

BHUPESH GAJBHIYE

GENAI & AGENTIC AI DEVELOPER | PYTHON & RAG ENGINEER

Nagpur, Maharashtra, India | +91-8459176208 | bhupeshgajbhiye88@gmail.com | linkedin.com/in/bhupesh-gajbhiye | github.com/Bhupeshgajbhiye

PROFESSIONAL SUMMARY

Innovative GenAI & Agentic AI Developer with over 3 years of experience architecting autonomous AI pipelines, multi-agent orchestration systems, and enterprise Retrieval-Augmented Generation (RAG) applications. Demonstrated expertise in leveraging Python, FastAPI, and leading LLMs (GPT-4, Gemini, LLaMA, Claude, DeepSeek) to automate complex workflows and build context-aware AI tools. Proficient in cutting-edge agentic frameworks including AutoGen, CrewAI, LangChain, LangGraph, and Model Context Protocol (MCP) to develop resilient, self-correcting agent architectures.

TECHNICAL SKILLS

- Generative & Agentic AI:** Agentic AI Workflows, AutoGen Framework, CrewAI, LangChain, LangGraph, Model Context Protocol (MCP), Autonomous Multi-Agent Systems, Guardrails, Function Calling
- LLMs & Frameworks:** GPT-4, Google Gemini, LLaMA 3, Claude 3.5 Sonnet, DeepSeek-R1, OpenAI API, Groq API, Hugging Face Transformers, Agno, OpenClaw
- RAG & Vector Databases:** Retrieval-Augmented Generation (RAG), FAISS, ChromaDB, Context Chunking Strategies, Vector Embeddings, Semantic Search, Re-ranking
- Programming & Backend:** Python (Advanced), FastAPI, Flask, SQL, PostgreSQL, RESTful API Architecture, JSON/Pydantic Parsing, Streamlit, Gradio
- Tools & DevOps:** Git, GitHub, VS Code, Cline (AI Agent), Docker, Linux/Unix Shell Scripting, Jupyter Notebook, Postman, CI/CD Fundamentals

PROFESSIONAL EXPERIENCE

Software Engineer – AI & Automation (NOC/SOC) | Phoneme Solutions Pvt. Ltd, Pune, India (Full-time) Oct 2025 – May 2026

- Architected and deployed autonomous multi-agent systems using **AutoGen and CrewAI** (Planner, Executor, Reviewer) to automate NOC/SOC alert triage and incident response, cutting manual execution time by **45%**.
- Engineered seamless tool-calling and enterprise data fetching pipelines via **Model Context Protocol (MCP)** with Claude as client, improving contextual retrieval accuracy by **35%**.
- Developed high-throughput, asynchronous microservices using **Python & FastAPI** for automated repository interaction, real-time log parsing, and system interoperability.
- Accelerated Python development and automated test debugging by integrating **Cline AI Agent** within the VS Code development lifecycle.

Associate AI & Software Engineer | VLax Tech, Pune, India Nov 2022 – Oct 2025

- Developed and deployed an enterprise LLM conversational AI chatbot using **Python and REST APIs**, processing **1,000+ daily queries** and reducing user wait times by **50%**.
- Architected high-performance RAG pipelines utilizing **FAISS and ChromaDB** vector stores with recursive document chunking, increasing retrieval precision by **38%**.
- Engineered system prompts and role-based guardrails across LLM families (GPT, LLaMA, Claude) to enforce output compliance and eliminate hallucinations.
- Connected GenAI workflows with relational SQL databases using Natural Language-to-SQL querying for automated reporting.

KEY PROJECTS

Automating Enterprise Email Triage & Response Generation | FastAPI, OpenAI API, Groq, Python GitHub Link

- Built an intelligent email triage pipeline using **FastAPI and Python** to classify corporate emails, parse semantic intent, and prioritize critical support tickets.
- Integrated **OpenAI & ultra-low latency Groq APIs** to generate context-aware draft responses, reducing manual inbox processing overhead by **60%**.

Agentic AI Multi-Agent Task Automation System | AutoGen, CrewAI, LangGraph, MCP, Claude GitHub Link

- Architected a multi-agent orchestration framework using **AutoGen and CrewAI** to execute complex, multi-step engineering workflows with dedicated planner, worker, and critic agents.
- Implemented **Model Context Protocol (MCP)** bindings and LangGraph cyclical graphs for dynamic tool execution, live environment feedback, and self-healing error correction.

Enterprise RAG-Based Document Analyzer Application | LangChain, ChromaDB, Python, Streamlit GitHub Link

- Built an interactive document analysis application using **LangChain and Streamlit** for multi-format document ingestion, recursive text chunking, and contextual multi-turn Q&A.
- Developed custom vector embedding pipelines using **ChromaDB** for low-latency semantic search with precise source citations and contextual re-ranking.

EDUCATION

Master of Computer Applications (MCA) | Dr. D. Y. Patil Centre for Management and Research, Pune (SPPU) 2023 – 2025

Bachelor of Science in Computer Science (B.Sc.) | DRB Sindhu College, Nagpur (RTMNU) 2019 – 2022

CERTIFICATIONS

- Oracle Certified Associate – Agentic AI Certified Foundations Associate** (Oracle) | **Python Developer** (NASSCOM)
- AWS Academy Machine Learning & Cloud Foundations** (AWS) | **SQL Advanced** (HackerRank)