

Kevin Shreve

GENERATIVE AI ENGINEER

kshreve@gmail.com · (810) 333-3117

SUMMARY

Generative AI Engineer shipping production LLM features end-to-end in TypeScript. 10+ years of full-stack engineering experience, now focused on building RAG systems, tool-using agents, and AI-powered product surfaces with end-to-end ownership from prompt and retrieval through React/Next.js UI to deployment on Vercel. Comfortable across the stack: TypeScript on both ends, Postgres + pgvector for retrieval, OpenAI APIs and the Vercel AI SDK for inference and tool use.

KEY TECHNOLOGIES

- **AI & LLMs:** OpenAI (GPT-4 / 4o / o1, embeddings), Vercel AI SDK, tool use / function calling, agents, Retrieval-Augmented Generation (RAG), prompt engineering, Claude Code, Warp
- **Languages:** TypeScript, JavaScript (ES2015+), HTML5, CSS/Sass
- **Frontend:** React, Next.js (App Router, SSR, ISR, SWR), Tailwind CSS, styled-components, React Native
- **Backend:** Node.js, Next.js API routes / Server Actions, REST APIs, WebSockets, Sequelize
- **Data & Retrieval:** Postgres + pgvector, SQL Server, MongoDB
- **Deployment & Infra:** Vercel (Functions, deployments, env mgmt), Docker, AWS, Buildkite
- **Tooling & Testing:** Git, Cypress, Webpack, JIRA

EXPERIENCE

First Harmonic, Generative AI Engineer

Remote · January 2025 – Present

- Build and ship production AI features end-to-end in TypeScript, owning prompt design, retrieval pipelines, React/Next.js UI, and Vercel deployment.
- Architected a Retrieval-Augmented Generation (RAG) chat application for internal knowledge Q&A and external-source research, built on OpenAI GPT-4/4o, OpenAI embeddings, and pgvector on Postgres.
- Built a custom note-taking application with tool-calling agents that reconcile deal context across **Attio** (CRM) and **Circleback** (meeting transcripts).
- Designed a custom rubric scoring system using structured LLM outputs to surface quantitative signal on potential deals from unstructured inputs.
- Built streaming chat surfaces with the **Vercel AI SDK** (streamText, generateObject, tool calling), deployed on Vercel Functions.
- Own the full deployment pipeline: preview deployments, environment management, and production rollout on Vercel.
- Use **Claude Code** and **Warp** daily for agentic feature scaffolding, AI-assisted code review, and AI-generated tests / bug repros.

Caesars Digital, Staff Software Engineer

Remote · April 2022 – January 2025

- Owned the betslip in a React + TypeScript / Redux-Saga codebase, refactoring it to support a wide variety of bet types in a high-traffic, highly regulated environment.
- Shipped **wager-to-win** and **quick-bet** features, cutting time-to-place-bet by ~3 seconds and reducing clicks from 7 to 4.
- Led the cross-platform mobile migration from Cordova to Capacitor.
- Configured jurisdiction rollouts (Kentucky, D.C., Mississippi) behind feature flags with automated test coverage.

- Introduced Tailwind into the company design system; upheld standards within the redux-saga toolkit.

Yello (WayUp), Staff Software Engineer

Remote · November 2017 – January 2024

- Led the frontend platform migration from legacy AngularJS to SSR React, substantially improving TTFB and FCP.
- Stood up the Next.js platform that still powers production today.
- Unified 13 separate platform navigation bars into one consistent cross-product navigation.
- Built shared bit.cloud components and DX tooling for production-parity local environments.
- Mentored an offshore team to ensure delivery quality and unblock execution.

EARLIER EXPERIENCE

- **FlashRecruit**, Senior Software Engineer · Remote · Aug 2016 – Nov 2017. Real-time WebSocket chat connecting candidates and recruiters; shipped iOS and Android apps in React Native.
- **Trinity Health**, Senior Software Engineer · Livonia, MI · Apr 2016 – Aug 2016. EmberJS and React / D3.js applications for bad-debt recovery and patient-cost analytics.
- **MeridianRx**, Senior Software Engineer · Detroit, MI · Aug 2015 – Apr 2016. Greenfield ES6 / React application integrating RESTful services.
- **Optum**, Senior JavaScript Developer · Southfield, MI · Jan 2015 – Jul 2015. Responsive AngularJS SPA integrating with CQ/AEM6 and REST services.
- **Urban Science**, Associate Software Developer · Detroit, MI · Jun 2012 – Jan 2015. Full-stack web work (HTML/JS/CSS/C#); built AJAX criteria tooling for Hyundai Motors.
- **Sircon, a Vertafore Business**, Software Development Intern · Okemos, MI · Oct 2011 – Mar 2012.

EDUCATION

B.S., Computer Science, Michigan State University, East Lansing, MI 3.59 GPA · 3.51 Technical GPA