

# David Dickinson

davidrd@alum.mit.edu · github.com/davidrd123 · linkedin.com/in/david-r-dickinson · Culver City, CA

ML engineer building generative video/image systems — diffusion inference optimization (Triton, CUDA, Nsight), LoRA training pipelines, and production tooling for creative teams.

## EXPERIENCE

---

### ML Engineer

Aug 2025 – Present

*Rude Elf Collective (Freelance)*

Culver City, CA

- Built a generative media platform for an Academy Award-nominated production designer — multi-provider image generation studio (31 routes, atomic state, provenance tracking), YAML-driven batch orchestration, and ComfyUI automation with graph patching (Python, FastHTML, HTMX)
- Built end-to-end LoRA training data pipeline: FFmpeg scene detection, frame-accurate segmentation, automated captioning (Gemini, Qwen 2.5 VL), NDJSON clip indexing with sidecar-based provenance
- Maintained 75+ tests and 49 ADRs across an 8-step modular refactor of the full platform

### Participant — Interactive AI Video Program

Dec 2025 – Feb 2026

*Daydream (Competitive cohort · \$12K prize pool)*

Remote

- Achieved 3.9× throughput (8.8→34.5 FPS) on Blackwell GPUs via profiling-first methodology — built CUDA-event profiling infrastructure, identified VAE decode as bottleneck, resolved runtime and tensor layout issues forcing slow Conv3d dispatch
- Integrated FA4 score\_mod attention, fused RoPE, and torch.compile with architecture-specific guards; systematic one-change experiments across three GPU architectures (SM90/SM100/SM120)
- Built pipeline-parallel + tensor-parallel inference architecture (PP1+TP) with server integration, NCCL control plane, and 37 validation runs across H200, B300, A100, and RTX PRO 6000
- Awarded 2nd place (Scope track, Jan cohort) and Honorable Mention for Creative Excellence (Feb cohort); credited for VACE implementation contributions to Krea Real-Time 14B

### ML Engineer

Mid-2024 – Aug 2025

*Fair Trade Models*

Culver City, CA

- Trained text-to-image and text-to-video generative models for creative production, iterating on LoRA fine-tuning across multiple architectures (Flux, Wan 2.2)
- Delivered an Allstate spec commercial using AI-generated visual content, from concept through final output
- Built training data pipelines (video scene detection, frame-accurate segmentation, multi-model captioning) and custom tooling for the full LoRA training lifecycle

## PROJECTS

---

### Umbra — umbracollab.net

2023

*Real-time collaborative browser-based code editor*

- Engineered a service-oriented backend (Node.js, Express, AWS EC2/RDS/Cognito, Cloudflare Workers/Durable Objects/R2) with real-time collaborative editing via Yjs CRDTs over WebSocket
- Automated deployment with GitHub Actions and Docker; authored technical case study; team of four across three time zones

## TECHNICAL SKILLS

---

**Languages:** Python, TypeScript, JavaScript, Go, SQL, HTML/CSS

**ML / AI:** PyTorch, Triton, CUDA, Nsight (ncu/nsys), LoRA, diffusers, Gemini API, OpenAI API, ComfyUI

**Web:** FastHTML, HTMX, FastAPI, Node.js, Express, React, WebSocket, WebRTC

**Infra / Tools:** AWS (EC2, RDS, Cognito), Cloudflare (Workers, Durable Objects, R2), Docker, GitHub Actions, PostgreSQL, Redis, FFmpeg, Git, pytest, Linux

## EDUCATION

---

### Massachusetts Institute of Technology

1998 – 2003

B.S. Electrical Engineering and Computer Science (6-3). Media Lab (HCI), tangible interfaces.

### Launch School

2021 – 2023

Full-stack software engineering. Mastery-based (exams + live coding interviews). Capstone: Umbra. | AWS Solutions Architect cert.

Previously: Private STEM Tutor (self-employed, 2009–2019) — math, physics, and introductory programming. Los Angeles, CA.