

Charlotte Zhu

Stanford, CA | LinkedIn | GitHub

Education

Stanford University

Stanford, CA

• M.S. in Computer Science, GPA: 3.9/4.3

Sep 2025 – Jun 2027

• B.S. in Symbolic Systems, GPA: 4.1/4.3 (Sterling Award for Scholastic Achievement – top 1%)

Sep 2021 – Jun 2025

Professional Experience

AI Engineer Intern – Good Humans Inc., San Francisco, CA

Jun 2026 – Aug 2026

- Built a full-stack **annotation platform (Flask, Supabase)** with auth and server-enforced blind review to systematically capture human judgments of AI companion behavior.
- Owned **LLM generation style evaluation** end-to-end across 10+ LLM APIs; Curated a **340+ row eval set** from team judgments.
- Identified human annotation as the key bottleneck; independently designed and built an **LLM-based judge system** reaching **96% agreement** with human raters.
- Applied LLM judge for **prompt optimization**, achieving **12.5%** improvement in style alignment; shipped best prompt to production.
- Iterated memory encoding and retrieval system across **8K+ chat sessions** and **4K+ voice-call transcripts** existing data.
- Built a **human-coded 50-session test set**; improved memory encoding accuracy **21%** and retrieval accuracy **18%**.
- Shipped a **redesign of the onboarding flow** (Day 1–3), raising three-day conversion from 6% to **14%** for all app downloaders.

Research Assistant – Stanford HCI Group

Sep 2024 – Present

- Led a study evaluating LLMs' ability to predict human emotional experiences on social media; Ran a **200 participant** user study, ran LLM predictions in batch and statistical models across a **400k+** tweet corpus for analysis.
- Designed and built **LoveSims**, a persona-conditioned multi-agent LLM simulation for relationship "what-if" scenarios via structured dialogue, benchmarked via a **4,560-interaction** study; extended abstract published at ACM CHI EA '25 [Link].
- Developed **Vera**, a RAG-based bias-warning system for educational contexts; ran an **18** participant user study (workshop paper).
- Contributed to the Social Media Feed Elicitation project, implementing a parallel-processing framework for LLM-based tweet-preference coding that cut latency by **28%**; full paper accepted to ACM CHI '26 [Link].

AI Product Management Intern – ADVANCE.AI, Singapore

Jun 2024 – Aug 2024

- Built a **RAG-based support chatbot** to streamline customer support agents' work with response auto-generations.
- Built the pipeline for document ingestion, chunking, retrieval and generation with **Langflow**.
- Iterated and tested the prototype on **3k+** support tickets and **1k+** internal documents.
- Shipped a prototype that automated **30%** of common support issues, cutting manual lookup time for agents.

Projects

Cross-Cultural Bias in VLM Emotion Recognition

Apr 2026 – Jun 2026

Python, PyTorch, LoRA, LLaVA-1.5-7B, Qwen-2.5-VL-7B, GradCAM

- Co-designed a **2x2 LoRA finetuning ablation** (LLM-only vs. LLM+projector; balanced vs. unbalanced data) across **LLaVA-1.5-7B** and **Qwen-2.5-VL-7B** to study cross-cultural bias in facial emotion recognition.
- Found LoRA finetuning corrected LLaVA's near-total neutral-expression failure, raising recall from **0.001** to **0.868**.
- Showed balanced training improved East Asian accuracy by **+5.5pp** without sacrificing Caucasian accuracy, narrowing a key cross-cultural fairness gap.
- Built a **GradCAM** interpretability pipeline (facial landmark segmentation, region-level attention analysis) across **3k+** test images.

Team Lead & Full-Stack Developer – RendezViewz

Sep 2023 – Jun 2024

React Native, Javascript, Supabase, PostgreSQL, OpenAI API, OMDb API

- Developed a full-stack social movie platform enabling users to rank movies/shows and interact with friends.
- Built scheduling for group-watch sessions and AI-assisted movie recommendations based on ranking preferences.
- Implemented authentication, DM functions, and main feeds. Connected backend to **Supabase**.
- Integrated **OpenAI** and **OMDb APIs** for LLM-personalized recommendations and movie rankings.
- Ran **30+ user interviews / testing** to iterate on the UI/UX and improve usability. Contributed to Figma design.
- Grew registered users from **25 to 100+** in **2 weeks** via referral loops. Hit **75%** week retention.

Skills

Languages: Python, C++, JavaScript, TypeScript, SQL, HTML/CSS, R

Frameworks & Libraries: React, React Native, Node.js, Next.js, Flask, Django, Tensorflow, Pytorch, Langchain

Databases & Cloud: PostgreSQL, Supabase, Firebase/Firestore, AWS, GCP/Cloud Functions

Tools & Platforms: Git, Docker, Figma, Langflow

ML/AI: LLM APIs, RAG, Prompt Engineering, Embeddings & Vector Search, Automated Test & Evaluation Pipelines