

SAI PRANEETH PEGADA

Software Engineer

+91-7989427037 | saipraneethpegada@gmail.com | LinkedIn | GitHub | Portfolio

SUMMARY

Software engineer with 4+ years of professional experience, including 3+ years building production web and AI applications. Currently the top contributor on a HIPAA-compliant clinical AI platform — building Next.js/TypeScript frontends, Python/FastAPI backends, and LangGraph agent pipelines on AWS Bedrock end to end. Created two shared infrastructure codebases used across production services; drive releases, code reviews, and engineering standards across a multi-repo, multi-team codebase.

TECHNICAL SKILLS

- **Languages:** Python, TypeScript, JavaScript, SQL, Bash
- **Frontend:** React, Next.js (App Router), Tailwind CSS, TanStack Query, TanStack Table, shadcn/ui, Radix UI, TipTap, Redux, Material UI, Vite
- **Backend:** FastAPI, Pydantic, SQLAlchemy (async), Alembic, Django, REST APIs, Server-Sent Events (SSE), Redis
- **AI / LLM:** LangGraph, LangChain, Anthropic Claude, AWS Bedrock (Knowledge Bases, Prompt Management), OpenAI, RAG, prompt engineering and versioning, LLM token/cost observability
- **Cloud & Infrastructure (AWS):** Lambda, ECS Fargate, API Gateway, SAM/CloudFormation, DynamoDB, S3, SQS, EventBridge, Cognito, CloudWatch, RDS (PostgreSQL), Secrets Manager, CodeArtifact, IAM, Amplify
- **DevOps & Quality:** GitHub Actions (CI/CD), Docker, pytest, ruff, ESLint, Trivy, CodeQL, Git
- **Practices:** HIPAA/PHI compliance, code review, release management, technical design documentation, mentoring

PROFESSIONAL EXPERIENCE

Software Engineer — Kinwits | Jul 2025 – Present

Core engineer for healthcare client bVital / Energy4Life Centers: built E4L-Prometheus, a HIPAA-compliant AI platform in production that generates Parkinson's treatment plans from EHR (Cerbo), CRM (GoHighLevel), and lab data — top contributor with ~56% of backend and ~68% of frontend commits.

Clinical AI platform (Python, FastAPI, LangGraph, AWS Bedrock, Next.js, TypeScript)

- Led the platform-wide LLM migration from OpenAI to Anthropic Claude on AWS Bedrock: designed the provider-abstraction layer (text, JSON, vision, and structured-output APIs) that became the single integration point for all AI calls, enabling environment-driven provider switching and safe model upgrades — a 119-file refactor.
- Architected a supplement recommendation engine using RAG over AWS Bedrock Knowledge Bases — clinical rules extracted from practitioner transcripts, deterministic pre-LLM selection over a 500+ product catalog with fuzzy matching — the platform's largest subsystem (14K+ lines), backed by 112 unit tests.
- Designed an asynchronous clinician-feedback pipeline (DynamoDB Streams → Lambda → SQS FIFO → LangGraph workers) that classifies provider corrections to AI-generated plans and proposes fixes, with conflict detection and an insights analytics API — closing the human-in-the-loop AI quality cycle.
- Built the nightly batch plan-generation service on ECS Fargate with EventBridge scheduling, two-level deduplication, and per-plan token/cost tracking with CloudWatch alarms.
- Hardened HIPAA compliance across the stack: PHI scrubbing in structured logs, customer-managed KMS encryption, immutable (WORM) audit archives, and routing LLM vision calls through AWS Bedrock to keep PHI within the compliance boundary.
- Resolved production reliability and performance issues: subprocess-isolated PDF rendering to survive native SIGSEGV crashes in nightly batches, Poppler-to-PyMuPDF migration (~17x faster document processing), ~90% reduction in CRM API calls via caching, and direct token streaming that eliminated a duplicate LLM inference per plan.
- Designed and built the provider-facing web app (Next.js, TypeScript, Tailwind CSS, TanStack Query): feedback-management dashboard, AI-insights analytics, scheduled-plan review with one-click EMR write-back, and a WYSIWYG treatment-plan editor (TipTap) with locked clinical sections and version history — authored ~30 of the app's 43 components.

Shared infrastructure and integrations

- Created bvital-core, a shared HIPAA-compliant Python library (unified LLM client, PHI scrubbing, audit logging, metrics, async SQLAlchemy ORM layer with ~20 models) consumed by two production services via AWS CodeArtifact — 187 tests and CI-gated CalVer releases eliminated cross-service code duplication.

- Built and own a serverless CRM-to-EHR integration service (FastAPI on AWS Lambda, deployed with SAM): automates PDQ-39 Parkinson's questionnaire scoring into clinical encounters, processes lab-result webhooks behind three authentication schemes, and ships with an 18-metric CloudWatch monitoring dashboard.
- Founded the patient-portal frontend later built out by a five-engineer team: scaffolded the Next.js 16 / React 19 application, design system (shadcn/ui + Radix), PWA setup, strict Content Security Policy, and AWS Amplify CI/CD pipeline.

Leadership and engineering standards

- Drove engineering standards across four repositories — introduced PR templates, CONTRIBUTING/SECURITY policies, ruff lint migration, and CI quality gates; own dev-to-main release integration (39 release merges); mentor an intern and junior engineers through code reviews with change-request authority.
- Authored client-facing technical decision documents, including the AWS account-ownership and HIPAA Business Associate Agreement transition guide; earned a company Rewards & Recognition award (Mar 2026).

Software Engineer — SidsFarm Pvt. Ltd. | Hyderabad, India | Jul 2024 – Feb 2025

- Built an interactive logistics map with the Google Maps API — driving/aerial distance calculation and Places Autocomplete for real-time location input — improving delivery-address accuracy (React, Redux, Django, PostgreSQL).
- Shipped multiple feature enhancements across a React + Django stack and conducted peer code reviews to maintain code quality and share knowledge.

Associate Software Engineer — CodeParva Technologies Pvt. Ltd. | Bengaluru, India | Jan 2023 – Jul 2024

- Led end-to-end development and deployment of the Vehicles module in a storage-management product, using Redux for API state and Formik FieldArray for dynamic form management.
- Developed and maintained 80+ client websites on an internal low-code builder — React with Strapi headless CMS and Netlify CI/CD with per-site environments.
- Centralized typography across the website-builder framework, eliminating per-site styling work and standardizing UI across all client sites.

Reliability Engineer — Cyient Pvt. Ltd. | Bengaluru, India | Jun 2021 – Aug 2022

- Analyzed aircraft incident reports and maintained reliability data in SAP before transitioning into software engineering.

EDUCATION

Bachelor of Science, Aircraft Maintenance Engineering — Telangana State Aviation Academy (Jun 2017 – Mar 2021)

CERTIFICATION

Full Stack Development — GUVI Geek Network Pvt. Ltd. (Feb 2022 – Nov 2022) — [Certificate](#)