

Prasad Firame

Pune, India | +91 7387930755 | pframe18@gmail.com | linkedin | github | blogs

Technical Skills

Languages: Python, Go, TypeScript, JavaScript, SQL

Backend: Node.js, FastAPI, REST APIs, Microservices, RabbitMQ, Redis

Frontend: React, HTML5, CSS3, Tailwind CSS

Databases: PostgreSQL, MySQL, MongoDB, ChromaDB, Pinecone

AI / LLM: OpenAI API, LangChain, Retrieval-Augmented Generation (RAG), Semantic Search, Vector Databases, AI Agents, Voice AI

Cloud & DevOps: AWS (Lambda, EC2, S3, SES, CloudWatch), Docker, Kubernetes, CI/CD, Linux, Git

Experience

CoReCo Technologies

Nov. 2024 – Present

Associate Software Engineer

Onsite

- Engineered scalable backend systems, async APIs, and microservices using **Node.js**, **FastAPI**, **PostgreSQL**, **Redis/Valkey**, and **RabbitMQ** for AI-driven product development platform.
- Owned end-to-end backend architecture across **authentication**, **dataset processing**, **AI workflows**, **experiment orchestration**, and **real-time systems** with modular service-oriented design.
- Designed and orchestrated multi-step **RAG pipelines** integrating **OpenAI** and **Perplexity APIs**, automating insights generation across 100+ products/day and reducing analyst turnaround time by **60%**.
- Built multi-tenant AI backend infrastructure for **Nexus**, securely supporting **500+ tenants** with **RBAC/ABAC authorization**, isolated workflows, and scalable retrieval pipelines.
- Developed ETL and data processing pipelines handling **50k+ product records/day** with LLM-based enrichment and semantic search workflows, achieving **95% data accuracy**.
- Architected production-grade **real-time voice AI agents** using **Twilio**, **Deepgram**, **ElevenLabs**, **WebSockets**, and **OpenAI**, processing **1k+ conversations/week** with low-latency streaming pipelines.
- Built async background job and task orchestration systems using **Redis queues**, **RabbitMQ**, and worker-based execution pipelines for long-running AI and data processing workloads.
- Refactored backend modules into **domain-driven** and modular architectures with clearly separated services, repositories, and schemas, improving maintainability and reducing engineering overhead.
- Implemented automated testing pipelines using **Pytest** and integration testing strategies for API validation, improving deployment reliability and reducing production regressions.
- Built and deployed serverless **AWS infrastructure** using **Lambda**, **SES**, **S3**, and **CloudWatch** to automate reporting workflows generating **500+ reports/month** and reducing manual effort by **80%**.
- Containerized services with **Docker** and automated deployments through **CI/CD workflows**, significantly reducing release turnaround time and improving deployment consistency.
- Developed scalable frontend systems using **React** and **TypeScript** for AI-powered internal tools, analytics dashboards, and operational workflows.
- Led a team of **2 engineers** in designing and shipping MVP features, conducting technical interviews, mentoring interns, and improving backend engineering practices.

CoReCo Technologies

July 2024 – Oct. 2024

Software Engineer Intern

Onsite

- Built context-aware LLM chatbot backend systems achieving 95% intent accuracy, improving automated customer support resolution workflows.
- Developed backend APIs and real-time operational modules for Wipro PARI's enterprise parking automation platform using scalable service-oriented architecture principles.

Projects

Orbit 🌀 | *Docker, WebSockets, Go, Node.js, Python*



- Ephemeral cloud compute engine that spins up isolated Docker execution environments for secure, on-demand code execution with real-time terminal access.

Shuriken 🌀 | *Python, FastAPI, Agentic Systems, Distributed Architecture*



- Agent orchestration SDK for building and running multi-framework AI agents with unified execution, scheduling, RBAC, and lifecycle management of memory, tools, and traces.

Education

Shivaji University

June. 2020 – May 2023

Bachelor of Science in Computer Science

Karad, Maharashtra