Jason Sutter

Applied AI Engineer

Sacramento, CA · Open to SF Bay, Austin, Boston, or Remote
jasonsutter87@gmail.com · linkedin.com/in/jasonsutter87 · github.com/jasonsutter87 · hackerone.com/mrblackkeys

Engineer with 7+ years shipping production software who now builds mostly with and on LLMs: agents with tool calling, RAG, evals, LoRA fine-tuning, and local inference with llama.cpp and whisper.cpp. I care about making model output trustworthy, with retrieval that cites its sources and declines when it doesn't know, and evals that decide what ships. Background in Angular, React, TypeScript, Rust, and Python, plus independent security research (accepted Twilio bug bounty).

Skills: LLM apps · RAG · Evals · Agents & tool calling · MCP · LoRA fine-tuning · llama.cpp · whisper.cpp · Claude API · AI coding agents (Claude Code) · Python · TypeScript · Rust · React · Angular · Node.js · Tauri · Docker · Testing (Vitest/Jest) · CI

SELECTED AI PROJECTS

- **Werbos**: a local AI assistant I built from scratch (2,100+ commits): agent loop with tool calling, RAG, memory, an eval harness, and LoRA fine-tuning. Runs entirely on open-weight models, with no hosted LLM calls.
- **Murmur and Hablo**: voice AI apps in Rust. Push-to-talk dictation and an offline Spanish tutor running Qwen 3B via llama.cpp with streamed speech.
- **Unfold**: an LLM pipeline engine with YAML recipes and resumable, editable runs. CLI, library, and HTTP API.
- **WRITEpad**: a Mac and Windows script-writing app (Tauri 2, Rust) with A/B drafts, local narration and dictation, and an MCP server so AI agents can edit manuscripts.
- **Tallybook**: a live bookkeeping app (Rust, React, Tauri) with bank sync, smart categorization, and tax prep.

EXPERIENCE

GROW — Web Developer → Frontend Developer

Auburn, CA · Oct 2019 – Oct 2026

- Angular developer on **Mandated Reporter Training**, a live, state-approved compliance-training LMS serving **5,000+ organizations** nationwide, with multilingual courses and QR-authenticated certificates. Also built an offline retrieval assistant over its training content (35 documents, 2,342 cited fact chunks) that answers in the browser and declines when nothing matches.
- **Designed and built Lectern**, a white-label course authoring tool similar to Articulate Storyline: TypeScript Turbo monorepo, React + Vite editor, a course player engine, SCORM 1.2/2004, xAPI, and PDF/HTML export, plus a Tauri 2 desktop app. **2,000+ tests** across packages.
- Used AI coding agents (Claude Code, multi-agent workflows) for most of the build, with my own code review, tests, and CI checks keeping the output production quality.
- Ran my own security audits on it: **110 audit reports, 663 findings, 661 fixed**, each with a severity, exact code location, repro steps, and a patch backed by CI checks.
- Built **50+ client websites** and led the **DNN → Hugo migration**: brought in real source control, made builds much faster, and set up a shared design system that raised the quality of client sites.

Independent Security Researcher

Sacramento, CA · Jan 2026 – Present

- Vulnerability research and responsible disclosure via HackerOne (**mrblackkeys**): source-level code review with reproducible, unit-tested proofs of concept.
- **First accepted finding in Twilio's bug bounty program**, triaged and rewarded.

Earlier — Freelance & Frontend Roles

Greater Sacramento / SF Bay · 2016 – 2019

- Built full-stack web apps for startups and small businesses, working directly with owners from scoping to launch: React/Angular frontends, Node.js backends, Hugo static sites, and deployment automation. Earlier frontend roles at InfoStar and a Bay Area startup.

EDUCATION

Dev Bootcamp, Computer Programming (2016)

Academy of Art University, B.F.A. in Rigging & Animation (2012)