+91 6307674245
Bengaluru, India
krish24090@gmail.com

Krishna Dwivedi
Full Stack Developer

GitHub
LinkedIn

Professional Summary

Full-stack developer and GenAI engineer with hands-on experience building production-grade web applications and AI systems, including MERN-based platforms, retrieval-augmented generation (RAG) pipelines, agentic workflows, and voice-enabled assistants. Strong background in backend engineering, authentication, and modern LLM frameworks (LangChain, LangGraph, Gemini, OpenAI), with a proven ability to design controlled, scalable, and user-focused AI products.

SKILLS

Languages	Python, JavaScript
Frontend	React, Streamlit, Form Validation (Yup)
Backend & APIs	Node.js, Express.js, FastAPI, RESTful APIs, express-validator, JWT-based Auth, Role-Based Access Control (RBAC), Media Handling (Cloudinary)
Databases	MongoDB, PostgreSQL
AI / GenAI	LangChain, LangGraph, Langsmith, Retrieval-Augmented Generation (RAG), AI agents (Checkpointing, Human in the Loop), Conversational AI Interfaces, Integrated models with Qdrant (Vector DB) and Neo-4J(Graph DB), OpenAI/Gemini APIs, Guardrails, Fine Tuning
DevOps & Tools	Git, GitHub, Docker, Environment Config, Secrets Management

PROJECTS

Tandem — Full Stack Project (MERN Stack)
tandem.builders
Backend GitHub
Frontend GitHub

- Architected a **project orchestration platform** featuring multi-tenant organizational structures, complex hierarchical task scheduling, and real-time state synchronization.
- Engineered a custom **stateless authentication framework** using **JWT-based rotating tokens**, implementing secure session persistence and granular authorization middleware.
- Developed a scalable **microservices-oriented backend** with **Node.js and MongoDB**, enforcing strict data integrity via **schema-level validation** and request sanitization layers.
- Implemented **hierarchical data modeling** for recursive task-subtask relationships, optimizing query performance for complex relational workflows.
- Integrated **Cloudinary** for high-availability binary object storage and optimized multi-part asset uploads.
- Developed automated **transactional messaging pipelines** via **Mailtrap-integration** for asynchronous account verification and lifecycle events.
- Built a high-performance frontend using **React**, implementing centralized **global state management** and client-side validation logic for low-latency user interactions.

Chat PDF AI — RAG-based Document Q&A System
Demo Video
Backend GitHub Link

- Designed and built a **Retrieval-Augmented Generation (RAG)** application enabling users to upload PDFs and query them using natural language.
- Developed a responsive **React-based frontend** to provide a seamless document upload and interactive question-answering experience.
- Implemented a high-performance backend using **FastAPI**, handling document ingestion, embedding generation, and query orchestration.
- Processed and chunked PDF documents using **RecursiveCharacterTextSplitter (LangChain)** to optimize semantic retrieval accuracy.
- Generated vector embeddings using **OpenAI Embeddings** and stored them in **Qdrant**, enabling efficient similarity search over large documents.
- Implemented semantic search to retrieve the most relevant document chunks based on user queries.
- Integrated **Gemini-2.5-Flash** as the language model to generate precise, context-aware responses using retrieved document context.
- Designed an end-to-end RAG pipeline emphasizing **scalability, low latency, and modular architecture**.

Autonomous Speech-Interface Agentic Environment
Demo Video
GitHub Link

- Engineered an **audio-driven terminal AI** leveraging real-time signal processing and low-latency synthesized speech for hands-free system interaction.
- Architected a **state-driven agentic workflow using LangGraph**, implementing cyclical graph logic with conditional tool execution and persistent memory.
- Integrated **Speech-to-Text (STT) processing** via Google Cloud APIs to transform natural language audio signals into structured model inputs.
- Optimized **Audio Stream Synthesis** using OpenAI streaming APIs, reducing latency for high-frequency interaction loops.
- Developed a **tool-augmented LLM orchestration layer** using **LangChain** for autonomous file system operations and directory traversals.
- Designed custom modular tools for **system-level command execution**, managed through LangGraph's ToolNode and adaptive routing.
- Implemented **typed state management** and message aggregation to handle complex, multi-turn reasoning and context retention.
- Utilized **Asynchronous I/O (asyncio)** to parallelize audio signal processing, LLM inference, and multi-threaded tool execution.

Persona-Based Coding Chatbot
Deployed App Link
GitHub Link

- Built a **persona-driven chatbot** using **Streamlit** and **Google Gemini (gemini-2.5-flash)** to answer **coding and IT career questions** in a controlled, instructor-style conversational tone.
- Engineered **prompt-based guardrails** to enforce domain restriction, maintain persona consistency, and politely decline non-technical queries across multi-turn conversations.

EDUCATION

B.Tech in Civil Engineering (CGPA: 8.37), *Guru Gobind Singh Indraprastha University* 2019 – 2023
Senior Secondary (XII), Rani Laxmi Bai Memorial School (91.8%) 2016 – 2018
Secondary (X), St. Thomas' Inter College (92.3%) 2016

ACHIEVEMENTS

- Secured **AIR 1420 (Top 1%)** in **GATE 2024 — Civil Engineering**; received an **admission offer from IIT Bombay**.