Yuchen Lee (李聿宸)

Hsinchu, Taiwan | eason.yuchen.lee@gmail.com
github.com/Windmill10 | linkedin.com/in/yuchen-lee-cs

EDUCATION

National Tsing Hua University (NTHU) Hsinchu, Taiwan
B.S. in Computer Science *Sep 2023 – Present; expected July 2027*
GPA: 4.01/4.30 | *TOEFL iBT:* 108 (2025)
Academic Excellence Award: Top two in class, Spring 2025.
Coursework: Operating Systems, Computer Architecture, Networks, Databases, Deep Learning, Storage Systems.

RESEARCH EXPERIENCE

SSD-Based KV-Cache Offloading — NTHU syslab
Undergraduate Researcher — Advisor: Prof. Yun-Chih Chen *Feb 2026 – Present*

- Characterized KV-cache write waste in one month of Claude Code and Codex traffic, identifying context-compaction orphans, partial-tail churn, and cache retention after session termination.
- Led the core method design: write-back DRAM staging, deterministic write suppression, and session-grouped ZNS placement; implemented in LMCache and validated with FEMU and a live vLLM/LMCache deployment.
- Simulation: **29.2%** (single-node) and **16.3%** (multi-node) lower system-level write amplification than LMCache.

Training Signals for Story Generation — NTHU DATALAB
Undergraduate Researcher — Advisor: Prof. Shan-Hung Wu *Mar 2025 – Present*

- Built a 130K-chapter Wattpad corpus and engagement reward model; showed text-only engagement prediction is infeasible (**89%** of vote-rate variance lies between stories).
- Showed that LLM "slop" in OLMo-2 7B comes from ChatGPT-derived SFT data, not from the training objectives.
- **RL-only** training of Qwen2.5-1.5B Base with verifiable rewards: valid stories **12%** → **91%**, judged on par with the official Instruct model; openings narrowed (top first-token prob. 0.07 → 0.47).

PUBLICATION

Yu-Chen Lee, Chih Yen Chen, Ming-Yi Kuo, Yun-Chih Chen.
"Multi-Agent LLM Serving with SSD-Based KV-Cache Offloading." *ASP-DAC 2027, accepted regular paper.*

ENGINEERING EXPERIENCE

Microsoft Taipei, Taiwan
Software Engineer Intern, Azure Maps *Sep 2026 – Present*

- Full-stack bug fixes and feature work for Azure Maps in Microsoft Fabric.

Industrial Technology Research Institute Remote
Research Intern *Jul 2026 – Present*

- Build an agent-teams pipeline as a Claude Code plugin: analyst, architect, developer, QA, and lead agents coordinate software tasks through a GitHub Projects board, with human approval gates.

NeuinX Hsinchu, Taiwan
Frontend Developer *Sep 2025 – Present*

- Build exercise-analytics UIs in React/TypeScript with Three.js court visualizations and Recharts dashboards.

LEADERSHIP

MeiChu Hackathon
Development Team Co-Lead *Sep 2025 – Sep 2026*

- Co-led the team building the event app; scaffolded the PWA, wrote the architecture guide, built the notification system and made the architecture decisions; about 300 people used it during the event.

TECHNICAL SKILLS

Languages: Python, Rust, C/C++, TypeScript, JavaScript | **Tools:** Git, Linux
Systems and ML: vLLM, LMCache, PyTorch, HuggingFace (TRL) | **Frontend:** React, Three.js, Recharts