

STEVEN TIKAS

Jacksonville, FL • 904-892-0025 • stevenmtikas@gmail.com
linkedin.com/in/steven-m-tikas • steventikas.online

AI SOLUTIONS ENGINEER

Summary

Builder of production AI applications end to end — LLM integration, multi-agent orchestration, retrieval, and the full-stack products wrapped around them — backed by 20+ years of professional experience across technical support, customer service, and manufacturing. Self-directed portfolio spans five model providers (OpenAI, Anthropic, Google Gemini, DeepSeek, and local inference via Ollama), schema-validated structured output, tool calling, RAG over vector search, and cost-controlled agent pipelines. Equally comfortable debugging a failing integration and explaining it in plain language to a non-technical stakeholder. B.S. in Information Technology; portfolio at steventikas.online.

Technical Skills

AI & LLM: OpenAI, Anthropic Claude, Google Gemini, DeepSeek, local models via Ollama; CrewAI multi-agent orchestration; structured output with response schemas; tool / function calling; RAG with vector embeddings (sqlite-vec, nomic-embed-text); prompt layering; model routing, caching, and cost control

Languages: TypeScript, JavaScript, Python, SQL, Shell / PowerShell

Web & App: React 18/19, Next.js 15, Django, FastAPI, Node.js, Tailwind CSS, React Native (Expo)

Data & Cloud: PostgreSQL, SQLite, Firebase (Auth, Firestore, Cloud Functions, Hosting), Microsoft Azure, AWS Textract, Render

Testing & Quality: Vitest, Cypress, pytest, golden-fixture regression harnesses, TypeScript strict mode, ESLint

Security & Ops: Role-based access control, AES-256-GCM encryption at rest, OAuth 2.0, session authentication, SSO / MFA, encrypted backup and restore, Active Directory

Selected Technical Projects

Worxstance — AI-Powered Career Workspace React, TypeScript, Firebase, Gemini

- Architected and built an AI-powered SaaS platform spanning eight of nine planned modules — job discovery, resume tailoring, skill-gap analysis, and interview prep among them — organized under a Feature-Sliced Design architecture.
- Drove every AI feature through Gemini 2.5 Flash with defined response schemas, so model output parses reliably as structured data instead of free text.
- Integrated the Adzuna and Jooble job APIs plus Google Search grounding into one match-scored feed with multi-factor scoring across skills, experience, role, and location; fully built, but currently disabled pending job-board legal clearance.
- Metered usage through Firestore-backed AI credit counters, and enforced per-user data isolation with Firebase security rules; instrumented with GA4 and PostHog.

Stack: React 19, TypeScript, Vite, Tailwind, Firebase Auth/Firestore/Functions, Gemini 2.5 Flash, Vitest, Cypress

SpecAnalyzer — Construction Drawing & Specification Takeoff Next.js, FastAPI, Claude, AWS Textract

- Built a document-AI pipeline that ingests architectural drawing sets, classifies every page with PyMuPDF, extracts window and door schedules via camelot with an AWS Textract fallback on low-confidence pages, and computes glazing quantities deterministically.
- Deliberately kept the language model out of the math: Claude reads rasterized elevation sheets, but quantities derive only from parsed schedule rows — unparseable dimensions stay null rather than being guessed.
- Attached source sheet, page index, bounding box, and confidence score to every extracted row so a human can verify any number back to its origin in the drawing set.
- Designed a per-trade profile registry so additional trades plug in without touching shared pipeline code; protected accuracy with a golden-fixture regression harness run against sample drawing sets.

Stack: Next.js orchestrator, Python FastAPI sidecar, PyMuPDF, camelot, AWS Textract, Anthropic Claude, pytest

Tikas Family Digital Home — Self-Hosted AI Platform Next.js 15, TypeScript, SQLite, Ollama + Claude

- Designed and delivered a nine-phase platform spec-first — module specifications, data model, decision log, and phased milestones written before implementation — covering calendar, tasks, documents and photos, search, notifications, administration, and an AI assistant.
- Built hybrid model routing: questions requiring private household data run entirely on local Ollama models so that data never leaves the network, while general questions route to Claude behind an explicit preview-and-confirm step showing exactly what would be transmitted.

- Implemented semantic search with sqlite-vec and local embeddings, role-based permissions with per-folder grants, AES-256-GCM encryption at rest, and encrypted backup and restore built on SQLite's online backup API.
- Shipped an Android companion app (Expo / React Native) that syncs photo metadata first and fetches full files on demand; 433 automated tests across the codebase.

Stack: Next.js 15, React 19, TypeScript, better-sqlite3, sqlite-vec, Ollama, Anthropic SDK, Expo, Vitest

YouTube Product Crew — Multi-Agent Content Pipeline CrewAI, Python, OpenAI / DeepSeek / Gemini

- Orchestrated parallel research agents across three model providers with a synthesis manager agent, producing scripts, SEO keywords, titles, and descriptions from a single product input.
- Cut repeat API spend with a configurable research cache (7-day default expiry) and tracked per-run generation cost, including automated DALL-E 3 product imagery at a known unit price.
- Made agents and tasks fully YAML-configurable, and added a Gradio management interface, a benchmark suite, and complete operator documentation.

Stack: Python, CrewAI, OpenAI GPT-4, DeepSeek, Gemini, DALL-E 3, Gradio, uv

Professional Experience

Technical Support Analyst — IT Support Center (Remote) Aug 2021 – Oct 2024; Mar 2025 – Present

- Resolve escalated technical issues across Windows systems, hardware, and business-critical software for internal and external users; consistently recognized as a top performer on customer satisfaction scores.
- Serve as Senior Technical Advisor — escalation point for complex cases, and trainer for new hires joining the support team.
- Administer user access, rights, and permissions across multiple systems, applying least-privilege practices to every provisioning request.
- Handle 15–35 support contacts per day across phone, chat, and email, translating technical root causes into plain language for non-technical users.

Technical Support Analyst — Availity (contract via Insight Global) Oct 2024 – Feb 2025

- Supported healthcare providers and payers on a national claims platform, diagnosing and resolving user issues across the insurance claim lifecycle.
- Provisioned and maintained secure system access, rights, and responsibilities for external customer organizations.
- Ranked a top performer in both resolution volume and average handle time at 25–40 contacts per day.

Customer Support Technician — Bank of America (contract via Diversant, LLC) 2017 – 2018

- Delivered advanced end-user support to a 20,000+ employee population: connectivity troubleshooting, hardware and software installation, and system upgrade rollouts.
- Ranked first of fifteen team members on overall performance KPIs; earned multiple “Top 3 Technician” awards.

Customer Service Representative — Lee & Cates Glass, Inc. 2007 – 2017; 2018 – 2019

- Primary point of contact for the accounts representing over 50% of company sales, coordinating quoting and installation scheduling across sales, plant, and service teams.
- Managed multi-million-dollar commercial glazing projects alongside high-volume smaller orders at an order error rate under 1%.

Education & Certifications

Bachelor of Science, Information Technology — University of Phoenix

Certificate in Advanced Business Analytics — University of Phoenix

United States Navy — Nuclear Machinist's Mate

Certifications: Microsoft Azure Fundamentals (AZ-900) · Google Cybersecurity Certificate · Advanced Business Analytics · Foundations of Project Management · Agile and Scrum Development · CompTIA Security+ (in progress)