

Priyansh Gupta

Senior Agentic AI Engineer | GenAI Lead | MLOps | Multi-Agent Systems

Senior Agentic AI & GenAI Engineer with 5+ years of experience delivering enterprise AI solutions on Azure, AWS. Expertise in RAG, multi-agent systems, MCP integrations, and Azure OpenAI. Proven success building production-ready GenAI platforms for Fortune 500 clients.

@ priyansh.9071@gmail.com

in priyansh19

🔄 priyansh19

☎ +971 506 479 919

WORK EXPERIENCE

Senior DevOps & AI Platform Engineer

🏢 Applied Information Services (AIS) 📅 Aug 2024 – Present 📍 Hyderabad, India

Keywords: Azure AI Foundry, Azure OpenAI, RAG, MCP, Azure AI Search, DevOps

- Deployed multimodal LLM models on Azure AI Foundry, configuring model endpoints, inference auto-scaling, and Azure AI Content Safety guardrails for production AI workloads.
- Architected enterprise RAG pipelines using document ingestion, chunking, and vector plus semantic hybrid indexing in Azure AI Search, improving grounded responses for Fortune 500 workloads.
- Delivered MCP integration for Azure Boards, enabling AI assistants to reference live tickets, sprint data, and work items for enterprise agentic tool-use.
- Led GenAI platform modernisation for FMGlobal, supporting zero-downtime migration of critical insurance workloads to secure Azure AI infrastructure.
- Standardised CI/CD governance across 20+ microservices and deployed JFrog Artifactory with Docker image promotion gates for container supply-chain auditability.

Senior GenAI / MLOps / Agentic AI Consultant

🏢 Freelance Consulting (Upwork) 📅 Aug 2022 – Present 📍 Remote – Global Clients

Keywords: LangChain, Pinecone, ChromaDB, Azure OpenAI, Claude, GPT-4o, MLOps

- Architected end-to-end RAG pipelines using LangChain with Pinecone and ChromaDB for semantic retrieval, top-k reranking, and grounded responses for production SaaS clients.
- Delivered production agentic AI solutions embedding Azure OpenAI, Claude, and GPT-4o into SaaS workflows with tool-use, function-calling, and autonomous decision-making.
- Deployed MLOps infrastructure on Azure ML and AWS SageMaker, orchestrating agentic pipelines with n8n and Python for model monitoring and drift detection.

DevOps Engineer

🏢 Applied Information Services (AIS) 📅 Oct 2022 – Aug 2024 📍 Hyderabad, India

Keywords: AKS, Helm, Kyverno, Istio, FluxCD, Azure Key Vault

- Engineered Helm Charts for AKS onboarding with rollback-enabled pipelines and implemented Kyverno admission control for security baselines, resource quotas, and image signing.
- Configured Istio service mesh with mTLS zero-trust peer authentication, FluxCD GitOps state management, and canary plus A/B traffic routing.
- Managed secrets via Kubernetes Secrets Store CSI Driver and Azure Key Vault, enforcing least-privilege access across AKS clusters.

Software Engineer

🏢 Applied Information Services (AIS) 📅 May 2021 – Oct 2022 📍 Hyderabad, India

Keywords: Terraform, Bicep, Azure DevOps, YAML, AKS, Key Vault

- Built reusable Terraform and Bicep modules for Azure VNets, Key Vaults, AKS, and Application Gateways.
- Authored Azure DevOps YAML CI/CD templates for build-test-deploy workflows across cloud-native services.

CERTIFICATIONS

AI-102 Azure AI Engineer Associate · AZ-400 DevOps Engineer Expert · CKA Certified Kubernetes Administrator · CKAD Certified Kubernetes Application Developer

AZ-700 Azure Network Engineer · AZ-104 Azure Administrator · AI-900 Azure AI Fundamentals · AZ-900 Azure Fundamentals

EDUCATION

M.Sc. Artificial Intelligence

🏛 Heriot-Watt University 📅 In Progress 📍 Dubai, UAE

- Focus on agentic AI systems and GenAI applications.

B.Tech. Computer Science

🏛 University of Petroleum and Energy Studies 📅 2017 – 2021

- Specialized in cloud computing and AI/ML coursework.

KEY PROJECTS

CrewSpace – Open-Source Multi-Agent AI Workforce Platform

[Github Link](#)

Hierarchical orchestration · Persistent memory graph · 3D Workspace · Agent Collaboration

- Architected and built CrewSpace from scratch, an AI workforce platform where a CEO agent autonomously decomposes user goals into domain-specific subtasks, delegates work across specialized AI agents, dynamically proposes new hires when required, and coordinates execution through multi-agent communication and approval workflows.
- Developed persistent per-agent memory, agent-to-agent and group chat, shared workspace and file permissions management, integrated terminal operations, real-time task boards, founder dashboards, and an interactive 3D office environment enabling users to monitor, collaborate with, and manage long-running AI organizations at scale.

OpenClaw — Open-Source Autonomous AI Agent Platform

[Github Link](#)

Active Open-Source Contributor · Multiple PRs Merged

- Active contributor to OpenClaw with multiple pull requests merged into the core codebase, improving agent runtime tooling, workspace file handling, and developer-facing diagnostics for autonomous AI agent workflows.
- Fixed image-tool workspace path resolution by making relative image paths resolve against the agent workspace directory instead of `process.cwd()`, preserving existing allowlist security checks and adding regression coverage.
- Improved write/read/edit tool validation errors by showing received parameter keys when required parameters are missing, making silent tool-argument drops easier to debug in long-running agent sessions.

AutoOps AI – Autonomous DevOps & Infrastructure Management Agent

GitHub + CI/CD Monitoring · LLM Root-Cause Analysis · Terraform Auto-Remediation

- Built an autonomous AI agent for monitoring GitHub repositories and CI/CD pipelines, performing multi-step LLM reasoning and root-cause analysis, raising PRs, and triggering Terraform-based infrastructure remediation on Azure.
- Implemented natural language IaC provisioning where plain-English commands are parsed via LLM function-calling, converted to HCL, validated, planned, and applied with auto-revert logic through Prometheus alerting.

RAG Production Chatbot – Local Ollama + Vector DB Pipeline

Zero External API Dependency · FastAPI REST Layer · ChromaDB · Full Audit Logging

- Shipped a production RAG chatbot on local Ollama inference with a ChromaDB vector pipeline for chunking, indexing, semantic search, and top-k retrieval.
- Built a FastAPI REST layer with structured audit logging, ensuring data-privacy compliance with zero external API dependency.

TECHNICAL SKILLS

Agentic AI & LLMs

LangChain · LangGraph · AutoGen · LlamaIndex · MCP · RAG · Tool-use · HITL · Guardrails · Chain-of-Thought · Agent Memory · Multi-Agent Orchestration

Azure AI Platform

Azure AI Foundry · Azure OpenAI (GPT-4o) · Azure AI Search · Azure ML · Prompt Flow · Azure AI Content Safety · Microsoft Fabric · Copilot Studio

Cloud, IaC & DevOps

Azure · AWS · GCP · Terraform · Bicep · ARM · Azure DevOps · GitHub Actions · ArgoCD · FluxCD

Containers & Orchestration

Kubernetes (AKS, EKS) · Docker · Helm · Istio · Kyverno · OPA

Vector & Search

Azure AI Search · Pinecone · ChromaDB · Qdrant · HuggingFace Embeddings · Top-k Reranking

Security & Observability

Prompt Injection Defense · RBAC · mTLS Zero-trust · Azure Key Vault · Prometheus · Grafana · Azure Monitor

Backend & APIs

Python · FastAPI · Node.js · TypeScript · REST APIs · Microservices · Event-driven Architecture