

Tianwei Ni

✉ twni2016@gmail.com  [twini2016.github.io](https://github.com/twini2016)  [LinkedIn](#)  [GitHub](#)  [Google Scholar](#)  Montreal, Canada

EDUCATION

Ph.D. in Computer Science, **Mila and University of Montreal**, Canada Sept 2021 - Present

Thesis submitted: *History-Dependent RL: Memory, Representation, and Adaptability*

Advisor: [Pierre-Luc Bacon](#)

M.S. in Machine Learning, **Carnegie Mellon University**, Pittsburgh, USA Aug 2019 - Dec 2020

Research topics: Reinforcement Learning, Human-Agent Collaboration

B.S. in Computer Science, **Peking University**, Beijing, China Sept 2015 - Jul 2019

EXPERIENCE

Applied Scientist Intern, **Amazon Web Services (AWS)**, Santa Clara, USA Sept 2024 - Jan 2025

Supervisors: [Rasool Fakoor](#), [Allen Nie](#). LLM post-training, reasoning, and test-time search (TMLR 2025).

Research Intern, **Allen Institute for AI (AI2)**, Seattle (remote), USA Jun 2021 - Aug 2021

Supervisors: [Jordi Salvador](#), [Luca Weihs](#). Embodied AI with RL (WACV 2023).

Research Assistant, **Carnegie Mellon University**, Pittsburgh, USA Jan 2021 - Jun 2021

Supervisors: [Benjamin Eysenbach](#), [Ruslan Salakhutdinov](#). Partially observable RL and meta-RL (ICML 2022).

SELECTED PUBLICATIONS

Tianwei Ni, Vineet Jain, Akash Karthikeyan, Pierre-Luc Bacon. [From Deployable Policies to Adaptive Priors in Offline Reinforcement Learning](#). *Preprint (2026)*.

Tianwei Ni, Esther Derman, Vineet Jain, Vincent Taboga, Siamak Ravanbakhsh, Pierre-Luc Bacon. [Long-Horizon Model-Based Offline Reinforcement Learning Without Conservatism](#). *Under review (2026)*.

Lu Li, **Tianwei Ni**, Yihao Sun, Pierre-Luc Bacon. [The Three Regimes of Offline-to-Online Reinforcement Learning](#). *Under review (2026)*.

Tianwei Ni, Allen Nie, Sapana Chaudhary, Yao Liu, Huzefa Rangwala, Rasool Fakoor. [Offline Learning and Forgetting for Reasoning with Large Language Models](#). *TMLR 2025*.

Tianwei Ni, Benjamin Eysenbach, Erfan Seyedsalehi, Michel Ma, Clement Gehring, Aditya Mahajan, Pierre-Luc Bacon. [Bridging State and History Representations: Understanding Self-Predictive RL](#). *ICLR 2024* and *NeurIPS Workshop on SSL (Oral)*.

Tianwei Ni, Michel Ma, Benjamin Eysenbach, Pierre-Luc Bacon. [When Do Transformers Shine in RL? Decoupling Memory from Credit Assignment](#). *NeurIPS 2023 (Oral)* and *NeurIPS Workshop on FMDM*.

Tianwei Ni, Benjamin Eysenbach, Ruslan Salakhutdinov. [Recurrent Model-Free RL Can Be a Strong Baseline for Many POMDPs](#). *ICML 2022*.

ADDITIONAL PUBLICATIONS

Ziyan Luo, **Tianwei Ni**, Pierre-Luc Bacon, Doina Precup, Xujie Si. [Understanding Behavioral Metric Learning: A Large-Scale Study on Distracting RL Environments](#). *RLC 2025*.

Michel Ma*, **Tianwei Ni**, Clement Gehring, Pierluca D'Oro*, Pierre-Luc Bacon. [Do Transformer World Models Give Better Policy Gradients?](#) *ICML 2024* and *ICLR Workshop on GenAI4DM (Oral)*.

Tianwei Ni, Kiana Ehsani, Luca Weihs, Jordi Salvador. [Towards Disturbance-Free Visual Mobile Manipulation](#). *WACV 2023*.

Huao Li, **Tianwei Ni**, Siddharth Agrawal, Fan Jia, Suhas Raja, Yikang Gui, Dana Hughes, Michael

Lewis, Katia Sycara. [Individualized Mutual Adaptation in Human-Agent Teams](#). *IEEE Transactions on Human-Machine Systems* 2021.

Tianwei Ni*, Harshit Sikchi*, Yufei Wang*, Tejus Gupta*, Lisa Lee, Benjamin Eysenbach. [f-IRL: Inverse Reinforcement Learning via State Marginal Matching](#). *CoRL* 2020.

SELECTED AWARDS

RBC Borealis AI 2023-2024 Fellowship (10 awardees in Canada)

NeurIPS 2023 Scholar Award

Wang Shengdi Scholarship, Peking University, 2018

Outstanding Research Award, Peking University, 2017 and 2018

TALKS AND BLOG POSTS

Blog post on [Adaptive Policy Prior in the Era of Experience: Rethinking the Goal of Offline RL](#)

Also a talk at Mila RL Sofa Reading Group, Feb 2026

Blog post on [RL Fine-Tuning Is a Stability-Plasticity Problem](#), Feb 2026

Invited talk on [Bridging State and History Representations: Understanding Self-Predictive RL](#)

Intel AI Lab (Host: Mariano Phielipp), Mar 2024

Invited talk on [When Do Transformers Shine in RL? Decoupling Memory from Credit Assignment](#)

NeurIPS 2023 Oral, Dec 2023

Guest lecture on [A Modern Take on Partially Observable MDPs](#)

Mila course on [Reinforcement Learning and Optimal Control](#) (Host: Pierre-Luc Bacon), Dec 2023

Invited talk on [Recurrent Model-Free RL Can Be a Strong Baseline for Many POMDPs](#)

Donders Institute (Host: Luca Ambrogioni), Jan 2022

OPEN-SOURCE PROJECTS

Transformer-based RL for POMDPs (github.com/twni2016/Memory-RL) with 70+ stars, 2023

Recurrent RL for POMDPs (github.com/twni2016/pomdp-baselines) with 340+ stars, 2021

SKILLS

Coding Python, Bash, Slurm, Git, LaTeX, HTML

Deep Learning NumPy, PyTorch, JAX (Equinox/Flax)

RL and LLMs OpenAI Gym, HuggingFace TRL (SFT, RLHF), vLLM, distributed training

Languages English (full professional), Mandarin Chinese (native)

ACADEMIC SERVICES

Conference Reviewer ICML (2023-), NeurIPS (2023-), ICLR (2024-), RLC (2025-)

Workshop Organizer RLC 2026 The Continual RL Workshop (under review)

Teaching Assistant 10-703 Deep RL, CMU, Fall 2020

Admission Member Mila (PhD and MSc, Fall 2024), CMU MLD (MS, Fall 2020)

Mentorship Mentored 3 PhD students and 2 MSc students