
SUMMARY

I build AI agent systems that run in production, not demos. My agent architecture starts with a question most teams skip: what happens when this fails at 2 AM, and who's responsible for it?

AI Agent Core: AI Agent Orchestration, Multi-Agent Systems, Context Engineering, Prompt Design, LLM Evaluation, Guardrail Architecture

Platform & Scale: Distributed Systems, Real-Time Processing, API Design, Observability, Load Testing, Reliability Engineering

Technologies: Python, TypeScript, AWS, Redis, Docker, Kubernetes

EXPERIENCE

Staff Software Engineer | AI Systems Architect, Unicity International | Provo, Utah

Jan 2025 – Jan 2026

Tech Stack: GenAI, Python, LLMs, Context Engineering, MCP, Multi-Agent Systems, AST Analysis

- **Multi-Agent Production System:** Engineered 5-specialized-agent platform with human-in-the-loop checkpoints. 289 migrations at <10 minutes each, 96% faster than manual process.
- **Parallel Agent Orchestration:** 34 AI agents running simultaneously with deterministic validation boundaries and full audit trails. 90% error reduction, 85-90% automated code generation.
- **AI Guardrail Architecture:** Designed Truth Source Architecture and Context Engineering pipeline. Agents consume structured documentation instead of pattern-matching. Zero hallucination in production outputs.
- **Ambiguity Navigation:** Reverse-engineered compiled legacy system without source access using AI-assisted bytecode analysis, producing 45+ specs (50,000+ lines) at 99%+ accuracy.

Engineering Manager / Technical Lead, Unicity International | Provo, Utah

Nov 2023 – Jan 2026

Tech Stack: AWS, Redis, DataDog, Kubernetes, Docker, API Architecture, Load Testing

- **High-Throughput Migration:** Led 289 API endpoint migration achieving 96% production traffic, 34% faster response times, 4x throughput (300 to 1,300 orders/5min) at 99.99% uptime.
- **Observability Platform:** Built real-time debugging system (Redis/Xdebug/BPF) reducing MTTR 95%, from hours to minutes. Adopted across engineering voluntarily.
- **Cost-Efficient Scale:** \$336K+ annual savings while improving reliability. Monitoring costs reduced 62.5% (\$12K to \$4.5K/month) through DataDog optimization.
- **Load Confidence:** Led company's first load testing program — 5 sessions, zero failures, 85-90% peak traffic confidence established.

Principal Engineer, ButcherBox | Remote

Feb 2022 – Oct 2023

Tech Stack: OpenAI, GPT-3, Azure, Kubernetes, Event-Driven Architecture, React

- **AI Pioneer:** Deployed GPT-3 in production (2022, pre-ChatGPT) automating support workflows. ~2 hrs/week/employee saved, NPS 57. One of earliest enterprise AI deployments.
- **Scale Architecture:** Platform processing \$2B+ transactions, scaled 100x over four years. 50% month-over-month volume spikes handled without incidents.
- **Cross-Team Delivery:** Directed 6-month Instacart flagship integration across engineering, product, and executive stakeholders.

Senior Lead Software Engineer, ButcherBox | Boston, Massachusetts

Jul 2019 – Feb 2022

Tech Stack: Azure, Kubernetes, Microservices, Event-Driven Architecture, Docker

- **Production Resilience:** Systems handled 10x COVID traffic surge with zero scaling incidents while competitors went down.
- **Engineering Standard:** Built platform recognized as internal benchmark for Testing, Stability, and Innovation.
- **Rapid Impact:** 87% application performance improvement in first 90 days.

VP of Engineering & Lead Systems Architect (Founder), O2O HQ | O2O Global Marketplace | American Fork, Utah

Jun 2015 – Nov 2018

Tech Stack: AWS, Docker, Distributed Systems, Multi-Region Architecture

- **Multi-Region Architecture:** Distributed systems spanning AWS (US) and Aliyun (China), solving Great Firewall latency transparently. The cross-channel thinking voice/SMS/chat platforms need.
- **Unified Integration:** 19+ payment gateways behind a single API. Integrate once, works everywhere. The pattern multi-channel agent platforms require.

Principal Engineer, Janeiro Digital | Boston, Massachusetts

Jan 2019 – Apr 2019

Tech Stack: Node.js, React, Real-Time Systems

- **Project Rescue:** Recovered failing insurance fintech 6 months behind schedule through rapid prototyping and decisive technical leadership.
- **Tim Berners-Lee Task Force:** Selected for Solid Protocol team building decentralized data systems. Trust models for decentralized data map directly to the trust infrastructure customer-facing agents require.

EDUCATION

Computer Software Engineering

Utah Valley University

2010 - 2013

Associate's Degree, Pre-Professional Associate in Science

Utah Valley University

2007 - 2009

PATENTS

Optimized Caching and Updating of Custom Assets RSS Algorithm (2012)

Process Patent for a native asset caching system that significantly improved load speeds across fragmented mobile platforms.

HONORS & AWARDS

Eagle Scout + 3 palms | New Century Scholarship | Dean's List