

Zihe (Zach) Liu

flowingpurplecrane.com | [GitHub](#) | [Google Scholar](#) | [LinkedIn](#) | zl559@cam.ac.uk

Distributed RL, Long-horizon Alignment

EDUCATION

University of Cambridge

Master of Engineering, Information and Computer Engineering

July 2026 - Present

Masters Thesis: *Second Order Forward Optimization for Natural Policy Gradients*

- Develop efficient Gauss-Newton optimizer for RL, on custom OCaml ML compiler and AD stack (JAX-like)

Bachelor of Arts, Information and Computer Engineering — First Class Honors

Sep 2023 - June 2026

Coursework: RL, Deep Learning, Probabilistic Inference, [Information Theory](#), Optimization, Computer Architecture

- David Newland [Prize](#) for Design (best undergrad project)

EXPERIENCE

ERA: AI

Cambridge, UK

Research Fellow — Mentor: Elizaveta Tennant, *Google DeepMind*

July 2026 - Present

- [AI research program](#) (top 0.5% of 6,000+ applicants) working on long-horizon RL for LLM alignment
- Develop [multi-turn RL algorithm](#) to reduce ethical violations 17% at $2\times$ training horizon (80K tokens, 128 turns)
- Scale distributed RL in verl (FSDP, async vLLM) on $4\times B300$. Fused GatedDeltaNet kernels raised MFU 1.75% $\rightarrow$ 27%
- Traced GPU hang to FlashInfer kernel race, fixing it enabled $4\times$ larger batch-sizes hence inference throughput

Amazon

London, UK

Software Development Engineer Intern

July - Oct 2025

- Built ML (*CatBoost*) triaging system to root-cause fatal Prime Video errors, with *AWS Bedrock* LLM log analysis
- Scaled throughput $100\times$ ($3\rightarrow 300$ sessions/cluster, 20GB logs) by batching SQL queries on *AWS Lambda*, *Athena*
- Refactored monolithic *AWS CDK* stack, and unblocked a *Fargate* pipeline stalling deployment for 4 months
- *Turned down full-time return offer to work on RL research*

UCL Centre for AI

London, UK

Student Researcher — Mentor: Peter Hayes, *Humanloop (now Anthropic)*

Feb - May 2025

- Developed *Critique Direct Preference Optimization*: generative critique for sample-efficient DPO | [Blog](#)
- Distributed training on *Qwen-2.5 3B* models with $16\times H100$ GPUs, on UCL HPC facility | [GitHub](#)

Singapore AI Safety Institute, Infocomm Media Development Authority

Singapore

Research Engineer Intern

Dec 2024 - July 2025

- Research industry evaluation metrics for LLM reliability and robustness | [Report](#)
- Benchmark hallucination detectors on *HaluBench* (simple CoT LLM-judge outperforms all) | [GitHub](#)

Alignment Research Engineer Accelerator ([ARENA](#))

London, UK

Research Engineer — 1 of 2 undergrads in a mostly-PhD cohort of ~ 20

Sep - Oct 2024

- Built GPT-2 LLM from scratch, reverse-engineered its internal reasoning circuits, and fine-tuned via RLHF (PPO)

Cambridge AI Safety Hub

Cambridge, UK

Operations Lead, Student Researcher

Oct 2023 - Present

- Organized AI safety workshops, hackathons, research reading groups on interpretability, alignment | [Rt Talk](#)

Changi Airport Group

Singapore

Machine Learning Engineer Intern

July - Aug 2023

- Trained P2PNet (object detection CNN model) to localize and count luggage trolleys (*MAE decreased by 5x compared to previous systems*). Quantized model with ONNX to deploy on edge (Nvidia Jetson Nano)

Republic of Singapore Air Force

Singapore

Security Trooper, Corporal First Class

Jan 2021 - Jan 2023

- Led Immediate Response Force team (2IC) to win APGC Eagles Challenge 2022. Best Airman Award.

PUBLICATIONS

Z. Liu, D. Le, E. Tennant. [Multi-Turn RL for Long-Horizon Moral Alignment](#). *NeurIPS 2026 AI4GOOD Workshop*
J. Robinson, L. Franceschetti, M. Lovett, D. Le, K. Sen, N. Narayanan, **Z. Liu**, E. Tennant. [Reasoning-Behavior Coherence Beyond the Rewarded Context](#). *NeurIPS 2026 AIWILD*
M. Saghir, N. Raghavendra, **Z. Liu**, E. Gunter. [NGD converges to less degenerate solutions than SGD](#). *arXiv*

AWARDS

SG Digital AI Scholarship | 500k\$+ Full undergraduate funding, for exceptional Singaporeans pursuing AI overseas
GB3 RISC-V CPU Competition 2026 | *1st of 8 teams*. Optimize tiny CPU with instruction cache, lookaheads | [GitHub](#)
SF2 Image Compression 2026 | *1st of 13 teams*. Built JPEG compression scheme with DCT, intra prediction | [Results](#)
Changi Airport Group Scholarship | Offered full undergraduate funding for fast-track graduate programme

SKILLS

Programming Languages: Python, C/C++, OCaml, Verilog/SystemVerilog, TypeScript/JavaScript, MATLAB, Java
ML Systems: verl, TRL, vLLM, PyTorch FSDP, DeepSpeed-ZeRO, NumPy, JAX
AWS/Cloud: CloudFormation, CDK, Fargate, Lambda, S3, Athena
Non-ML: HPC (Sun Grid Engine), FPGA/Arduino/Embedded Systems, Git, SQL