

Jack Fan

317-774-4456 | jackfan@college.harvard.edu | jackfan.dev
linkedin.com/in/jack-fan-dev/ | github.com/itsjackfan

Technical Skills

Programming Languages: Python, Rust, Go, R, C, C++, Javascript (React, Next, Node), Typescript, HTML/CSS, SQL, NoSQL, Cypher

Frameworks & Tools: TensorFlow/Keras, PyTorch, LiteLLM/OpenAI/Gemini/Claude, FastMCP

Databases & Cloud: MongoDB, Supabase, Firebase, GCP, AWS, Pinecone, Turbopuffer, HelixDB

Development Tools: Git, Docker

Experience

Researcher

Kempner Institute at Harvard University

Dec 2025 - Present

Cambridge, MA

- Applied AI/ML research under Yilun Du and Embodied Minds Lab, working with Zhenting Qi and Karly Hou. Design and implementation for a novel, tree-based decentralised multi-agent architecture that improves Tau2-bench performance by 10% compared to baseline single- and multi-agent systems.

Forward-deployed/Research Engineer

ThirdLayer (YC W25)

Sep 2025 - Present

San Francisco, CA

- **Product and forward-deployed work** (Next.js/TypeScript backend): built custom Asana and Attio integration infrastructure (OAuth/API interfacing + custom tools/schemas, subagent prompts, evals/tests) + custom MCP subagent architecture (frontend interfacing, Vercel AI SDK-powered OAuth and authenticated MCP operations, prompting). **2x overall user count + DAU/WAU**

- **Applied AI research/evals work:** built agentic memory evaluation framework (Python) based on Meta's Agentic Research Environment (ARE). Task, data, and agentic harness design for tool-calling benchmark focussing on learning behaviours for effective extraction of and execution based on idiosyncratic user preferences.

Founding Lead Engineer

Series

Apr 2025 - Aug 2025

New York, NY

- Solo engineer for alpha version. Fully built and deployed Python backend (server logic, OpenAI/Gemini applied AI stack + prompts + tools, iMessage communication logic, MongoDB, deployed on GCP), and React/TypeScript frontend (deployed on Vercel). **From 3k to 14.8k total signups, 7k users on the platform in 2 months.**

- Lead engineer for beta version (v3) including build/deploy for Python/Go backend (FastAPI, MongoDB, OpenAI/PydanticAI agent layer, deployed on GCP) and Next.js 15.3.3 frontend (deployed on Vercel); **From 0 to 200,000+ matchable profiles and 0 to 10,000+ signups within 3 weeks** (left to remain at school for Fall '25 semester).

Senior Software Engineer

Harvard Tech for Social Good

Oct 2024 - Jan 2025

Cambridge, MA

- Lead engineer for 5-person team. Developed comprehensive impact statistics dashboard + weekly cron job update workflow used by 90+ Ersilia contributors, tracking 20+ key metrics, 180+ models across 17 countries and 200+ partner organizations. Full-stack app built on Plotly Dash, Python/GitHub Actions to aggregate and calculate data from Airtable and WHO APIs.

Software Engineer

Harvard Tech for Social Good

Sep 2024 - Dec 2024

Cambridge, MA

- Built LLM-powered webscraper on Gemini 1.5 Flash to automate AmpleHarvest's maintenance of food pantry directory; 650x reduced scanning time across 8000+ pantries in the directory; reduced hallucinations by 6x, pipeline runtime by 1.5x, and increased accuracy by 2x through prompt engineering and efficient caching.

Research Assistant

Columbia University Medical Centre

Sept 2022 - Apr 2023

New York, NY

- Computational neuroscience research; Python/R scRNAseq data analysis targeting potential effects of SETD1A gene on psychiatric disorder expression. Preprint publication on Biorxiv: <https://www.biorxiv.org/content/10.1101/2025.03.26.645419v1.full.pdf>.

Projects

Custom Autodiff/Backprop/MLP

GitHub

Dec 2025

- Toy vectorised autodiff engine implemented in Numpy, inspired by `autograd`; manual backprop and gradient descent computations supported. Optimised 2-layer MLP on MNIST to ≤ 0.4 loss.

Personal Knowledge Management System

GitHub

Mar 2024 – Nov 2024

- FastAPI-based backend with MongoDB integration; custom PageRank/GraphRAG pipeline for knowledge graph maintenance and organisation.

Publications

Qi, Z., **Fan, J.**, Hou, K., Han, S. S., Lakkaraju, H., Du, Y. (2026). *Scaling Agentic Intelligence with Principled Multi-Agent Decentralization*. Submitted to International Conference on Machine Learning (ICML), pre-review.

Sun, Z., Zhu, H., He, X., Lendemeijer, B., **Fan, J.**, et al. (2025). *Genomic and transcriptomic signatures of SETD1A disruption in human excitatory neuron development and psychiatric disease risk*. bioRxiv preprint, doi:10.1101/2025.03.26.645419.

Fan, J. (2023).

A *Quantum Leap Forward*. ISBN: 979-8387719516.

Education

Harvard University

Computer Science

Cambridge, MA

Sep 2024 – May 2028

Relevant Coursework: Compositional AI Systems, Intro to Systems, Data Structures and Algorithms, Software Engineering in the Age of Generative AI, Advanced Topics in Data Science (applied ML and deep learning), Intro to Probability Theory, Linear Algebra