

José Miguel Hernandez

Staff Software Engineer | Applied AI & Agent Platforms

info@jose.io · 831-818-5505 · github: jose-m-hernandez · linkedin: josemiguelhernandez · Morro Bay, CA (remote)

SUMMARY

I make AI agents safe to run and worth using across an engineering org: platform, guardrails and adoption. At AdRoll, I classified our internal LLM agent as critical-tier and built its kill switch, its CI gate on unclassified write endpoints, and PR provenance auditable against our SOC 2 controls. 30+ colleagues, from the CTO to customer support, used it in its first month to resolve support and triage issues faster. I wrote the rules our shared agent skills repo is built to, planned migrations across 28 repositories with agents, and gave 800+ repositories a graph of what depends on what, who owns it and who to ask.

EXPERIENCE

AdRoll (NextRoll), San Francisco · Remote

Jun 2017 – Jul 2026

Staff Software Engineer

Hired as Senior Software Engineer, promoted to Staff. AdRoll has been used by over 140,000 brands and handled 120 billion ad auction requests a day. Advertisers' budgets and creatives become campaigns through our web apps and services. I kept those running, so our bidders kept spending at scale for our customers. Took PagerDuty on-call shifts and supported the systems I owned at night and on weekends.

Agent Safety and Internal Agents

- Owned the safety and governance layer of our internal LLM agent, built on the Claude Agent SDK, and became its second-largest contributor. Turned our AI agent security policy into a shared skill that classifies any agent by risk tier, and wrote classifications for four systems, including that agent, rated critical-tier under the lethal trifecta model (private data, untrusted input, ability to act). Built its mandatory kill switch in its Python broker, disabling writes while read-only investigation kept working, and a CI test that fails on unclassified write endpoints, which caught four new endpoints as the broker grew. Added requester-and-rationale provenance on agent PRs, auditable against our SOC 2 controls.
- Dozens of colleagues used the agent, from the CTO, SVP of Engineering and VP of Product to design, data, support, business systems and security. In its first month, 47 of its pull requests from Slack merged across five repositories. Closed two security-review findings and required signed Slack requests.
- Built and shipped 411Bot, a second Slack agent on the same runtime that answers who owns what from the service catalog, ownership records and org chart, promoting my prototyped skills into the shared repo.
- Largest contributor to our shared agent-skills repo. Wrote its eval-before-merge rule (each new skill's PR shows an invocation, input, output and pass/fail judgment) and its linter-enforced S/M/L sizing convention, which 125 of 128 skills carry. Audited the repo against the SVP of Engineering's nine AI principles, wrote its auto-invoke policy, and disabled auto-invoke on 54 skills.
- Evaluated skills against our real codebase before promoting them, re-running after each fix until clean.
- Audited skill-use telemetry for skills that never fired or ran on too small a model, and tested our internal agent by hand before and after every change to it or the shared skills.
- Took the Gong, Jira, Zendesk and Cube insight dashboard our Head of Product and Business Operations prototyped to ready for handoff in a week: moved its credentials to 1Password, added a server test layer, prepared its CI, purged customer data from its docs, and brought it into our product-AI repo. Turned its sources into shared agent skills, and support ran its Zendesk analysis from Slack.
- Defined the SOUL.md and wrote most of the guardrails and rapid-prototyping skills for Monet, a Slack design agent that built and deployed UX prototypes, bound to the design system's tokens, patterns and precedents.

Agent Platform and Developer Tooling

- Grew a dependency crawler, pass by pass, into a knowledge graph of the company's ~800 repositories, their dependencies, owners, Jira boards, Slack channels and org chart, and kept it from drifting: it generated upgrade PRs and Jira tickets, derived ownership from CODEOWNERS once manual mappings drifted, and held a schema-backed org chart in a source-of-truth repo shared with security, with every mapping reconcilable against other systems. Served it as an MCP server with seven tools, then as a skill in our internal agent.

- Built and ran the React 19 program while other engineers landed the individual upgrade PRs. Agents sequenced each upgrade along its dependency chain from the derived dependency and ownership data, and sized it against the owning team's past execution and known blockers. CI pipelines delivered it to that team as a PR ready for review and deploy, and no app ever got more than one major ahead. Designed its tooling, agent skills and scripts.
- Built the design system's canonical source and implementation gates so AI-assisted prototypes stayed aligned with downstream teams' components. A full-product Playwright snapshot gallery checked visual drift.
- Made design-system prototypes reviewable per pull request: a branch became a live preview URL inside its PR. Wrote the Terraform for keyless GitHub Actions deploys to AWS (OIDC federation, IAM trust policy, ECR), onboarded the repo to our Kubernetes (EKS) preview platform, and fixed the platform's cleanup job so PR activity kept an environment alive.
- Audited all 815 AdRoll production feature flags with a custom API client and corrected 300 mislabeled statuses. Gave each owning team a view of its stale flags and presented the results at AdRoll's Global Town Hall at the CEO's request. Cut our feature-flag vendor's seat count nearly in half from one Slack conversation with our internal agent, checking seats against team and ownership data I centralized.

Applied AI and Ad Products

- Built the bridge that let AdRoll move ad creation to the new ads service without a cutover: bootstrapped our shared library for managing ads, and created the service's embed view, the iframe message bus that let the legacy app host the new UI, and the old and new ad data models with the conversion layer between them.
- Built the on-brand AI copy and image-outpainting interfaces for Create Web Ads, our HTML5 ad builder. By fall 2025, about 250 advertisers were spending on ads built with it.
- Prototyped an AI ad builder that generated a templated HTML5 ad from a customer's site data in one shot.
- As an early partner for OpenAI's ads API before it was public, in two days mapped OpenAI's beta Ads Manager onto our ads and assets models from screenshots, wrote backend, frontend and data-model specs checked against our codebase, and built a working prototype on a mocked API. It proved the integration feasible and unblocked planning.
- Took a colleague's Media Planner prototype to production as an internal AI media-planning assistant for sales reps preparing pitches: built its front end and tool integration and wired in its planning heuristics. Wrote and ran its test protocol, checking every prompt edit against three sample briefs and a pass/fail checklist, and wrote its rule that the agent never invents a number when benchmark data is missing. Checking the prompt against a live API response turned up eight schema mismatches and drove the next iteration.
- In 2026, in the ads service's Python backend (Flask, SQLAlchemy), wrote the data migration correcting about 14,000 misclassified ads, and triaged all 100 unresolved Sentry issues into six tickets in a day. Diagnosed an N+1 query affecting 309 users and fixed three other bugs, including a silent failure that stopped syncs for whole advertisers and a label-creation race behind 11,000+ error events.
- Enabled WebP creatives across our ad platform and its pipelines: added WebP asset handling to the ad sync pipeline, removed its WebP-to-PNG fallback, and fixed the transparency loss it exposed in our Go image service, a hardcoded background in the govips C bridge.

Technical Leadership

- Ran Front End Core from 2018 through 2026 and spearheaded web performance and observability for AdRoll's customer-facing apps: set the Core Web Vitals SLA, drove Datadog real-user monitoring across the product teams' apps, and ran the remediations that reduced page load times across many of them. Wrote the first Enzyme-to-RTL test conversions.
- Shaped the 2026 company roadmap: mapped the Creatives team's AI work to company-level bets and translated product goals into technical specs and effort estimates.
- Served on NextRoll's Eng Culture committee: produced our twice-yearly Hack Week 14 times and moderated its brainstorming. Placed in seven Hack Weeks, winning four. Mentored five interns on the Ads and Front End teams.

Pandora, San Francisco

Dec 2015 – Mar 2017

Senior Software Engineer, Consumer Electronics

- Enabled Pandora's 80 million active listeners to stream on smart TVs and inside Uber: built account linking in Uber's app, supported the LG smart-TV app, and architected and built the lab rig that verified cross-device communication and handoff.

- Shipped the Chromecast receiver UI for the Pandora Premium launch, adding repeat, shuffle, and queue controls, and profiled its performance on first-generation hardware.
- Worked on integrations for Samsung, Android TV, and Windows 10, plus shared client and message-bus infrastructure for Chromecast and Google Home.

Rdio, San Francisco

Jun 2014 – Dec 2015

Senior Software Engineer

- Managed Rdio's Roku partnership and built account-linking and subscription-billing flows.
- Built Rdio's shutdown export, letting listeners download their collections or transfer stations to Pandora, with a personal listening-history timeline.

TastemakerX, San Francisco

Mar 2013 – Jun 2014

Senior Software Engineer

- Relaunched TastemakerX's music-discovery client as a Backbone single-page app, building activity-feed aggregation, a media-embedding engine, responsive UI, and Facebook onboarding. Co-built and shipped convertmyplaylist.com, a playlist converter from Spotify to Rdio.

Telly (formerly TwitVid), San Francisco

Dec 2011 – Sep 2012

Lead Front-End Engineer

- Led TwitVid's front-end redesign into Telly, rebuilding its PHP app as a hybrid SPA with History.js navigation.

50 Cubes, San Francisco

Jan 2010 – Oct 2011

Director of Activation

Owned platform integrations and distribution for two top-100 Facebook games, Mall World (5.6M MAU) and Fashion Designer (3.4M MAU).

- Owned the JavaScript/HTML platform and payment layer around both games: Facebook integration and viral flows, syndication across MySpace, Orkut, hi5, Plinga, and Viximo.
- Built a SWF-to-SVG/PNG conversion pipeline for ~14,000 Flash assets, then built JANICE, a mobile-browser avatar runtime that consumed the converted catalogue.

AOL, San Francisco

May 2008 – Nov 2009

Senior Software Engineer

- Built admin tools, templates, and reporting for self-counting ad placements deployed across ~50 locations, including most of Bebo's front page at ~9M daily views.

Bebo, San Francisco

Sep 2006 – May 2008

Senior Software Engineer, Ad Operations

Hired into ad operations to build sponsored profile pages, then moved to engineering and promoted.

- Hand-built large-brand sponsorships, profile takeovers, and contests in HTML and Flash, including full IAB size matrices with Flash units and GIF fallbacks. Homepage takeovers were worth \$10K–\$50K a day.
- Ad targeting on every page load was slowing a site that grew past 200 million page views a day. Coming from ad operations, I engineered the front end of the fix and led its product side: wrote the spec for a counter service that split targeting and analytics out of the page, designed the front end's data schema, designed and built its dashboards, and co-built the service while keeping the existing ad setup running until launch.

SKILLS

Applied AI · AI agents, large language models (LLMs), agentic workflows, Claude Code, Codex, Cursor, Claude API and Agent SDK, Gemini, local LLMs (Ollama), MCP, tool calling, prompt engineering, evals, guardrails, provenance

Developer Tooling · CI/CD and merge gates (Buildkite, GitHub Actions), dependency and ownership graphs, migration automation, codemods, MCP servers, agent skills, feature-flag hygiene (Split), Node.js, shell scripting

Languages · Python, TypeScript, JavaScript, Go, Java (JSP), PHP, SQL, GraphQL, Swift, Objective-C

Backend and Data · Flask, Django, SQLAlchemy, Postgres, Trino, Redash, OAuth, payments, billing

Infrastructure · AWS (S3, ECS, ECR, IAM), Terraform, Docker, observability (Datadog, Sentry)

Testing · Playwright, Jest, React Testing Library, Appium (device labs)

Frontend and Design Systems · React, Next.js, Redux, Webpack, design systems, component libraries, Figma, design tokens, accessibility

EDUCATION

University of California, Santa Cruz · BA, Film and Digital Media