# Raymond Tsao

Github: github.com/Raymond112514
Linkedin: linkedin.com/in/r112514
Personal Website: raymond112514.github.io
Email: raymond1123581321@gmail.com

## Education

**University of California, Berkeley** June 2022 – Present
*B.A. Computer Science Honors, B.A. Applied Math Honors, expected May 2026* *GPA: 3.961*
M.S. Electrical Engineering and Computer Sciences, expected May 2027

***Research interests**: policy pretraining for efficient robotic adaptation and sample-efficient online RL*

***Relevant coursework**: Natural Language Processing*, Combinatorial Algorithms and Data Structures*, Computer Vision, Introduction to Artificial Intelligence, Computer Architecture, Operating Systems and System Programming*
**Graduate coursework*

## Experience

**Robotic AI & Learning Lab (RAIL), UC Berkeley** Jan. 2025 – Present
*Researcher – Sergey Levine, Andrew Wagenmaker*

- Co-authored *Posterior Behavioral Cloning*, running experiments and analysis on policy pretraining for sample-efficient RL fine-tuning; accepted as an **ICML 2026 Spotlight (top 2.2% of submissions)**.
- First author of *Learning Process Rewards via Success Visitation Matching for Efficient RL*; developed a reward learning method that transforms sparse success signals into dense process rewards for more efficient robot policy fine-tuning (under review).
- Training **7B-parameter vision-language models** for multitask value/advantage estimation on large-scale robot datasets including DROID and RoboArena, using **distributed training on Google Cloud TPUs** for generalist policy improvement.

**Berkeley Function Calling Leaderboard (BFCL), Berkeley RISELab** Sep. 2024 – Jan. 2026
*Undergraduate Researcher – Joseph Gonzalez, Huanzhi Mao*

- Led development of the **BFCL-v4 Web Search Benchmark**, building its multihop web-search dataset, data generation and evaluation pipelines, and search API used to evaluate **100+ open-source and proprietary LLMs**.
- Developed automated generation and evaluation pipelines for challenging multihop web-search tasks, enabling scalable evaluation of LLM tool-use and search capabilities.
- Designed visualization tools and custom evaluation metrics for BFCL-v3 to characterize function-calling error modes and identify systematic failure patterns across LLMs.

**National Center for High-Performance Computing** Jul. 2024 – Aug. 2024
*Research Intern, Data Science Team*

- Developed a **ReAct-based LLM agent** for natural-language SQL querying and multi-step reasoning over agricultural GIS data, supporting **1K+ domain-specific queries** including flood-damage estimation.
- Curated and processed **500GB of agricultural GIS data** across 5 crop types, resolving irregular field boundaries and data-quality issues for downstream geospatial analysis.
- Fine-tuned **DeepLabv3 and SegFormer** for agricultural field segmentation, improving pixel accuracy from **43% to 82%**.

## Technical Skills

**Technical Skills:** Python, PyTorch, TensorFlow, JAX, Git, distributed training on Google TPUs, GPU cluster training.