

SWAYAM GOSAVI

[+91 9822687804](tel:+919822687804) swayamgosavi1410@gmail.com [Portfolio](#) [LinkedIn](#) [GitHub](#) [LeetCode \(750+\)](#)

EDUCATION

Pune Institute of Computer Technology
B.E. in Electronics and Telecommunication

Sept 2023 – June 2027
Pune, Maharashtra

PROFESSIONAL SUMMARY

Software engineer specializing in Backend engineering, Applied AI engineering, DevOps and Infrastructure and distributed architectures. A two-time Founding Engineer, Co-Founder and an avid Open Source Contributor at the age of 20, with experience taking products from concept to production in fast-paced startup environments while owning architecture, implementation, deployment, and technical decision-making. Passionate about building CLIs, web-apps and stuff which solve real problems and shipping them.

EXPERIENCE

Founding Engineer
PensionIntel

May 2026 – Present
Hybrid

- Working directly with the founding team in a 0→1 fintech SaaS startup with technical ownership across backend engineering, infrastructure, and product development.
- Leading technical execution and streamlining product workflows while coordinating across distributed engineering teams and refactoring legacy business logic into scalable backend systems.
- Designing hierarchical multi-tenant SaaS architecture, cloud-native infrastructure, and scalable backend systems for secure customer onboarding and tenant isolation.
- Implementing spreadsheet-driven financial computation workflows and UK NHS pension-calculation systems involving domain-specific business logic and forecasting models. Achieved a 99.27% accuracy for the pension calculations
- Collaborating with financial practitioners and consulting stakeholders to streamline operational workflows while developing AI-driven lead funnel systems.

Founding Software & AI Engineering Intern
Tecnviron Pvt Ltd

Jul 2025 – Jan 2026
Remote

- Designed and developed a B2B client management platform integrating WhatsApp (WABA), Facebook, voice-bot, and payment workflows using TypeScript and Python.
- Built automated lead-generation and chatbot workflows across WhatsApp and Meta platforms, reducing manual operational overhead and supporting business automation.
- Worked directly with clients to gather requirements and iterate on product workflows for production deployments.
- Implemented CI/CD pipelines using GitHub Actions and Docker with automated Azure deployments, reducing deployment time by 30%.

Co-founder
janus

Jul 2026 – Present
Remote

- Building janus, an infrastructure layer for agents and agentic workflows in stealth
- More at janushq.dev

Webmaster
PICT IEEE Student Branch

Jan 2025 – Present
College

- Lead the web-development team of my student branch end-to-end, overseeing, building and shipping websites for main website and multiple club events.
- Built and deployed the central festival website, along with a codeforces style online judge from scratch for internal events, multiple quiz platforms and oversaw the development of the trading event platform.
- Collaborated with the executive team of the club to translate organizational requirements into web solutions building personal skills like ReactJS, JavaScript, TypeScript, Python, Go, ExpressJS, FastAPI, SQL, MongoDB
- Managed the DevOps flow of the multiple websites gaining skills like AWS, GCP, CI/CD, Docker, Nginx, CDN, Linux, Redis, BullMQ, GitHub Actions
- Mentoring juniors of college over writing clean, scalable and production ready code. Lead SIGs(Student Interest Groups) and discoursed sessions on Backend Development, Databases, DevOps, AI Engineering

PROJECTS

Distributed Remote Code Execution Engine — Python, Go, JavaScript, Docker, Redis, PostgreSQL, GCP

- Designed the architecture and implemented a remote code execution system using Docker-based sandboxing, Redis queues, WebSockets, and Python/Go workers, handling 300+ concurrent users with latency comparable to LeetCode and Codeforces.
- Implemented gRPC-based communication between the submission gateway and Go worker nodes, enabling low-latency task dispatch and result aggregation across distributed execution environments.
- Deployed backend on GCP with CI/CD automation, NGINX reverse proxies, and RSA-encrypted database connections; system was used for competitive coding events run by PISB.

sentry — Python, LangGraph

- Engineered a Python-based Service Level Agreement (SLA) framework for AI agents, featuring execution tracing, budget tracking, and stateful replay capabilities
- Developed modular framework adapters to seamlessly integrate the SLA system with the Claude SDK, LangGraph, and raw operational loops.
- Architected an automated benchmarking harness with judge heuristics, taxonomy classification, and real-time metric tracking to evaluate agent performance.

Clio — React, TypeScript, Supabase, BullMQ, Node.js, AI Agents

- Building an AI-native version-control platform for professional documents such as resumes, cover letters, essays, and research papers, bringing full Git-style workflows, branching, committing, pushing, pulling, merging, rebasing, stashing, reverting, and cherry-picking, into a document domain with no real prior tooling.
- Designed a DAG-based version engine with section-level merging and per-block conflict detection on top of a content-addressed, normalized document store, so distributed-systems-grade versioning stays hidden behind a simple interface any non-technical user can pick up.
- Built a dual-mode AI collaborator directly into the editing workflow: an advisory "ask" mode that suggests edits, and an autonomous "agent" mode that writes and applies changes on its own.
- Pre-launch with validated product leverage; owned the build end-to-end across frontend, backend, and AI-orchestration as sole technical lead.

doc2html — Python, Pip

- Building an open-source, cli-installable conversion pipeline that turns image-heavy, tabular, and color-dense PDFs, Word documents, and PowerPoint decks into clean, semantically structured HTML, aimed at cutting token overhead for downstream LLM use.
- Built a multi-modal document parser that handles image regions, tables, and color-coded layout elements, hitting roughly 90.2% conversion accuracy across mixed document formats in early testing.
- Early benchmarks show a 30–70% cut in LLM token consumption compared to feeding in raw documents, giving a theoretically grounded and empirically validated efficiency win for AI document-processing pipelines.

Phason — TypeScript, Node.js, Ollama (Llama3), Git Hooks

- Developed and shipped an LLM-powered CLI tool as an npm package that generates context-aware commit messages from staged git diffs using a locally-run language model, integrated via git hooks.
- Architected the tool around local Ollama inference (Llama3:8b), ensuring zero data leaves the user's machine, making it safe for use with private codebases and proprietary repositories.
- Achieved 100+ installs within the first month of release and sustained 5+ daily installs, demonstrating shipping discipline and practical AI tooling experience.

TECHNICAL SKILLS

Languages: Python, TypeScript, JavaScript, Go, Java, Bash, C++

Cloud/Infra: AWS, GCP, Azure, Docker, Kubernetes, EKS, Terraform, NGINX, GitHub Actions, Jenkins

Backend: FastAPI, Flask, Express, PostgreSQL, Redis, RabbitMQ, MongoDB, WebSockets