Declan Javers

linkedin.com/in/declan-javers | github.com/declanjavers | declanjavers.com | declanjavers@gmail.com | 301-642-3823

EDUCATION

Northeastern University, Khoury College of Computer Sciences **Boston, MA**

Bachelor of Science in Computer Science, Concentration in Artificial Intelligence *December 2026*

Relevant Coursework: Foundations of AI (Graduate Level), Fundamentals of Software Engineering, Algorithms and Data Structures, Object-Oriented Design

EXPERIENCE

Zeta Global **June 2026 - August 2026**

Core AI Engineer Intern *New York, NY*

Worked on the improvement loop for a deployed production AI agent: finding real failures, fixing agent behavior, and turning checks into repeatable LLM evaluations.

- Classified production LangSmith traces into three recurring failure types: silent failures where a broken backend tool still reported success, bad tool parameters, and tools called out of order. Built a labeled failure dataset that set fix priority and became the baseline for measuring improvement.
- Rewrote tool descriptions and schemas and wrote agent instructions and skills, spelling out each tool's purpose, required inputs, expected outputs, and when not to use it, to target the wrong-tool and wrong-order calls found in the traces.
- Built smoke, regression, and progression LLM evaluation suites with custom Python evaluators and promptfoo that check backend state instead of trusting a successful response. Promoted consistently passing progression tests into regression so new agent abilities could not silently break.
- Split one large Model Context Protocol (MCP) server into two focused FastMCP servers on AWS Bedrock AgentCore, with a bridge that imports shared tool code from an internal package so the logic lives in one place.
- Fixed the platform SDK's local evaluation pipeline, which was producing misformatted MCP client spans that left trajectory evaluators grading empty traces. Evaluators now grade real agent runs.

Capitol AI **May 2025 - May 2026**

Software Engineer Intern *Washington, DC*

Worked across the stack on an enterprise AI platform built on retrieval-augmented generation (RAG) over proprietary data and agentic workflows. Owned the document-generation agent for a Big 4 consulting client's isolated EU tenant, built on Claude Opus through the Anthropic API.

- Fixed problems the client flagged during UAT, including inconsistent output between runs, missing chart axes, and overlapping text, by tightening prompt and skill files, pinning the model to Claude Opus 4.6, and adding citation rules tested across seven output formats.
- Rewrote the descriptions of 13 agent tools and separated white-label themes by tenant so the client's agent used its own labels and branding.
- Shipped React and TypeScript features and FastAPI REST endpoints across the consumer product and the visual composer where users build AI agents, and built a Playwright end-to-end testing harness covering eight services.

Northeastern University, Khoury College **June 2024 - December 2024**

Web Content Coordinator Co-op *Boston, MA*

- Built and managed WordPress pages for the launch of a redesigned college website. Automated content updates with Python scripts, and ran UX research, A/B testing, and accessibility audits that found and fixed usability problems.

Delta Kappa Epsilon International Fraternity **May 2023 - May 2025**

Executive Board: Vice President, VP of Health and Safety, Secretary *Boston, MA*

PROJECTS

Pulse Pitch, pulse.declanjavers.com **2026**

Personal Project *TypeScript, Python*

- Built and deployed a browser arena sports game in TypeScript and React with online 2v2 to 4v4 play on an authoritative WebSocket server: 60 Hz simulation, 30 Hz snapshots, client prediction and reconciliation, and mid-match rejoin.
- Shared a single simulation module between browser and server so physics stays identical, and built passwordless email sign-in and cloud saves on Cloudflare Workers, D1, and Drizzle that store only hashed codes and session tokens.
- Built a reinforcement learning pipeline on a NumPy copy of the simulation, parity-tested against the TypeScript version to 1e-6 per tick, training policies with behavior cloning on 2 million samples and PPO self-play for a custom TypeScript inference runtime.

Tidesworn, declanjavers.itch.io/tidesworn **May 2026 - June 2026**

Personal Project *Unity*

- Built an open-source generative AI pipeline that turns text prompts into game-ready 3D assets for a Catan-inspired multiplayer Unity game, chaining a text-to-image model (Z-Image) into image-to-3D generation (Trellis) on Google Colab.

TECHNICAL SKILLS

AI and LLMs: LLM evaluation (evals), AI agents, agentic workflows, RAG, tool calling, prompt and context engineering, observability and tracing
AI Tools: MCP, FastMCP, LangChain, LangSmith, promptfoo, Anthropic API, Claude Code, Codex, Hugging Face
Languages and ML: Python, TypeScript, Java, SQL, C, PyTorch, reinforcement learning (PPO, behavior cloning), NumPy, scikit-learn
Engineering: FastAPI, REST APIs, React, Next.js, WebSockets, PostgreSQL, Cloudflare Workers, Docker, Sentry, Playwright, pytest
Interests: Fitness, Soccer, Reading, Cooking