

Memo Ozdincer

memo@cs.toronto.edu · [GitHub](#) · [Google Scholar](#) · [LinkedIn](#) · [Website](#)

LLM Post-Training · Reinforcement Learning · Agent Robustness

First-author accepted papers: **COLM 2026** · **ICML 2026** **AI4Good**

Research availability: May 2027 – Aug 2028

AI RESEARCH EXPERIENCE

Jinesis Lab, Vector Institute · AI Researcher

Sep 2025 – Present

Advised by Zhijing Jin & Bernhard Schölkopf

Designed a Training Method Reducing Prompt-Injection Attack Success to 0–4.1%

First Author · [COLM 2026](#)

- Designed Odile, a weight-level defense with lower prompt-injection attack success (0–4.1%) than every published defense we evaluated, including Meta SecAlign. [Release](#)
- Retained $\geq 94\%$ of baseline capability using only 184 paired injected/clean traces.

Post-Trained Expert LLMs, Improving 120B Nemotron by 7.4% on Clinical Tasks

On-premise deployment @ RWTH-UKA

- Improved Nemotron-120B by 7.4% and Qwen3-32B by 9.2% on RWTH's held-out renal oncology tasks.
- Designed and built the full RLVR stack (distributed rollouts, environments, verifiers) for 12K+ patient-history tasks, with expert-feedback loops on a 40K+ H100-hour allocation.

Built Agents That Resist Poisoned Context, Reducing Failed Trajectories by 81%

First Author · [AI4Good @ ICML 2026](#)

- Cut task failure from 59% to 11% across 8B–80B agents when conflicting or poisoned instructions derail them.
- Used representation fine-tuning on an 8K-trace poisoned- and conflicting-context harness (flight booking, coding, web search) to teach agents to prioritize true instructions.

Developed an RL Algorithm Matching GRPO's Performance with +67.7% Distinct Correct Solutions

- Identified that RL on Lean collapses onto a few dominant tactics, and designed Restriction-RL to block them and force exploration of alternative proof strategies. [Results](#)
- Matched GRPO's performance with 67.7% more distinct correct proofs and 31.9% shorter proofs on average.

Matter Lab, University of Toronto · AI Research Assistant

Aug 2025 – May 2026

Advised by Alán Aspuru-Guzik

- Designed custom numerical optimizers that beat Sella by 34% on transition-state finding across 33M+ simulations. [Paper](#) · [Code](#)
- Built RL-guided diffusion samplers (PyTorch/CUDA) to steer sampling toward chemically valid transition states.

SERIS, National University of Singapore · ML Engineer Intern

May 2025 – Aug 2025

Advised by Erik Birgersson

- Built Transformer surrogates in PyTorch that replace $\sim 4,800$ -second drift-diffusion simulations with millisecond inference, $100,000\times$ faster. [Paper](#) · [Code](#)
- Approximated the solver to 0.104% MAE on 115K held-out device configurations, with CUDA-level debugging.

PROJECTS

Built a Proof-Checking Library, Speeding Up RL Verification 27 $\times$

- Released a Python/Lean 4 library (prefix batching, certificate reuse, isolated persistent subprocesses) that speeds up repeated proof tactics 27 $\times$ on average, up to 325 $\times$, and powers the Restriction-RL experiments. [PyPI](#)

EDUCATION

University of Toronto · BAsC, Engineering Science (Engineering Physics)

Expected: May 2029 (after 16-month work term)

Skills: Python · PyTorch · JAX · vLLM · FSDP · Ray · CUDA · C++ · W&B