Ethan Cai

EECS student at UC Berkeley. Some things I care about: verified AI, applied AI (mostly in healthcare), model training, reinforcement learning, trying new food, and chatting it up with cool people :)

ethan_cai@berkeley.edu · GitHub · LinkedIn · ethancai.com

RESEARCH

Metabolomics-Based ML Diagnostics of PASC

UC San Diego / San Diego Supercomputer Center · May 2024 – Dec 2025

First-author paper in Metabolites · 0.899 AUC-ROC

Machine learning pipeline for diagnosing Post-Acute Sequelae of SARS-CoV-2 Infection (Long COVID) using metabolomic data. A proof of concept that molecular descriptors calculated using E-Dragon can work to build assay-agnostic diagnostic algorithms with disease differentiating power. Differentiates PASC from 4 of the 5 PASC-similar diseases tested (ME/CFS, Lyme disease, POTS, and IBS), with only fibromyalgia proving indistinguishable. First-author paper published in Metabolites.

Tech: Python, Machine Learning, Bioinformatics

Project page · Code · Paper

Prioritized Level Replay in RL Driving Environments

UC Santa Cruz · May 2025 – Jun 2026

Demonstrates that prioritized level replay, an automated curriculum learning method for RL, improves PPO performance in highway driving scenarios with reactive traffic. Scenes are sampled from Scenic scenario specifications into MetaDrive, scored by positive value loss, and resampled by rank. In a controlled 30-seed comparison, PLR-trained agents beat the unmodified PPO baseline in 21 of 30 seeds (sign test, $p < 0.05$).

Tech: Python, Reinforcement Learning, PPO, Autonomous Driving

Project page · Code

Do Face Embeddings Predict Attractiveness?

Independent research · Jun 2026 – Aug 2026

Benchmarked 11 face-embedding families (deep identity CNNs, classical methods, CLIP, DINOv2) against 3 human attractiveness-rating datasets totaling ~7,800 rated faces, using frozen embeddings and ridge regression. Found that embeddings trained to capture identity did consistently worse at predicting attractiveness than those that were not, pointing to a tradeoff between pose invariance and attractiveness predictivity. CLIP led every dataset ($r = 0.94$ on SCUT-FBP5500), while Fisherface—the most aggressively identity-separating method tested—finished last at a mean $r = 0.32$.

Tech: Python, PyTorch, scikit-learn, Computer Vision

Project page · Code

TOOLS

Vigil

Talk to Users Hackathon, SF Tech Week · Oct 2026

Grand prize (Bright Data track) · F1 0.97 vs. 0.57 for a single-prompt LLM

Grand prize (Bright Data track), Talk to Users Hackathon (SF Tech Week 2026), built in 12 hours. Checks an investment adviser's website, blog, and third-party pages for marketing that may break the SEC Marketing Rule, quotes each flag for human review, and drafts rewrites for approval. Led the project and built its two-pass, cross-referenced inference pipeline and an eval suite drawn partly from 12 real SEC enforcement orders. On 125 eval cases: F1 0.97 with 4 false flags, vs. 0.57 and 111 for a single-prompt LLM.

Tech: Python, TypeScript, LLMs, Bright Data

Project page · Code · Live demo

Grant Test Auto

Cal Hacks 13.0 · Oct 2026

Built at Cal Hacks 13.0 for the Hyphenate track, after our PI talked at lab meeting about how much of his time goes to keeping funding instead of research. Since 2025, a grant can be canceled, restored by a court, and still not pay a cent, and the official records don't agree on which is true. On top of that, every purchase has to clear obscure funding rules, like no supplies bought in a grant's last 90 days, and someone has to check all of it by hand, constantly. Grant Test Auto automates that grunt work. It sits between a lab's researchers and its grant office: researchers ask for things in plain words, agents check each request and invoice against the grant's real balance and every spending rule, and one human grant manager approves each one with the reasons in front of them. "Awarded," "restored" and "paid" are three different facts, so a court order alone doesn't unfreeze a grant; only money actually arriving does. Claude reads and explains but never decides, since the rules live in the database where no model can talk its way past them. Worked on the backend, feature ideation and market research. Tested only on real cases: caught 38 of 38 freezes on 37 NIH grants and flagged 28 of 28 unallowable items from real audits.

Tech: TypeScript, SpacetimeDB, React, Claude, Apify

Project page · Code · Live demo

Anchor Lines

Jul 2026 – Aug 2026

Google AI-powered document intelligence app that transforms dense financial-aid documents and syllabi into clear, decision-ready summaries. Every insight stays anchored to the original source, with direct citations that make claims easy to trace and verify through a sleek, modern interface designed to support an expanding range of document types.

Tech: TypeScript, Next.js, Gemini

Project page · Code

Human Benchmark 3D

May 2025 – May 2026

1,800+ downloads on SideQuest

Virtual reality edition of classic block-tapping memory tests utilizing blocks of various heights. Uses Firebase to record user performance and provide comparative feedback to other users of the same age bracket. Downloaded over 1,800 times on SideQuest.

Tech: C#, Unity, VR, Firebase

Project page · Code · Live demo

Shambo

May 2026 – Jul 2026

Featured five-star extension on the Chrome Web Store. Blocks distracting sites and auto-closes tabs, with support for custom subpaths. Utilizes psychology (optionally) to fight procrastination by leaving snarky comments each time users try to open unproductive sites on their custom blocklist.

Tech: JavaScript, Chrome Extensions

Project page · Code · Live demo

Toto Simple

May 2026 – Aug 2026

Five-star Chrome Web Store extension designed around minimizing friction, with proactive reminders, one-click snoozing, urgent-task flags, and cloud sync. Originally built around AI task organization, but it made everything cluttered. Kept as a dear lesson that AI doesn't belong in everything.

Tech: TypeScript, React, Chrome Extensions

Project page · Code · Live demo

GAMES

DADS: Distraction-Aware Driving Simulator

2024

Winning submission to the Games for Change Game Jam in the 'Own the Road' category. A driving game that warns about the dangers of distracted driving by making users contend with simulated distractions while they try to drive safely on the road. Features summaries of risky driving caused by distracted driving.

Tech: C#, Unity

Project page · Code · Live demo

Party Box

Nov 2025 – Feb 2026

Most party game apps are ruined by forced ads and paywalled categories. Party Box bundles party games such as charades and imposter into one app without any content restrictions. It runs offline on Android, and in any browser on a phone or laptop with no signup or install: just a link.

Tech: Kotlin, Android, JavaScript

Project page · Code · Live demo · Google Play