Raj Kumar

+917780684490 | raj315920@gmail.com | linkedin | portfolio | Github

ABOUT ME

Software Engineer with 2+ years building cloud-native systems, AI agent platforms, and secure backend services from the ground up. I work across the full stack, from React interfaces to PostgreSQL and AWS infrastructure, and I ship AI-driven products that handle real user load. I care about correctness: deterministic evaluation, secure payment and CRM integrations, and failure modes designed up front rather than patched later. I use AI coding tools daily and review every file they touch. Open to relocating to New York.

EDUCATION

Andhra Loyola Institute of Engineering

Bachelor of Engineering *Graduation Date: Jul 2023*

PROJECT EXPERIENCE

DriftShield: RL Gym for Adversarial Agentic Systems huggingface.co/spaces/raj23211/driftshield

- Investigated whether GRPO-trained LLM policies can learn resistance to adversarial inputs, schema drift, and poisoned tool outputs with no secondary LLM in the reward path.
- Built an OpenEnv-compatible RL gym simulating a six-service enterprise environment with four adversarial task families mapped to the OWASP LLM Top 10.
- Designed a fully deterministic grader (5 reward components, 4 DriftShield bonuses, 13 penalties) with reproducible results across seeds and no LLM-as-judge.
- Fine-tuned Qwen3-1.7B with GRPO + LoRA over 640 episodes and measured a 2.7x higher reward ceiling on schema-drift than on prompt-injection tasks.
- Committed proof artifacts (reward logs, curves, Colab reproducibility notebook) and deployed live on Hugging Face with Docker.

VulnGate: Security Gate for AI-Generated Code (Go)

- Offline CLI that gates merges of AI-agent-generated code, built in 24 hours. Composes semgrep, gosec, and gitleaks and owns only the merge-gate decision instead of reinventing scanners.

AegisGate: LLM Security Middleware (Azure AI Foundry, Semantic Kernel, Python)

- Azure-native real-time security layer that intercepts LLM and agentic tool calls before execution, with a classifier pipeline, YAML policy enforcement, and a forensic audit trail. Enforces behavioral constraints at the tool-call boundary.

AgentGuard: Multi-Agent Safety Evaluation System (2025 UiPath AgentHack, Track 3)

- Three-agent framework in Python and LangChain: an Adversary agent probes the target, an Evaluator agent scores behavior against safety policy, and a Healer agent patches detected failures before production.

WebSentry AI: Agent Governance Platform websentry-ai.vercel.app

- Governance layer for AI web agents with real-time prompt-injection defense, behavioral risk scoring, and live session monitoring.

WORK EXPERIENCE

Merchant Nation California

Full Stack Developer *Oct 2025 - Present*

- Built LLM, RAG, and workflow-automation platforms for conversational AI handling high interaction volume, with multilingual retrieval pipelines and knowledge-base orchestration.
- Built backend integrations with external APIs, CRM platforms, payment systems, and event-driven webhooks for dependable data exchange.
- Built automation services for data processing, user qualification, routing, and decision support, reducing manual work.
- Applied authentication, access control, and secure communication across distributed AI services and third-party integrations.
- Built monitoring, logging, and analytics for performance, user behavior, and reliability; shipped production platforms with cross-functional teams.

Plexora Bangalore, India

Founding Engineer *Feb 2025 - Sep 2025*

- Architected a real-time analytics platform (React, Node.js, PostgreSQL) with sales, inventory, and customer insight dashboards used by 150+ restaurants.
- Built an LLM-powered business intelligence assistant for natural-language queries of business metrics, cutting manual analysis time by 50%.
- Deployed on AWS with Docker and CI/CD, holding 99.7% uptime through 200% user growth.
- Shipped customizable visualization modules that raised client engagement by 40% and operational visibility by 35%.
- Built predictive analytics and inventory-loss detection that helped restaurants cut food waste by 20%.

Andhra Pradesh Technology Services — Vijayawada
Full Stack Developer — *Jul 2023 - Nov 2024*

- Migrated legacy C++ and .NET vehicle management systems to secure, cloud-native architectures on Microsoft Azure.
- Built secure CI/CD pipelines with Azure DevOps and Docker, improving release reliability and reducing operational risk.
- Integrated DevSecOps into the SDLC: static analysis, dependency auditing, vulnerability scanning, and security hardening.
- Led large-scale SQL migrations to Azure SQL and Oracle Cloud while preserving data integrity and availability.
- Ran internal workshops on secure coding and DevSecOps for cross-functional engineering teams.

Octazen
Software Engineering Intern — *Jul 2022 - Dec 2022*

- Helped build a cloud file synchronization engine across AWS S3, Azure Blob Storage, and Google Drive, with REST APIs for sync and version control.
- Built the React interface for sync status and failures, and wrote test scripts that cut production sync-related bugs by 30%.

SKILLS & INTERESTS

Languages: Python, TypeScript, JavaScript, Go, SQL
Backend and data: FastAPI, Node.js, Express, PostgreSQL, pgvector, Prisma ORM, BullMQ, WebSockets, Elasticsearch
Frontend: React, Next.js, Vue 3, Tailwind CSS
Cloud and infra: AWS, Azure AI Foundry, Azure DevOps, Azure Content Safety, Docker, Kubernetes
AI and agents: LangChain, Semantic Kernel, Mastra, RAG, FAISS, Pinecone, Hugging Face, TRL, vLLM, OpenEnv, GRPO, PPO, LoRA/PEFT, reward modeling, agent evaluation, prompt injection defense, OWASP LLM Top 10
Security: bug bounty research, API security, authentication and access control, DevSecOps
Interests: Chess, Design, Philosophy, AI, Open Source, Strategy, Psychology, Storytelling, Ethics

ACHIEVEMENT

- Winner, Microsoft Build AI Hackathon India (AegisGate).
- Top 4 Finalist at AiroThon 2025, an Agentic AI Hackathon organized by Airo Digital Labs; awarded a certificate and a reward for an innovative AI solution.
- Participant in Bay2BayHacks 2025, recognized with a certificate for an innovative project.
- Reported 30+ vulnerabilities through coordinated bug bounty programs: triaged reports, validated findings, and assisted remediation with engineering teams.