

Aman Kumar

Kanpur, Uttar Pradesh | 7007219431 | amankumar.ak0012@gmail.com | [GitHub](#) | [LinkedIn](#)

PROFESSIONAL SUMMARY

AI Software Engineer with 3+ years of experience building production-grade AI platforms, scalable backend systems, and cloud-native applications for Fortune 500 enterprises and 0-to-1 startups. Expertise in **Python (FastAPI)**, **Java (Spring Boot)**, **LangChain**, **LangGraph**, **Model Context Protocol (MCP)**, enterprise **RAG**, and **real-time Voice AI**. Reduced AI inference costs by **45%** through intelligent model routing and delivered **99.99%** platform uptime as the sole Founding Backend Engineer. Experienced across **AWS**, **Azure**, **GCP**, **Terraform**, and microservices architecture

WORK EXPERIENCE

R SystemsNoida, Uttar Pradesh

Software EngineerNov 2025 - Present

- Built **production-scale AI platforms** for a **Fortune 500 enterprise** using **Python (FastAPI)**, microservices, and scalable backend architectures.
- Developed **custom AI agents** for **VS Code**, **Cursor**, **Claude Code**, and **IntelliJ IDEA**, automating code generation, repository analysis, requirement analysis, and SDLC workflows, improving developer productivity by **45%**.
- Designed and implemented **multi-agent systems** using **LangChain**, **LangGraph**, and **Model Context Protocol (MCP)**, integrating **Azure DevOps**, **GitHub**, and enterprise tools to automate development workflows.
- Built **production-grade RAG pipelines** supporting multi-format document ingestion (PDF, DOCX, PPTX, XLSX), vector search, and semantic retrieval, improving AI response accuracy by **40%**.
- Developed AI-powered solutions for **user story generation**, **manual test case creation**, and **requirement analysis**, reducing manual engineering effort by **65%**.
- Engineered an intelligent LLM routing strategy that leveraged GPT-5 mini for low-complexity tasks and premium models for advanced reasoning, reducing inference costs by 45% while maintaining response quality and improving overall system efficiency.
- Implemented **end-to-end AI observability** using **LangSmith** and **Datadog**, including prompt versioning, tracing, evaluation, and **P95 latency SLO** monitoring, reducing agent debugging time by **40%**.
- Automated **multi-cloud Infrastructure as Code (IaC)** using **Terraform** and **Terragrunt**, reducing new environment provisioning time from **~2 days to under 4 hours**.

CodewingletSurat, Gujarat

Software Engineer (Contract)Jun 2025 - Nov 2025

- Built a **production-scale AI Agent Platform** enabling users to create, deploy, and manage custom AI agents using natural language prompts.
- Architected a **Supervisor Agent** using **LangChain** and **LangGraph** to orchestrate specialized agents, improving workflow execution efficiency by **40%**.
- Designed a visual **AI workflow automation platform** similar to **n8n** and **Relevance AI**, enabling users to build intelligent multi-step workflows with reusable tools and integrations.
- Built **production-grade RAG pipelines** using **ChromaDB**, semantic search, and vector retrieval, improving AI response accuracy by **30%**.
- Implemented **real-time Voice AI** using **LiveKit**, **OpenAI Realtime/WebRTC**, and **ElevenLabs**, reducing end-to-end response latency by **40%**.
- Developed scalable asynchronous AI processing using **Azure Service Bus** and **MongoDB**, increasing workflow throughput by **45%**.
- Integrated enterprise tools including **SharePoint**, **Calendar**, **Email**, and custom APIs, automating business operations and reducing manual effort by **50%**.
- Leveraged **Claude Code** to build AI-powered developer assistants for code generation, repository analysis, debugging, and automated code reviews.

BXKG - Kiroku NotesSingapore, Contract

Software Engineering Team Lead (Contract)Apr 2025 - Nov 2025

- Led the backend engineering and AI architecture for an **AI-first scheduling platform**, serving **200+ automated workflows** across intelligent calendar and event management.
- Engineered **50+ production-grade REST APIs** using **Python (FastAPI)**, improving scheduling efficiency by **40%** through travel-aware planning and intelligent automation.
- Designed and implemented **supervisor-worker multi-agent systems** using **LangChain** and **LangGraph**, reducing manual scheduling conflicts by **60%**.
- Integrated **Google Maps**, Calendar APIs, and location intelligence to optimize travel routes and automate meeting coordination.
- Led a cross-functional team of backend, AI, and frontend engineers, driving architecture decisions, code reviews, and engineering best practices.
- Collaborated with product and design teams to deliver scalable AI-powered scheduling solutions with high availability and low-latency APIs.

Getaligned
Founding Engineer (Backend)

Bengaluru, Karnataka
Aug 2023 - Jul 2025

- **Served as the sole Backend Engineer**, architecting and building the entire backend infrastructure for an **AI-powered SaaS platform** from **0 to production**, owning system design, API development, cloud infrastructure, and deployments.
- Built an **AI Meeting Assistant** using **GPT, Mistral, and Llama**, automating meeting transcription, summaries, follow-ups, objection handling, and CRM updates, increasing post-meeting engagement by **60%**.
- Developed **AI agents** integrating with **Slack, Jira, ClickUp, Linear, Gmail, Google Calendar, Zoom, and Microsoft Teams**, automating cross-platform workflows and reducing manual task management by **45%**.
- Designed and developed scalable backend services using **Java (Spring Boot), Python (Flask), and Node.js**, powering enterprise-scale AI applications and REST APIs.
- Architected cloud infrastructure across **AWS (EC2, ECS, ECR, S3, CloudWatch)** and **GCP (Cloud Run, Firestore, BigQuery)**, achieving **99.99%** platform availability while reducing infrastructure costs by **15%**.
- Containerized microservices using **Docker** and optimized deployments on **Amazon ECS** and **Google Cloud Run**, improving deployment efficiency and scalability.
- Built CRM integrations and AI-powered workflow automation for **HubSpot, Salesforce**, and internal business systems, reducing manual sales operations by **45%**.
- Configured **Nginx** as a reverse proxy and load balancer, improving application performance, scalability, and fault tolerance.

EDUCATION

Swami Vivekanand Subharti University	Meerut, Uttar Pradesh
Bachelor of Computer Applications Computer Science (BCA)	Graduation Date: Feb 2024
Masai School	Bengaluru, Karnataka
Full Stack Web Development Program	Graduation Date: Jul 2023

SKILLS & INTERESTS

- Skills:**
- **Languages:** Java, Python, JavaScript, SQL
 - **Frameworks:** Spring Boot, Flask, FastAPI, Django, Node.js, React, Next.js (used in AI tool interfaces)
 - **Cloud & DevOps:** AWS, Azure, GCP, Docker, Terraform, Terragrunt, Nginx, Azure DevOps, Git, GitHub
 - **Databases:** PostgreSQL, MongoDB, MySQL, Redis, ChromaDB
 - **Technologies:** REST APIs, Microservices, Serverless Architecture, OAuth2, Authentication & Authorization, GitHub
 - **AI / LLM:** Agentic AI, AI Agents, Multi-Agent Systems, LangChain, LangGraph, Claude Code, NLP, LiveKit, WebRTC, RAG, Hugging Face, AI Observability, Vector Databases, MCP, n8n, AWS Bedrock, Azure AI

PROJECT EXPERIENCE

AI-Powered SDLC Assistant Platform

- **Tech Stack:** Python, FastAPI, LangChain, LangGraph, OpenAI, Azure DevOps, MCP, PostgreSQL, Redis
 - Built an enterprise AI platform that automates the Software Development Life Cycle (SDLC), including requirement analysis, user story generation, manual test case generation, and Azure DevOps work item creation.

- Designed a **multi-agent architecture** using **LangChain**, **LangGraph**, and **MCP**, enabling intelligent task orchestration and seamless integration with Azure DevOps and GitHub.
- Developed production-grade **RAG pipelines** supporting PDF, DOCX, PPTX, XLSX, and Markdown documents with semantic search and contextual retrieval.
- Built AI-powered developer assistants capable of analyzing codebases, generating technical documentation, and improving software delivery workflows, reducing manual effort by **50%**.

Enterprise AI Agent Platform

- **Tech Stack:** Python, FastAPI, LangChain, LangGraph, ChromaDB, LiveKit, OpenAI WebRTC, ElevenLabs, Azure Service Bus, MongoDB
 - Built a production-scale **AI Agent Platform** with **Supervisor Agents**, multi-agent orchestration, and workflow automation using **LangChain** and **LangGraph**.
 - Developed **RAG pipelines** with **ChromaDB**, semantic search, and vector retrieval, improving AI response accuracy by **30%**.
 - Implemented **real-time Voice AI** using **LiveKit**, **OpenAI WebRTC**, and **ElevenLabs**, enabling low-latency speech-to-text, text-to-speech, and conversational AI experiences.
 - Integrated enterprise tools including **SharePoint**, **Calendar**, **Email**, and asynchronous processing with **Azure Service Bus** and **MongoDB**, improving workflow throughput by **45%**.