

Faique Ahmad Khan

+923336118864 | faiqueahmadkhan@gmail.com | [Linkedin](#) | [Github](#) | <https://faique.dev>

Summary

Full Stack Engineer specializing in AI-integrated products, with 3+ years building features used by 200K+ users. At Vyro AI, shipped an AI search engine that drove a 35% revenue lift and integrated 27+ LLMs into a production chat platform handling real-time streaming at scale. Strong in React/Next.js, Node.js, and LLM orchestration. Open to remote roles where product impact is measurable.

Skills

Core: TypeScript, React.js, Next.js, Node.js, PostgreSQL

AI & LLMs: Claude API, OpenAI API, Gemini API, LLM orchestration, RAG pipelines

Infrastructure: AWS, Vercel, CI/CD, Git

Familiar with: MongoDB, Firebase, Mapbox, Nest.js, Prisma

Experience

Software Engineer | Vyro AI

Islamabad, Pakistan | 12/2024 - Present

- Built **AI Docs** from the ground up, a production-grade rich-text editor with deeply integrated Claude API, supporting streaming inline completions, context-aware content transformation, and multi-block editing. Designed the document state model to handle concurrent AI edits without corrupting user content.
- Shipped an **AI Search Engine** with real-time web retrieval, semantic query handling, and fallback ranking logic. The feature became the product's top conversion driver and contributed to a **35% lift in subscription revenue** within 3 months of launch.
- Architected a **Strapi-powered CDN pipeline** to dynamically generate SEO landing pages, reducing manual marketing page creation overhead.
- Architected the **LLM orchestration layer** for Chatly AI (200K+ users), designing a unified abstraction over 27+ models including GPT, Claude, Gemini, and DeepSeek. Built dynamic model routing based on cost, latency, and capability tiers, with fallback logic that reduced failed completions by 20% and kept **p95 response time under 8 seconds**.

Full Stack Developer | Geokits — *product agency, client delivery role*

Remote | 04/2024 - 12/2024

- **Led a 3-person engineering team** (including 2 junior engineers) on a GIS platform from architecture to production delivery. Introduced PR review standards and a branching workflow that cut integration conflicts by roughly half and allowed the team to ship on a fixed client deadline with no scope cuts.
- Architected a full-stack **GIS platform** from schema design to production, using **Next.js, Node.js, PostgreSQL** for geospatial queries, and Firebase for real-time sync. Designed the event ingestion pipeline to handle **10K+ concurrent location events** with sub-second UI updates, optimizing PostgreSQL spatial indexes to keep query latency under 100ms at load.
- Built and maintained **CI/CD pipelines** across Vercel and AWS, cutting deployment friction and ensuring zero-downtime releases for production environments.

Associate Software Engineer | Developers Den LLC

Islamabad, Pakistan | 06/2023 - 06/2024

- Owned **end-to-end frontend delivery** across **12+ client projects** at a software agency, driving architecture decisions, component design, and production deployments independently with minimal oversight.
- Introduced **TypeScript** strictly and established reusable component patterns using **Redux and Zustand** that were adopted across the team, reducing redundant code and cutting estimated feature delivery time on repeat project types by ~30%.
- Consistently delivered pixel-perfect **Figma-to-production implementations** across diverse client stacks (**Tailwind, Material UI, Radix, Ant Design**), becoming the team's go-to engineer for high-fidelity UI work on tight deadlines.

Projects

Guardian Space — *Early Warning System for Natural Disasters* · Supervised by the **European Space Agency**

- A real-time early warning platform for natural disasters, developed under direct supervision of the European Space Agency. Aggregated live seismic, wildfire, and weather data from **USGS and third-party feeds** into PostgreSQL, with a global **Mapbox** interface enabling rescue agencies to visualize and respond to active events.
- Built inter-agency coordination via **Firebase Cloud Messaging** and an OpenAI-powered in-app chatbot for contextual disaster guidance.

Chatly AI

- Architected the **LLM orchestration layer** powering **200K+ users**, building a unified abstraction over **27+ models** with dynamic routing across cost, latency, and capability tiers, keeping **p95 response time under 8 seconds** at production scale.
- Engineered the **AI Search Engine** feature with real-time web retrieval and semantic query handling, which became the platform's top conversion driver and contributed to a **35% lift in subscription revenue** within 3 months.
- Built **AI Docs**, a production-grade rich-text editor with streaming Claude API integration, handling concurrent AI edits through a custom document state model that prevented content corruption under simultaneous user and AI mutations.

Education

Software Engineering

NUST, SEECS — May 2024

Languages & Certifications

English — C1 (IELTS 7.5)