Michael Nakayama

Staff engineer · applied AI, agents & full-stack product

Nashville, TN · mnaks.me · github.com/sportnak · linkedin.com/in/michael-nakayama-20827924

EXPERIENCE

MagicSchool AI · Staff Engineer

Jan 2025 – Present · Remote

Raina AI teaching assistant (2026)

- Led the re-architecture of Raina, an AI teaching assistant handling **~1.1M conversations/week**, from a single-turn LLM form flow into a tool-calling agent: tool search and execution over **100+ tools**, knowledge-base retrieval, sub-agents with streamed output, and a code-orchestrated createArtifact tool that took tool selection away from the model to improve output quality.
- Contributed to eval-driven development in Braintrust (scorers, LLM judges, 150+ case experiment arms, regression-gated merges) and set up decision-model judges that run **10–50× cheaper and faster** than LLM judges.

Studio Mode (2025–2026)

- Built and shipped Studio Mode, a collaborative document workspace inside Raina producing **100,000+ AI-generated documents per school day** (about **one in four** assistant turns): Markdown streamed as Yjs updates over websockets into a block-based editor, with document context engineered into the conversation so the agent edits what the teacher sees.

Search chat threads (2026)

- Led full-text retrieval over teachers' chat history from spec to production: Postgres tsvector search with GIN indexes over **18M threads** plus **~41K new threads per day**, hash-partitioned by user with row-level security.

Data Analysis Tool & platform (2025)

- Built the **Data Analysis Tool from zero to launch** as one of two core engineers: a conversational AI product where school administrators upload spreadsheets and get insights and charts; owned the NestJS file and query API, LLM query generation, and AWS deployment (ECS, Lambda, GitHub Actions, Secrets Manager).

Pursuit · Founding Full Stack Engineer

Feb 2024 – Nov 2024 · Remote

- Conceived and built a sales-forecasting product that won **three customers in its first month (>$100K revenue)**: a Temporal-orchestrated ML pipeline (scikit-learn) that confidence-scored prospects, reaching 70–80% confidence on 3K test records.

Olo · Senior II Engineer

Mar 2022 – Feb 2024 · Remote

- Achieved a **5x throughput increase** on an event platform handling **1M+ daily events** through message deduplication and optimized distribution.

Faithlife · Senior Engineer · Team Lead

Aug 2015 – Mar 2022 · Bellingham, WA

- Managed **4–8 direct reports**, leading growth conversations, technical design, and code review.

PROJECTS

Fantasy football decision engine · Python, LightGBM, PyTorch, scikit-learn, FastAPI, Next.js, Modal

2026

- Built a layered Monte Carlo NFL season simulator on **4.2M rows** of 2012–2026 NFL data that has generated **600M+ rows** of season trajectories, graded against Sleeper projections with calibration, CRPS, and coverage metrics by player tier.
- Designed a draft engine that treats each pick as a portfolio decision, maximizing the odds of a top-2 finish rather than projected points, using player trajectory models, an ADP-based model of opponents' picks, and value-over-next-available ranking, evaluated on held-out real drafts.
- Built it agent-first with Claude Code as a context-engineering exercise: layered context files (global, project, and per-area design docs), **9 custom skills** for recurring workflows, and specialist agents as review gates.

SKILLS

Languages: Python, TypeScript, SQL, JavaScript, Go

AI engineering: Agent loops & tool calling, Context engineering, Prompt engineering, Evals & LLM-as-judge, Structured output, MCP, Braintrust, Bedrock

Frameworks: React, Next.js, Node.js, NestJS, FastAPI, Playwright

Data & infra: Postgres/Supabase (full-text search), AWS (ECS, Lambda), Modal, Trigger.dev, Temporal, Redis, Datadog

Writing: mnaks.me: technical posts on model evaluation, error discovery, and building AI products

EDUCATION

B.S. Computer Science, Vanderbilt University · 2015