

SUMANTH KATTA

Hyderabad, India | +91 7075010405 | sumanthkatta7075@gmail.com | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

SUMMARY

AI/ML Engineer & Full-Stack Developer (B.Tech CSE – AI & ML, CMRCET, 2027). Built and shipped production Gen-AI and Agentic AI applications LLM API integration, n8n automation pipelines, and edge-AI prototypes with real users. Experienced in rapid prototyping, AI pilot documentation, and cross-functional knowledge-sharing. Passionate about applying AI to streamline real-world workflows, directly aligned with Zycus Merlin AI's agentic procurement automation mission.

EDUCATION

CMR College of Engineering & Technology (CMRCET)
B.Tech CSE – Artificial Intelligence & Machine Learning

Hyderabad, India
2023 – 2027 (Expected)

TECHNICAL SKILLS

- **AI and GenAI:** Python, OpenAI API, Gemini API, LLM Integration, RAG (Retrieval-Augmented Generation), Agentic AI, Prompt Engineering, LangChain, LlamaIndex, Chatbot Development, TensorFlow, PyTorch, ESP32 Edge AI
- **Automation:** n8n (Advanced), Workflow Automation, Low-code Tools, Make, Zapier, Docker, CI/CD, GitHub Actions
- **APIs and Data:** REST APIs, JSON, FastAPI, Node.js, Express.js, Firebase, Supabase (PostgreSQL), SQLite, Python data pipelines
- **Frontend:** React.js, Next.js, TypeScript, JavaScript, Tailwind CSS, HTML5/CSS3
- **Cloud and Tools:** Microsoft Azure, Vercel, Netlify, Git, Linux/CLI, Ollama, LM Studio, Figma, Postman

PROJECTS

LinkSnap AI – Gen-AI Bookmark Manager

Open-Source Lead | 2025–Present | [linksnapai](#)

- Integrated Gemini 1.5 Flash Vision API for auto-metadata extraction, cutting manual tagging by ~90% and eliminating a key friction point in the bookmark workflow
- Maintained contributor backlog, CI/CD pipeline, and bi-weekly releases; authored prompt libraries and technical docs to support adoption.

Override & Airlock – Agentic Browser Automation

Independent | 2025

- Override: Agentic AI Chrome extension using LLM tab-intent classification to reduce knowledge-worker context-switching in real time.
- Airlock: Client-side Data Loss Prevention (DLP) sentry intercepting sensitive data streams zero backend, rapid-prototyped in vanilla JS.

Phyto-Cognition – Edge AI Platform

Independent | Hackathon Finalist | 2025–2026

- Offline-first ESP32-S3 AI platform with 7 modules (Gemini API assistant, CNN inference <300ms, IoT sensors) — produced full docs + stakeholder pitch deck.

Last Seat – Real-Time PWA

Independent | 2025–2026

- React + TypeScript + Firebase PWA with dual GPS roles and sub-second latency prototype to production in a single sprint.

EXPERIENCE

AI Developer & Open-Source Lead Maintainer (Self-Directed)

2026 – Present

- **Engineered 6 n8n agentic workflow automation pipelines** integrating OpenAI and Gemini APIs via REST/JSON eliminating ~4 hours/week of manual processing per workflow and directly replicable for enterprise procurement automation and PO routing.
- Deployed local LLM inference (Qwen2.5-Coder-7B via Ollama) on a 60/40 GPU-CPU split, achieving **300 tokens/min at ~0.75s latency and \$0 API cost** vs. cloud API pricing demonstrating end-to-end on-device AI pilot ownership.
- Managed 2 open-source repositories (LinkSnap AI, Data Refinery Core) across **full release cycles**: code reviews, GitHub backlog, feature releases, prompt libraries, and knowledge-transfer documentation mirroring enterprise AI pilot support workflows.

CERTIFICATIONS & ACHIEVEMENTS

Columbia Engineering – Prompt Engineering | **Salesforce** – Agentforce (Autonomous AI Agents) | **Deloitte** – Software Strategy & Data Viz
Tata Group – Data Risk & Cybersecurity | **McKinsey** – Strategic Problem Solving | **Agricultural-AI Hackathon Finalist** (Phyto-Cognition, 2025–26)

EXTRACURRICULAR & INTERESTS

AI-generated content creation (video, art, generative imagery). | Journaling & reflective writing. | Drawing, painting, and mixed-media crafts. | Evaluating and stress-testing new AI developer tools and workflows.