAG pipelines and prompt-engineered integrations — with structured outputs, schema validation and deterministic fallbacks so AI features fail safely in production.
- Ship secure, scalable **REST APIs in Python FastAPI** — authentication, Pydantic validation, background jobs and business logic — across several production applications, plus Django/DRF services with role-based permissions.
- Design and optimise data across **PostgreSQL**, MySQL and MongoDB — schema design, efficient queries, migrations and data integrity across systems.
- Deliver full-stack features with **React / Next.js** front ends against FastAPI services, and real-time **WebSocket** features for live chat and notifications.
- Automate internal engineering workflows (task routing, approvals, compliance reporting) in Python, and deliver in **Agile** teams with Git/GitHub code review, sprint planning and weekly client meetings.

Wisdom Software House

Jun 2023 – Aug 2023

Software Developer Intern

Lahore, Pakistan

- Built full-stack web applications with a **Python FastAPI** backend and optimised **MySQL** queries for performance.

Devsloop

Dec 2022 – Jan 2023

Software Developer Intern

Lahore, Pakistan

- Web development fundamentals, version control, debugging and team workflows.

Education & Certifications

University of the Punjab, Lahore — BSc Computer Science (BSCS)

2020 – 2024

Certifications: Udemy — Python for FastAPI: Advanced Backend Development · Python for Computer Vision & Deep Learning

Awards: Battle Byte Programming Competition — Runner-Up (university level)

Languages: English (fluent) · Urdu (native) · Punjabi (native) · **Driving Licence**: Pakistan (LTV)

References available on request

Projects

AI Systems — self-built products · full write-ups at arslantabish.com/#projects

Cold Outreach Pipeline — AI Lead-Generation SaaS [Live](#) · [Code](#)

Python, FastAPI, Claude, PostgreSQL, Stripe

- Designed a **9-stage, status-driven, resumable AI pipeline** (extract, enrich, AI filter, gate, email, verify, AI copy, QA, sync) where every lead advances only on a valid stage transition, so runs crash and resume without duplicating work or LLM spend.
- Built a defensive **structured-output layer** for all AI calls — fence stripping, JSON recovery, schema validation and retry with a corrective nudge before a typed `ParseError` — plus deterministic guardrails and an **automated QA gate** that blocks weak copy before it reaches a prospect.
- Integrated **Apify, Aircscale, MillionVerifier and Instantly** into a single enrichment-and-outreach flow with webhook-driven status sync and per-provider fallbacks.
- Implemented multi-tenant isolation, BYOK keys encrypted at rest (AES-256-GCM), Stripe billing with plan limits, **per-call token-cost attribution**, and a zero-spend demo mode using a mock LLM behind a Python Protocol.

Dukandar AI — Voice-First Conversational Shopping Agent [Code](#)

Python, FastAPI, Claude, SQLAlchemy, React

- Built a **catalogue-aware conversational agent** that walks shoppers through a facet-driven dialogue (category → kind → colour → size → price) and mirrors its filters live onto the storefront grid.
- Constrained Claude slot extraction to facet values that exist in the live catalogue, **eliminating hallucinated and out-of-stock suggestions**, with a deterministic keyword parser as a \$0 fallback on missing key, timeout or malformed JSON — the conversation never breaks.
- Built the session engine, faceted search, a TTS proxy for voice output and the storefront front end on async **SQLAlchemy 2.0 + PostgreSQL**.

GoldWatch — AI Gold (XAUUSD) Trading Analyst [private repo](#) — [walkthrough on request](#)

Python, Claude, market-data APIs

- Built a market-analysis assistant that computes market structure and indicators, renders charts, gathers macro context and asks Claude for a read — with one hard safety rail: it never places an order, it prints a full briefing (entry, stop, target, lot size, money at risk) and waits for the human.
- Architected the system to **resist LLM over-eagerness**: a deterministic scanner scores setups first and an API call is spent only above a configurable threshold, keeping cost near zero on quiet days.

Full-Stack & Frontend — client delivery

Khaayaal — Elder-Care Platform [Live](#) · [Code](#)

TanStack Start, React 19, TypeScript, Tailwind v4

- Designed and shipped the complete six-route marketing site on **TanStack Start + Router** with React 19 and Tailwind CSS v4, on a custom design-token system, Radix UI primitives and a reusable brand component library.
- Implemented an animated **3D hero with React Three Fiber**, scene-based scroll animation and an interactive pricing calculator over tiered care plans; React Hook Form + Zod, deployed on Vercel.

Leads In The Pipe — B2B Agency Website [Live](#) · [Code](#)

React 18, TypeScript, Tailwind, Framer Motion, FastAPI

- Translated a Figma design into a production single-page front end on Vite + React 18 + TypeScript (strict) with Tailwind design tokens and Framer Motion, backed by a Python FastAPI serverless API for form handling on Vercel.
- Enforced quality with Vitest + React Testing Library + vitest-axe, a **WCAG 2.1 AA** pass and SEO work (meta/OG, JSON-LD, Core Web Vitals).

Professional Projects — Intagleo Systems (client and product names withheld; internal systems, not public)

Internal Engineering Operations Portal

FastAPI, PostgreSQL, GitHub API

- Built the full **FastAPI backend** for project tracking, developer task management and automated workflow processing — multi-department access with role-based permissions, real-time status updates, an automated code-review routing and approval workflow with GitHub integration, and Excel compliance reporting.

Admin Panel Backend — enterprise client

Django, Django REST Framework

- Built secure REST APIs with **Django and DRF** (ORM, CBVs and FBVs) for complex business logic and workflow automation, with role-based access control, file uploads and automated email; wrote unit tests, optimised ORM queries and ran weekly client meetings tracked in GitLab.

Multi-Database Service Platform

FastAPI, MongoDB, MySQL, AWS S3

- Developed a layered FastAPI service (routes / controllers / services / repositories) over MongoDB (Motor) and MySQL (aiomysql), with **TOTP two-factor authentication** (pyotp + QR provisioning), transactional email, AWS S3 file handling via boto3 and structured logging.

Digital Signage Layout & Live-Preview Tooling

Node.js, Express, WebSocket, HTML/CSS

- Contributed to a layout and live-preview workspace for data-driven menu boards deployed by **national quick-service restaurant chains** — an Express + WebSocket dev server hot-reloading the board in an iframe preview, with data-bound 1920×1080 / 1080×1920 layouts joining HTML/CSS to spreadsheet product data.