

Saurabh Kamboj Software Engineer

github.com/saurabhkamboj | linkedin.com/in/saurabhkam | saurabhkamboj@icloud.com | Bengaluru, KA

Experience

Co-Creator and Software Engineer, *Chunkwise*

10/2025 – Present

Chunkwise is an open-source, self-managed platform for evaluating chunking strategies and deploying ETL pipelines for RAG applications. It supports experimentation with visualization and corpus-specific benchmarking, alongside deployment of document processing pipelines within a single cohesive system.

- Implemented a distributed backend that isolates compute-heavy chunking and evaluation into containerized services, coordinated through a RESTful orchestration API with shared Pydantic models. (Python, FastAPI, ECS)
- Adapted Chroma's token-level metrics framework for evaluating chunking strategies, enabling corpus-specific benchmarking with reusable, LLM-generated ground-truth references. (ChromaDB, OpenAI API)
- Designed and implemented the client interface for experimentation and deployment experience, supporting real-time visualization optimized using debouncing, trackable deployment state using SSE, local caching of reusable metadata, and runtime request/response validation. (TypeScript, React, Server-sent Events, Web Storage API, Zod)
- Prototyped the ETL pipeline handling RDS provisioning, secrets management, and S3 ingestion with real-time deployment state streaming via SSE; designed secure, local-only credential retrieval using ARNs. (Boto3, PostgreSQL, pgvector)
- Scaled the ETL pipeline to parallelize document processing while managing external API rate limits to prevent throttling. (AWS Batch (Fargate), OpenAI API)
- Implemented client and server side encryption of S3 credentials using ephemeral, in-memory RSA key pairs to avoid plaintext transmission in a non-TLS, locally run client setup. (cryptography, Web Crypto API)

Software Engineer, *Self employed*

01/2023 – 06/2025

- Request Bin: Developer tool for receiving, persisting, and inspecting HTTP requests in real time via a generated URL endpoint. (JavaScript, Express, React, PostgreSQL, MongoDB, WebSockets, AWS)
- Wine-Review Chatbot: Chat loop using structured outputs and RAG to enable semantic and keyword-based wine search and recommendations over a 130k-review dataset. (Python, PostgreSQL, pgvector, OpenAI API, MemO)
- Forum: Message board with a simple database design. (Ruby, Sinatra, PostgreSQL)

Solutions Manager, 'Cue for Good' Solutions

01/2021 – 06/2022

Led cross-functional development of digital commerce stores for early-stage brands with ARR up to \$5M, collaborating with designers, developers, and quality analysts. Built internal tools for ads reporting (using Facebook's Graph API), billing, and revenue, and tracked sales performance to inform ongoing optimization and scaling efforts.

Co-founder, *dontknow.health*

03/2020 – 12/2021

Designed and built a health information platform combining a storefront and learning management system to distribute micro-courses authored by health professionals. Collaborated with a fully remote team distributed across Nepal, India, Brazil, and Germany. Created both educational resources and medical illustrations covering 17 health topics. Maintained an average of 62+ weekly active users over five months and acquired 30+ paying customers. (WordPress, HTML, CSS, Stripe API)

Head of Platforms and Marketing, India, AIESEC

06/2018 – 05/2019

Volunteered since 2015; selected for the national team in 2018. Led a multidisciplinary team of designers, developers, and marketers to build and manage digital platforms, landing pages, and marketing campaigns for a global non-profit operating across more than 30 cities in India.

Built internal tools to support operations, including cost calculators, lead planners, and progress trackers, and led the integration of digital payments and digital contracts. Implemented GDPR compliance for India across national digital platforms. Drove growth in key operational KPIs across two focus products, and received an internal global award.

Skills

Languages and Frameworks — TypeScript, JavaScript, Node.js, Express, Ruby, Sinatra, Python, SQL, React, HTML/CSS

AI Tooling — OpenAI API, LangChain and Chonkie, Context Engineering, RAG, Vector Databases, Embeddings, Claude Code

Cloud — AWS infrastructure design (ECS, EC2, Lambda, Batch, RDS, DocumentDB, S3, Secrets Manager, CDK), Docker

Other — RESTful APIs, PostgreSQL, MongoDB, Jest, Vitest, minitest, WebSockets, SSE, Nginx, HTTP/S, Git/Github