

Gokul JS

jsgokul123@gmail.com | <https://gokuljs.com/> | <https://github.com/gokuljs> | <https://www.linkedin.com/in/gokul-js/>
https://twitter.com/gokul_js029

PROFILE

Software engineer with startup experience building and operating production systems used by real customers. Owned end to end development of real time voice infrastructure, backend platforms, frontend applications, and LLM powered tools that supported core customer workflows at early stage and YC backed companies. Ran revenue critical systems, owned on call response, and worked directly with customers to resolve failures, improve reliability, and influence product decisions.

PROFESSIONAL EXPERIENCE

Rime | Senior Software Engineer 07/2025 – 02/2026 | San Francisco, CA (USA)

- Built and owned end-to-end real-time voice pipelines for Rime Text-to-Speech using LiveKit and Pipecat plugins over WebRTC.
- Integrated Rime TTS plugins across LiveKit and Pipecat to deliver low-latency, production real-time audio for customer use cases.
- Diagnosed and fixed latency, reliability, and failure issues across the WebRTC stack in LiveKit- and Pipecat-based deployments.
- Resolved customer-facing issues by making coordinated frontend and backend changes required for real-time audio stability.
- Served as on-call owner for real-time audio incidents and breaking changes, shipping fixes to prevent regressions.
- Built and supported customer-facing voice agents on top of these systems, driving production adoption and revenue.

Teamble | Lead Engineer 06/2024 – 03/2025 | New York City (NYC), USA

- Architected and maintained high performance web applications using Node.js and React
- Designed and implemented a frontend design system from scratch, enabling faster and more consistent feature delivery
- Optimized backend APIs by refactoring logic and migrating inefficient SQL queries, significantly improving response times
- Built LLM powered agents using LangGraph for internal and customer facing workflows
- Developed a conversational AI interface with role based permissions and real time data access
- Implemented a multi agent RAG based system for scalable performance review generation
- Built a One on One AI agent that aggregates performance and quarterly review data to generate actionable insights for managers

Aerotime, Y Combinator (W21) | Founding Engineer 11/2022 – 02/2024 | San Francisco, California

- Spearheaded the complete management and enhancement of frontend infrastructure, significantly improving web application performance.
- Contributed to the development of a JavaScript SDK, facilitating third-party integrations and expanding application functionalities.
- Engaged directly with customers through discovery calls and feedback sessions, driving user-centric product development.
- Communicated critical user feedback to company founders, influencing strategic decision-making and problem-solving.
- Deployed and managed open-source large language models like LLaMA and Stable Diffusion, enhancing the company's AI capabilities.
- Employed advanced techniques from the Hugging Face library, including DreamBooth and Textual Inversion, to optimize AI model performance.
- Implemented comprehensive analytics and telemetry systems to analyze user interactions, leading to improved UI/UX design and user flow.
- Developed a sophisticated table with virtualization, efficiently managing large datasets while maintaining performance and responsiveness.

EDUCATION

CMR Institute Of Technology, B.E Computer Science and Engineering

2017 – 2021 | Bangalore, India

SKILLS

Languages: JavaScript, TypeScript, Python

Frontend: React JS, Next JS, SolidJS, Redux, React Query, Tailwind CSS, Material UI, Recharts, Chart.js, Vite, Zustand, Bootstrap, Gatsby JS, shadcn/ui, GSAP, Radix, Jotai, Storybook, HTML5, CSS, SASS.

Backend: Node.js, Express, GraphQL, SQL, Prisma, Redis, Django, Fast API, tRPC, gRPC, ElasticSearch, Logstash, Kibana, Suricata, Zeek, Bull MQ, TypeORM, Drizzle, Clerk(Auth)

Databases: PostgreSQL, MySQL, MongoDB, DynamoDB,

Large Language Models and AI: LangGraph, LlamaIndex, Langchain ,Retrieval-Augmented Generation (RAG), Stable Diffusion, Prompt Engineering, Prompt Evaluation, ComfyUI, AUTOMATIC1111, IP-Adapter.

Cloud and DevOps: AWS (S3, ECR, EC2, etc...), Docker, GitHub CI/CD, Vercel (serverless)

Realtime: LiveKit, Pipecat, WebRTC, Twilio(SIP+SDK)