

SHENGGE (DEMPSEY) ZHANG

Toronto, Ontario | 778-363-3748 | zsg258852@gmail.com | [GitHub](#) | [LinkedIn](#) | [Portfolio](#)

SUMMARY

Full-stack AI product engineer who ships AI agents into real products: a voice agent students talk to, a reporting agent that safely queries production data, and now a retrieval assistant for counselors, guided by evals. 560+ merged PRs on an AI counseling platform used by 2,000+ students at 20+ schools; project lead on 11 projects, from data model to deploy.

TECHNICAL SKILLS

AI / LLM: agents and tool calling (LangGraph, LangChain), voice agents (OpenAI Realtime, WebRTC), text-to-SQL, RAG (Qdrant), evals (LLM-as-judge, recall tests), guardrails, tracing, AssemblyAI

AI-native workflow: Claude Code and coding-agent skills authored for the team

Front-end: TypeScript, React, Next.js (App Router, Server Actions), Tailwind CSS, ShadCN, Radix UI, Redux

Back-end: Python, FastAPI, Celery, Redis, Node.js, WebRTC, Twilio, REST

Data and Infra: PostgreSQL, Prisma, Qdrant, Docker Swarm, GitHub Actions, Turborepo, GCP, Sentry, Vercel

Testing: Jest, React Testing Library, Pytest, Playwright

EXPERIENCE

Addie (getaddie.com)

Jan 2025 – Present

Full-Stack Developer (AI Product)

Remote, Toronto

- Merged 560+ PRs across a Python backend and three Next.js apps for an AI counseling platform serving 2,000+ students and 70+ active counselors at 20+ schools; **project lead on 11 projects** (6 shipped); top contributor to the repo (451 commits, Jul to Oct 2026).
- Agentic RAG assistant for counselors (lead):** counselors ask about any student and get sourced answers from meetings, notes and files. Built a LangGraph agent with 9 retrieval tools, streamed answers with sources, and server-side data scoping. Steered the design with a recall eval: lifted recall from 3% to 43% of planted facts (current chat: 67%) using a third of its tokens.
- Real-time voice agent (lead):** students talk to Addie to log college visits; the agent stays on goal, wraps up on time and saves structured notes to the student's profile. OpenAI Realtime API over WebRTC with guarded tool calls; fixed echo that made the agent hear itself on speakerphone.
- Natural-language reporting agent (lead):** school admins ask questions in plain English and the agent queries production data. Made it safe to ship: read-only, scoped to the admin's school, SQL guardrails and adversarial tests that keep student transcripts out of reach, every query audited.
- AI conversation workflows:** counselors design guided conversations that Addie runs with students over web, SMS, WhatsApp and voice (13,000+ assigned). Built the builder, multi-step journeys, reminders, and an LLM-as-judge scoring system calibrated against human labels.
- Meeting transcription (lead):** counselors record meetings and get speaker-labeled transcripts and summaries. Chose AssemblyAI after a vendor test (99.4% accuracy, half the cost) and built the consent flow state recording laws require. Live at 5 schools.
- AI student products:** built college search and profiles with an Ask Addie panel; leading Writing Studio, an AI essay workspace grounded in each student's own records.
- Reliability:** found deploys reporting success while running 3-day-old code; rebuilt the pipeline with verification, real rollback and alerts, and cut the slowest CI build from 13.7 to 6.6 min.
- Security and privacy:** fixed a critical admin login bypass; enforced school-level data isolation (FERPA) with audit logs; kept student PII out of error tracking.
- Platform:** led the Turborepo monorepo migration and student and counselor UI revamps.

Vosyn.ai

Sep 2024 – Jan 2025

Full Stack Engineer Intern

Remote, Toronto

- Led the homepage in React; built Redux search filtering (25% faster) and moved search results to Next.js SSR, cutting page load time by 30%.

EDUCATION

University of Toronto | *Master of Industrial Relations and Human Resources*

Sep 2018 – May 2020

University of British Columbia | *Bachelor of Management*

Aug 2014 – May 2018