

# Anikait Singh

✉ [anikait@stanford.edu](mailto:anikait@stanford.edu) | [asap7772.github.io](https://github.com/asap7772) | [asap7772](https://www.linkedin.com/in/asap7772) | [@Anikait\\_Singh\\_](https://twitter.com/Anikait_Singh_) |  Scholar

## Research Interests

---

My research aims to understand and tackle the core bottlenecks in scaling reinforcement learning to foundation models. I investigate this through the intersection of data and algorithms in two key questions: (1) How can we design simple, scalable, and predictable learning objectives for RL fine-tuning? and (2) Can we systematically curate data distributions to teach models dynamic, adaptive behaviors that supervised methods alone cannot? I explore these questions across a range of domains, from reasoning in mathematics and coding to enabling interaction in robotics and agent-based systems, and developing personalized, preference-aligned models.

## Education

---

<b>Stanford University</b> , Palo Alto, CA	Sep 2023 – Present
Ph.D. in Computer Science	Advisors: Prof. Chelsea Finn and Aviral Kumar
<b>University of California, Berkeley</b> , Berkeley, CA	Aug 2019 – May 2023
B.A. in Computer Science (Highest Distinction)	GPA: 3.986; Technical GPA: 4.0
Selected coursework: Deep Learning, Reinforcement Learning, Convex Optimization, NLP, and Info Theory.	

## Experience

---

<b>Microsoft Research</b> , New York, NY	Jun 2025 – Sep 2025
Research Scientist Intern.	
<b>Toyota Research Institute (TRI)</b> , Los Altos, CA	Jun 2024 – Sep 2024
Research Scientist Intern.	
<b>Google DeepMind Robotics</b> , Mountain View, CA	Apr 2023 – Feb 2024
Student Researcher.	
<b>X, the moonshot factory</b> , Mountain View, CA	Dec 2022 – Apr 2023
AI Resident.	

## Selected Publications & Preprints

---

(\* equal contribution; selection)

- ◇ **Learning to Discover Abstractions for LLM Reasoning.** Y. Qu\*, **A. Singh\***, Y. Lee\*, A. Setlur, R. Salakhutdinov, C. Finn, A. Kumar. Reasoning, Attention & Memory Workshop, CoLM 2025 (Oral).
- ◇ **Few-Shot Preference Optimization of Synthetic Preferences Elicits Effective Personalization to Real Users.** **A. Singh\***, S. Hsu\*, K. Hsu, E. Mitchell, T. Hashimoto, A. Sharma, C. Finn. NeurIPS 2025.
- ◇ **Cognitive Behaviors that Enable Self-Improving Reasoners.** K. Gandhi, A. Chakravarthy, **A. Singh**, N. Lile, N. D. Goodman. CoLM 2025 (Oral).
- ◇ **Improving Test-Time Search in LLMs with Backtracking Against In-Context Value Verifiers.** **A. Singh**, K. Arora, S. Keh, T. Hashimoto, C. Finn, A. Kumar. ICLR 2025 Workshop (Best Paper Runner-Up).
- ◇ **Big-Math: A Large-Scale, High-Quality Math Dataset for RL in LLMs.** A. Albalak, D. Phung, N. Lile, R. Rafailov, K. Gandhi, L. Castricato, **A. Singh**, *et al.* CoLM 2025.
- ◇ **Personalized Preference Fine-tuning of Diffusion Models.** M. Dang\*, **A. Singh\***, L. Zhou, S. Ermon, J. Song. CVPR 2025.
- ◇ **Test-Time Alignment via Hypothesis Reweighting.** Y. Lee, J. Williams, H. Marklund, A. Sharma, E. Mitchell, **A. Singh**, C. Finn. Preprint, 2024.

- ◇ **Towards System 2 Reasoning in LLMs: Learning How to Think With Meta Chain-of-Thought.** V. Xiang, C. Snell, K. Gandhi, A. Albalak, **A. Singh**, C. Blagden, D. Phung, R. Rafailov, N. Lile, D. Mahan, L. Castricato, J.-P. Franken, N. Haber, C. Finn. Position Paper, 2024.
- ◇ **Adaptive Inference-Time Compute: Large Language Models Can Predict If They Can Do Better, Even Mid-Generation.** R. Manvi, **A. Singh**, S. Ermon. Preprint, 2024.
- ◇ **Preference Fine-Tuning of LLMs Should Leverage Suboptimal, On-Policy Data.** **A. Singh\***, F. Tajwar\*, A. Sharma, R. Rafailov, J. Schneider, T. Xie, S. Ermon, C. Finn, A. Kumar. ICML 2024.
- ◇ **D5RL: Diverse Datasets for Data-Driven Deep Reinforcement Learning.** R. Rafailov\*, K. B. Hatch\*, **A. Singh**, A. Kumar, L. Smith, I. Kostrikov, P. Estruch, P. J. Ball, J. Wu, S. Levine, C. Finn. RLC 2024.
- ◇ **Robotic Offline RL from Internet Videos via Value-Function Pre-Training.** C. Bhateja\*, D. Guo\*, D. Ghosh\*, **A. Singh**, M. Tomar, Q. Vuong, Y. Chebotar, S. Levine, A. Kumar. ICRA 2024.
- ◇ **Open X-Embodiment: Robotic Learning Datasets and RT-X Models.** Open X-Embodiment Collaboration. CoRL 2024.
- ◇ **Offline RL With Realistic Datasets: Heteroskedasticity and Support Constraints.** **A. Singh\***, A. Kumar\*, Q. Vuong, Y. Chebotar, S. Levine. NeurIPS 2023.
- ◇ **Cal-QL: Calibrated Offline RL Pre-Training for Efficient Online Fine-Tuning.** M. Nakamoto\*, Y. Zhai\*, **A. Singh**, M. S. Mark, Y. Ma, C. Finn, A. Kumar, S. Levine. NeurIPS 2023.
- ◇ **Pre-Training for Robots: Offline RL Enables Learning from a Handful of Trials.** A. Kumar\*, **A. Singh\***, F. Ebert\*, M. Nakamoto, Y. Yang, C. Finn, S. Levine. RSS 2023.
- ◇ **RT-2: Vision-Language-Action Models Transfer Web Knowledge to Robotic Control.** Google DeepMind Robotics. ICRA 2023.
- ◇ **When Should We Prefer Offline Reinforcement Learning Over Behavioral Cloning?.** A. Kumar, J. Hong, **A. Singh**, S. Levine. ICLR 2022.
- ◇ **A Workflow for Offline Model-Free Robotic Reinforcement Learning.** A. Kumar\*, **A. Singh\***, S. Tian, C. Finn, S. Levine. CoRL 2021 (Oral).
- ◇ **A Mobile Application for Keyword Search in Real-World Scenes.** S. Pundlik, **A. Singh**, G. Baghel, V. Baliutaviciute, G. Luo. *IEEE J. Transl. Eng. Health Med.*, 2019.

## Teaching & Mentorship

---

Mentoring Program Organizer, Stanford CS Undergraduate Mentoring Program (2024-Present).  
 Program Coordinator & Mentor, Deep Learning Portal (2024).  
 Graduate Student Instructor, CS224r (Spring 2025).  
 Undergraduate Student Instructor, CS285 (Fall 2021, Fall 2022), CS188 (Spring 2022).

## Honors

---

NSF Graduate Research Fellowship (2023–2028).  
 CRA Outstanding Undergraduate Researcher Award Finalist (2022).  
 Dean’s List, UC Berkeley (2019–2023).

## Invited Talks and Service

---

Invited Talks at BeNeRL, Salesforce Future AI Forum, and ServiceNow  
 Reviewer (w/honors) for ICLR, NeurIPS, ICML, CoRL, ICRA (RAL), AAI, and CoLM.