

LEON LIU

TECHNICAL DEVELOPER ADVOCATE · AI AGENTS · EVALUATION · OPEN-SOURCE TOOLING

San Francisco Bay Area | English + Mandarin | leonwenhao@gmail.com | 239-908-8278
leonwenhao.com | [LinkedIn](#) | [GitHub](#) | [Writing](#)

PROFILE

Technical developer advocate and hands-on AI builder who turns real product use into open-source tools, tutorials, talks, and actionable product feedback. Achieved the highest OfficeQA accuracy in Sentient Arena Cohort 0 at \$1.85 per run; builds agent-evaluation infrastructure and works natively across U.S. and Chinese AI ecosystems.

EXPERIENCE

Founder & AI Builder | Dolores Research 2026 – Present
Agent evaluation, open-source tooling, developer education | San Francisco ↔ China

- Built Teller during [Sentient Arena Cohort 0](#) and achieved the highest OfficeQA accuracy: 71.5% (176/246) at \$1.85 per run with MiniMax M2.5. Sentient Arena is backed by Founders Fund, Pantera, and Franklin Templeton.
- Produced and presented an [EvoSkill setup tutorial](#) published by Sentient Labs, translating an agent self-improvement workflow into a reproducible developer walkthrough.
- Reported and isolated two Fireworks/OpenCode integration failures in EvoSkill [issue #48](#), including a verified complete MiniMax run; Fireworks support shipped as a first-class provider in [v1.3.0](#).
- Built Dolores Autocurricula, a local-first engine for generating, sandbox-verifying, solver-calibrating, scoring, and exporting software tasks; validated it through a 50-task, 250-attempt Fireworks calibration.
- Spoke at [Open AGI Builders Day](#) (Alibaba × Sentient × Fudan) and was featured by Sentient Labs at [MU Shanghai](#) discussing agent evaluation and China's open-source AI ecosystem.

SELECTED PUBLIC PROOF

Teller | *Open-source grounded reasoning agent* [GitHub](#)
• Published the competition methodology, reusable agent configuration, and technical report for enterprise document QA.

Yume | *1st place, Film & Simulation track – Worlds in Action* [Devpost](#)
• Built the backend and end-to-end generation pipeline for turning children's drawings into explorable 3D worlds in WebXR.

Public ecosystem work | *Speaking, interviews, and media* [Article](#)
• Lead builder profiled in 36Kr's [职场Bonus](#) (36Kr ≈ China's TechCrunch) coverage of the muShanghai residency; represented open-source agent work across English- and Mandarin-speaking audiences.

PRIOR EXPERIENCE

Product Lead & Business Development | Devermint Dec 2024 – Dec 2025
Enterprise Web3 studio | Remote

- Led product strategy and go-to-market for Aptos AI Layer and SUI Agents, coordinating engineering, design, and ecosystem stakeholders.
- Authored technical documentation, grant applications, and developer guides used in the Aptos Foundation accelerator and ecosystem showcases.
- Built partnerships and ran developer- and user-facing communications, including AMAs, technical content, and launch messaging.

Founder & Lead Product Builder | Anna Fantasy

Jan 2025 – Nov 2025

Generative AI video platform | San Francisco

- Designed and shipped an AI video platform with structured prompting, camera-scene controls, and community feedback loops; grew to 3,000+ monthly active users across 50+ countries.
- Turned complex generative-video workflows into product language and educational content for non-technical creators; generated 500K+ organic X impressions.

Head of Communications | DeVol Network

Apr 2022 – May 2023

Decentralized options protocol on Solana | New York

- Supported a \$3.5M seed raise through investor materials and outreach; built partnerships with Polygon, Circle, and Solana and translated DeFi concepts into technical and community content.

Web3 Innovation Consultant | Nike, Inc.

Jun – Oct 2021

Remote

- Advised senior Nike stakeholders on NFT strategy, blockchain-based digital collectibles, and the CryptoKicks initiative.

ADDITIONAL OPEN-SOURCE PROJECT

DeepRepo | *Recursive codebase analysis*

[GitHub](#)

- Built a CLI that coordinates a root model and focused worker models to analyze repositories and produce reusable technical context for coding agents.

SELECTED WRITING

[The Idea Guy Learned to Ship](#) | [Post-Training Is Just Dog Training](#) | [Chinese AI with American Characteristics](#)

CAPABILITIES

Developer Relations: technical tutorials, demos, documentation, talks, workshops, product-feedback loops, community and ecosystem partnerships

AI & Evaluation: agent harnesses, prompt and behavioral engineering, benchmark workflows, solver panels, verifier design, open-model inference

Tools: Python, JavaScript/TypeScript, React/Next.js, Git, Docker, CLI tooling, CI/CD, Fireworks, OpenRouter, Claude Code, Codex

EDUCATION

Emerson College – Visual & Media Arts / Film Production