

Christopher Souza

christopherlitobu@gmail.com • 213-205-0153 • Billerica, MA 01821 • in/chrystophersouza7

Senior Software Engineer with 7+ years of experience delivering scalable backend systems and AI-driven workflows. Led development of a cloud-native advertising platform that cut time-to-first-ad by 6x and reduced campaign setup errors by 70%, and built a multi-agent OpenAI orchestration that shortened creative production from hours to minutes. Skilled in API design, cloud services (AWS, Azure, GCP), and AI/ML integration, with a track record of mentoring engineers and driving high-performance, cost-effective engineering solutions.

Skills

AI Automation & Agents: Agentic AI | Agentic Applications | Multi-Agent Orchestration | Agentic Workflows | AutoGen | AI Agent Development | AI Workflow Automation | AI-Assisted Software Development | AI-Powered Process Automation | Applied LLM Engineering

AI/ML Tools & Platforms: LLMs, RAG Pipelines | LangChain | LangGraph | LangSmith | Langfuse | OpenAI | Claude | Anthropic LLMs | Generative AI | Multimodal AI | Text-to-Image | Text-to-Speech | Speech-to-Text | Deepgram | LiveKit | n8n | Embeddings | Vector Databases | Cursor | Claude Code | Prompt Engineering | AI Integration | Pinecone

Programming Languages & Frameworks: Python | JavaScript | TypeScript | Java | React | Next.js | Node.js | FastAPI | Flask | SQL | Nuxt.js | Django | Vue.js | PostgreSQL | GraphQL | C# | .NET | Spring Boot

Professional Experience

Senior AI Software Engineer

11/2023 – 06/2026

Laptop Freedom

Orlando, FL

- Architected the Rocketeer ads platform as a scalable software system using Java Spring Boot microservices and domain-driven data modeling on PostgreSQL, onboarding thousands of non-technical entrepreneurs and delivering 6x faster time-to-first-ad with sub-10-minute campaign launches.
- Integrated the Facebook Marketing API and Mapbox to enable end-to-end campaign management with geographic targeting, reducing campaign setup errors by 70% and eliminating reliance on Meta Ads Manager.
- Built an n8n workflow orchestrating 5-7 OpenAI agents with AutoGen and prompt engineering to automate ad copy and creative generation using Nano Banana text-to-image, cutting production time from hours to under 2 minutes.
- Applied LLM engineering to evaluate emerging models such as OpenClaw and NemoClaw with LangSmith, benchmarked ad quality using Claude Code and LangGraph, and defined model-selection criteria that guided the platform's long-term AI investment strategy.
- Built containerized infrastructure on MCP servers with CI/CD pipelines using Docker, Terraform, and GitHub Actions, reducing deployment time by 80% and supporting weekly releases with near-zero production incidents.
- Mentored junior developers through code reviews, pair programming, and architecture guidance on Node.js and React projects, cutting new-hire ramp-up from one month to one week while lowering bug rates and boosting team velocity.

Full Stack Engineer

05/2021 – 10/2023

RethinkFirst

New York, NY

- Built and optimized backend services and RESTful APIs with C#, .NET, and SQL, enabling high-throughput, low-latency financial analytics and secure authentication that reduced processing delays and safeguarded student data.
- Led full-stack development and defined the architecture for AI-powered features in education applications using React, Vue.js (Nuxt.js), Next.js, Python, TypeScript, and GraphQL, streamlining learning workflows and improving student-instructor communication.
- Implemented multi-threaded backend services using FastAPI and WebSockets to support real-time voice streaming, speech-to-text transcription, and concurrent audio sessions at production scale for live classroom environments.

- Expanded generative AI learning tools using Anthropic LLMs, the OpenAI SDK, and Cursor, applying agent development to reduce educator preparation time by up to 45% through automated lesson planning and content summarization.
- Built multimodal, voice-enabled AI learning assistants for student onboarding, real-time lecture transcription, and conversational tutoring using Whisper, Deepgram, LiveKit, and custom LLM pipelines, driving faster onboarding and higher engagement.
- Designed vector database architectures and RAG systems with large-scale multimodal data handling and embeddings to deliver context-aware learning insights while maintaining FERPA compliance.
- Deployed production AI systems on AWS SageMaker and Bedrock using Terraform across Azure and GCP, reducing inference costs by 40% while scaling to millions of monthly requests.
- Built advanced summarization systems generating lecture notes, study guides, and session summaries, reducing instructor documentation time by 40% across 10,000+ monthly learning sessions.
- Established automated workflows in Nuxt.js using LLM-driven agents to triage student inquiries, extract learning insights, and streamline instructor documentation, integrating CRM data via Docker to cut administrative effort and accelerate grading and feedback.

Full Stack Developer

06/2019 – 04/2021

Axiom Learning

Cambridge, MA

- Engineered full-stack applications with React, Next.js, Node.js, and Python that modernized learning management and academic operations, speeding up course setup and reducing administrative workload.
- Delivered custom software that improved enrollment forecasting accuracy by 15%, streamlined tracking of 5,000+ learning resources across multiple campuses, and strengthened program oversight for multi-site institutions.
- Designed scalable backend services on AWS using C# and .NET Core with multi-threaded architecture to securely handle FERPA-aligned student data, delivering 99.9% uptime.
- Built responsive interfaces with Tailwind CSS that improved data-visualization clarity for academic dashboards and reduced new-user onboarding time by 40%.
- Established CI/CD pipelines with GitHub Actions to automate testing, code review, and deployment, cutting release cycles from weeks to days and improving code quality across distributed teams.
- Collaborated with UX/UI designers through structured design reviews and Figma handoffs, translating pedagogical requirements into WCAG-compliant React interfaces that increased student and instructor engagement.

Education

Bachelor's degree

08/2015 – 05/2019

University of Massachusetts Lowell

Lowell, MA