

Shawn Liu

Software Engineer · LLM applications · M.S. CS, Columbia
(510) 565-2523 · xl3704@columbia.edu · LinkedIn · github.com/ShawnnnLiu · shawnnnliu.github.io
New York, NY · U.S. Permanent Resident, no sponsorship required

EDUCATION

Columbia University · M.S. in Computer Science · New York, NY Expected December 2027
University of California, Irvine · B.S. in Computer Science, GPA 3.84 / 4.0 · Dean's Honor List · Irvine, CA September 2022 - June 2026

PROJECTS

Loop · LLM study-planning agent, live at loop-study.com · Python, FastAPI, SQLite, Anthropic API, Docker, Fly.io May 2026 - present

- Shipped an LLM agent turning a career goal and free hours into a Google Calendar study plan: five LLM nodes propose, deterministic validation and human approval decide, every write verified and rolled back. ~30 monthly active users, ~100 plans generated.
- Grounded plans in RAG over a curated 242-document corpus: BM25 retrieval scored recall@5 0.76 and MRR 0.71 on 75 labeled queries. A dense-embedding hybrid scored +0.08 MRR and was not shipped.
- Gated the build on offline evals in GitHub Actions: CI re-grades committed LLM recordings and makes no model calls, failing below floors for schema validity, repair rate, and citation coverage; an LLM-judge score is reported, never blocking.

Foothold · transfer-credit navigator, live at foothold-transfer.com · Featured in Stellic Pathfinders 2026, top 20% of 169 August 2026

- Launched, solo in three weeks, a tool telling community-college students which courses transfer to a UC or CSU: a rules evaluator over 352,024 ASSIST rows with nine typed finding codes, every result citing the agreement it came from.
- Drafted appeal letters with Claude behind a validator: every course code in a draft is checked against the findings, a template fallback lands after two failed repairs, and repeat requests reuse a finished draft instead of calling the model again.
- Deployed it on Fly.io as a two-stage Docker image behind a 30-second health check, after an ingestion run of 34,975 requests at one per second with zero rate-limit errors; pinned to one machine once a second machine's cloned volume split sessions.

RESEARCH EXPERIENCE

Columbia University · Graduate Research Assistant · New York, NY September 2026 - present

- Verified C/C++ security patches as ground truth for evaluating coding agents (pinned environments, AddressSanitizer replay, per-test regression diffs, line-ablation minimization); wrote the Python and shell tooling that runs those checks.

BioIntelligence Lab, UC Irvine · Undergraduate ML Researcher · Irvine, CA December 2025 - July 2026

- Built a graph convolutional network over ESMFold-predicted conformations to predict antimicrobial peptide MIC; ablations against QSAR descriptors alone isolated what 3D conformation adds: 22% to F1 at the chosen threshold.

BiasLab, UC Irvine · Undergraduate ML Researcher · Irvine, CA June 2024 - June 2026

- Trained a PyTorch pose-estimation and symbolic-reasoning model for U.S. Navy UAV carrier landing and defined its metrics (flight time, deck-crew collision, impact speed): 65% fewer failures than legacy optical markers in adverse-weather simulation.
- Co-authored 4 accepted papers: WACV 2026, ISLPED 2026, NeurIPS 2025 NeurReps Workshop, Frontiers in AI 2026.

EXPERIENCE

AdamsFoods · Full-Stack Developer · Los Angeles, CA Summers 2024, 2025

- Developed a wholesale e-commerce platform (React, Node / Express, AWS S3, JWT admin routes) and a 300+ SKU inventory system on PostgreSQL that replaced handwritten logs and saves staff 5-7 hours a week.
- Collaborated in a team of 4 through agile sprints, gathering requirements directly from warehouse staff.

SKILLS

Languages: Python, TypeScript / JavaScript, SQL
LLM applications: Anthropic API, RAG, structured outputs, output validation and repair, LLM-as-judge evals
Evaluation: offline eval harnesses, CI gates, retrieval metrics (recall@k, MRR), BM25
Backend and web: FastAPI, Node / Express, REST APIs, PostgreSQL, SQLite, React, AWS S3, JWT
Deployment: Docker, two-stage builds, Fly.io, GitHub Actions CI/CD, health checks
ML: PyTorch, graph neural networks, deep learning