

Nathan Gamache

Lead Developer: LLM agents, MCP, retrieval, fine-tuning, evaluation

Michigan, USA nathanpaulgamache@gmail.com
nathangamache.com github.com/nathangamache
linkedin.com/in/nathan-p-gamache

SUMMARY

Lead developer who builds production LLM agent systems (MCP, retrieval, human-in-the-loop approvals, evaluation), has fine-tuned an SLM with QLoRA and trained reward models for reasoning. Invented PANoodle, the AI operations platform a digital agency runs on, and co-founded CatalogXpert, an AI sales assistant venture. B.S. Computer Science (AI concentration), expected Dec 2026.

SKILLS

AI and ML: LLM agents and tool use, MCP, RAG and vector search (Chroma, LLM reranking), prompt and context engineering, LLM evaluation, reward modeling (process and outcome), human-in-the-loop design, QLoRA/LoRA fine-tuning (PEFT, TRL), GGUF quantization, PyTorch, Hugging Face Transformers, scikit-learn, ONNX Runtime, computer vision (SigLIP 2, V-JEPA 2, pose estimation), Whisper, llama.cpp, Ollama, Anthropic Claude and OpenAI APIs

Software and data: Python, TypeScript, PHP, SQL, Java, FastAPI, PostgreSQL (Row-Level Security, PostGIS), Redis, GraphQL, WebSockets, React, Docker, GitHub Actions, Linux, Nginx, Cloudflare, WordPress, WooCommerce, Stripe, HubSpot, IBM i (DB2)

EXPERIENCE

Lead Developer · Perfect Afternoon, a digital agency supporting 260+ active websites for 100+ clients Mar 2026 – Present

- Invented and built **PANoodle**, the agency's AI operations platform (94% of its 779 commits). The team's Claude and ChatGPT connect to it over MCP for 1,300+ tools across 200+ connections, with a memory for every client and site.
- Its ClickUp agent runs maintenance and audit tickets end to end, with a person approving every live-site change: 181 tickets in its first 90 days. PANoodle helped build itself, shipping 52 of the 54 improvements its own agents requested.

WordPress Developer · Perfect Afternoon Jun 2022 – Mar 2026

- Handled 3,200+ tasks for 60+ client accounts and internal projects since June 2022; wrote the scopes others build from.
- Led **Corrigan Companies'** 11-site multisite (launched 4 of 5 live sites) and built both sides of **Terminal Supply Co's** WooCommerce to IBM i (DB2) REST integration; took MJS Packaging's mobile PageSpeed from 47 to 90+.

Co-founder · CatalogXpert, a venture of JDW Ventures and Gamache Technology Solutions Mar 2026 – Present

- Wrote the original demo and first sales engine of its AI sales assistant for technical-product stores: bundle planning, Chroma vector search, and LLM reranking with rule-based guardrails. One pilot running; demo at demo.catalogxpert.com.
- Wrote the build and migration specs and reviewed the work before merging: multi-tenant Row-Level Security, an LLM gateway with provider fallback and per-call cost tracking, and a 20-persona end-to-end eval raised from 13/20 to 20/20.

Founder · Gamache Technology Solutions LLC Feb 2024 – Present

- For main client **SJD Taxi**, turned an unfinished booking system into one REST hub for 4 sites (work split with Perfect Afternoon) with 5-currency Stripe checkout, and hardened a colleague's HubSpot sync, which has not missed a deal since.

Moderator, then Administrator · The Outpost, a Minecraft server with about 10,000 total players Sep 2019 – 2022

SELECTED AI/ML PROJECTS

Streaming Violence Anticipation · solo senior capstone, AI-assisted Fall 2025 – May 2026; rebuilt Oct 2026

- Traced the capstone-era models' high scores to leakage and rebuilt the evaluation (7,604 videos, grouped folds, causal scoring). Of pose, motion, V-JEPA 2 and SigLIP 2, only SigLIP 2 prompt scores passed fixed gates; their probe warned 30+ s ahead for 31% of 344 events at ~2 false alarms/hour vs 10% for a no-video baseline (screening folds, not yet confirmed).

Detroit Pulse · solo, Deep Learning course Spring 2026

- A local speech-to-LLM pipeline on one RTX 4070 Ti Super that turns scanner radio into a live incident map (Silero VAD, Whisper large-v3, grounded LLM extraction, an LLM judge). Fine-tuned Qwen2.5-7B-Instruct for address extraction with QLoRA (4-bit NF4, PEFT, TRL), evaluated against the base model on exact match, F1 and hallucination rate.

Process and Outcome Reward Models · co-authored Machine Learning course paper Fall 2025

- ChatGPT labelled a 3B SLM's chain-of-thought steps and answers; fine-tuned two 1.5B reward models on them (step-level 98% validation accuracy) to pick the best of several sampled solutions.

EDUCATION

Lawrence Technological University · B.S. Computer Science, Concentration in Artificial Intelligence Expected Dec 2026