

PULKIT SARAF

✉ pulkitsaraf.dev@gmail.com 📞 +91-8882655977 📍 New Delhi 🌐 psudokit.in 🔄 Pulkit7070 📧 Pulkit7070 🌐 linkedin.com/in/pulkit-saraf-893213290

Software Engineer | AI Agents, LLMs, Full Stack & Distributed Systems

SUMMARY

Software engineer with 1+ year building **AI agent tooling**, **LLM services** and **scalable backends** in TypeScript, Python and Rust. CNCF open-source contributor (21+ PRs across Jaeger, Meshery and OpenFGA) and 10x national hackathon winner.

SKILLS

Languages: Python, TypeScript, JavaScript, Rust, C++, Go, SQL

Frameworks: React, Next.js, Node.js, Express, PyTorch, TensorFlow, LangChain

Infrastructure: AWS, GCP, Docker, GitHub Actions, PostgreSQL, Redis, MongoDB, GraphQL, gRPC, WebSocket

AI & Agents: LLMs, RAG, AI Agents, Multi-Agent Orchestration, Prompt Engineering, Vector Databases (pgvector), NLP

Systems: Distributed Systems, Microservices, Event-Driven Architecture, Observability (OpenTelemetry, Sentry)

Soft Skills: Technical Leadership, Project Management, Cross-functional Collaboration, Communication, Problem Solving

EXPERIENCE

ArmorIQ.ai

AI Engineer

Full Stack Developer Intern

Remote

Aug 2026 - Present

Feb 2026 - Aug 2026

- Architected a 5-package Turborepo codebase (frontend, backend, tooling) with parallel CI/CD builds, **cutting deploy time by 40%**.
- Built policy-enforcement plugins for **4 coding agents** (Claude Code, Codex, OpenCode, Antigravity) gating tool calls on intent and policy.
- Wrote a Go hook client for Antigravity adding **~0.5 ms** over a bare shell, meeting a 30 ms p95 budget (Bun added 20 ms).
- Fixed token-usage accounting from agent transcripts that naive summing inflated **~900x**, powering per-session cost dashboards.

Agent Orchestrator (ao)

Core Contributor

Bengaluru, India

May 2026 - Oct 2026

- Shipped **86 merged PRs** (36K+ lines) to a desktop IDE for coding agents used by ~1,000 daily users and 10,000+ session spawns/day.
- Cut received-text lag **93% (2,857 to 213 ms)** via bounded event batching and incremental streaming, plus backoff SSE reconnects.
- Rebuilt the cross-platform auto-updater (macOS repair installer, progress watchdog, bounded retries), shipping hotfixes **within 15 min**.
- Built PostHog/Sentry telemetry for desktop and mobile (DAU heartbeat, retention, rate caps), cutting costs **from \$500 to \$0/month**.

PreciQube (IIT Madras Startup)

Full Stack Developer & Research Intern

Chennai, India

Oct 2025 - Nov 2025

- Delivered a research-grade 2D optics simulation platform with physicists that saves scientists from building simulators from scratch.
- Integrated ray tracing into the simulation engine, improving responsiveness by **~20%**.
- Translated **4 optics models** from peer-reviewed papers (refraction, reflection, Snell's law, Fresnel) into the engine for reproducibility.

Rabbit AI

AI Engineer & Technical Project Manager

Remote

Jun 2025 - Oct 2025

- Co-developed a TensorFlow ML platform serving SLM inference via REST APIs, lifting **model accuracy 20%** and **engagement 30%**.

OPEN SOURCE

CNCF · Jaeger · Meshery · OpenFGA · Falco Playground

Mar 2026 - Present

- Contributed **21+ PRs**: gRPC error handling, Badger span-kind storage (closing 8-year-old TODO #1922) and Postgres pool warm-up.
- Shipped 4 PRs to **C2SI/CodeLabz**: MUI v5 migration, Firestore notifications, RBAC and Vitest test coverage.
- Authored **multigravity-pro** (23 stars, 5 forks), a CLI for unlimited isolated IDE profiles with templates, a dashboard and export/import.
- Published **get-cooked** on npm, tracking token usage and rate limits across 3 AI coding agents: Claude Code, Cursor and Copilot.

PROJECTS

LazyClip: Conversational Video Editor | TypeScript, React, Convex

www.lazyclip.buzz | Jul 2026 - Present

- Co-founded a video editor that turns natural-language requests and YouTube links into shorts; awarded **\$5,800** at GrowthX.
- Launched a Hermes-powered Telegram clipping workflow and a TypeScript/React editor on Convex, drawing **100+ demo-day signups**.

agmonverse: AMM Strategy Prototype | Python, Solidity, LLMs

github.com/Pulkit7070/agmonverse | Feb 2026

- Prototyped an AI AMM strategy engine that generates adaptive Solidity fee policies, cutting arbitrage losses **7.9%** vs. a 50 bps baseline.
- Implemented a generate-validate-simulate loop that evolves strategies across 10 focus areas, scored by loss-versus-rebalancing.

Blink AI: Multi-Tool Desktop Assistant | TypeScript, Electron, Deepgram

github.com/Pulkit7070/blink-ai | Mar 2026

- Built a desktop AI assistant with live Deepgram transcription, screen OCR and plugins routing chat to **240+** Composio and MCP tools.

ACHIEVEMENTS & LEADERSHIP

GrowthX Buildathon Winner (Jul 2026): won \$5,800 in credits with LazyClip on Nous Research's Hermes agent; 100+ signups.

10x National Hackathon Winner: 30+ hackathons; selected for ETHGlobal, ETH Mumbai and the Algorand Ecosystem.

Technical Head, E-Cell R&D at VIPS: led startup incubation; pitched at Aarambh'26, IIT Delhi.

EDUCATION

Vivekananda Institute of Professional Studies (VIPS-TC), GGSIPU

New Delhi, India

B.Tech in Computer Science and Engineering (CGPA: 8.32/10)

Sep 2023 - 2027 (Expected)

Coursework: Data Structures, Algorithms, OOP, Operating Systems, Computer Networks, DBMS, Computer Architecture, Theory of Computation, Compiler Design, Discrete Mathematics, Software Engineering, AI, Machine Learning, Distributed Systems