Minghao (Mark) Liu

The Hong Kong University of Science and Technology, Hong Kong SAR, China
mliuby@connect.ust.hk | mark-minghao-liu-web.github.io

Education

Hong Kong University of Science and Technology (HKUST) *2026–*
PHD in Computer Science

Hong Kong University of Science and Technology (HKUST) *2022 – 2026 (Expected)*
BEng in Computer Science | *Minor in Mathematics*
CGA: 3.902 / 4.3 (Top 2%)

Washington University in St. Louis *Fall 2024*
Exchange Student, McKelvey School of Engineering
GPA: 3.94 / 4.00

Research Interest

- **NLP** with emphasis on **interpretable**, **reliable**, and **socially aware AI systems**.
- Developing **interpretable methods** to guide (vision-)language model behavior and enhance **real-world reliability**.
- Diagnosing **limitations** of language models via **explainable, systematic benchmarking**.
- Modeling **social and economic phenomena** (public policy, financial markets) using **(multi-)agent simulations**.
- Designing **dynamic, scalable learning environments** and **robust evaluation metrics** for LLM agents to learn, reason, and perform beyond human-level competence.

Publications

- **LeanForPhysics: Comprehensive Reasoning Framework for University-level Physics in Lean4**
  Yuxin Li[*], **Minghao Liu**[*], Ruida Wang[*†], Ji Wenzhao, Zhitao He, Rui Pan, Junming Huang, Tong Zhang, Yi R. Fung
  ICLR 2026 (in submission)
  [*]Equal contribution

- **MedEBench: Diagnosing Reliability in Text-Guided Medical Image Editing**
  Minghao Liu, Zhitao He, Zhiyuan Fan, Qingyun Wang, Yi R. Fung
  Findings of EMNLP 2025

- **Scaling Environments for LLM Agents: Fundamentals, Approaches, and Future Directions**
  Yuchen Huang, Sijia Li, Zhiyuan Fan, **Minghao Liu**, Wei Liu, Yi R. Fung
  SEA @ NeurIPS 2025

- **A Benchmark for Evaluating Purchase Intention Comprehension Abilities of Large Language Models in E-commerce**
  Wenxuan Ding[*], Weiqi Wang[*], Sze Heng Douglas Kwok, **Minghao Liu**, Tianqing Fang, Jiaxin Bai, Xin Liu, Changlong Yu, Zheng Li, Chen Luo, Qingyu Yin, Bing Yin, Junxian He, Yangqiu Song
  Findings of EMNLP 2024
  [*]Equal contribution

Projects & Research Experience

UROP, HKUST *Jun–Aug 2023* *Advisor: Dan Xu*

- Developed a **conditional diffusion model** for monocular depth estimation from RGB images.
- Built an **Efficient-UNet** with residuals and optimized up/down-sampling to preserve depth structures.
- Designed a two-step **depth infilling algorithm** for handling missing values in NYU Depth V2.
- Applied **Step-Unrolled Denoising (SUD)** and masked losses to mitigate distribution shift on incomplete maps.

- Finetuned under compute limits to improve depth completeness and global scene consistency.

UROP, HKUST *Sep–Dec 2023* *Advisor: Yu Hu*

- Modeled **brain-wide neural dynamics in zebrafish** using a recurrent **Firing Rate Network**.
- Implemented **firing rate evolution equations** with Poisson inputs and synaptic filtering.
- Simulated large-scale circuits and trained synaptic connectivity via **Partial In-Network Training (PINning)** with Recursive Least Squares.
- Analyzed structural patterns in connectivity to identify potential subnetworks and functional motifs.

KnowComp Group, HKUST *Feb 2024 – Sep 2024* *Advisor: Yangqiu Song*

- **BrainASER (Led by Shi Haochen)**: Studied correspondences between neural activity and **knowledge graph** structures.
- Aligned fMRI data (Narratives dataset) with story-based stimuli to analyze brain-language interactions.
- Developed brain-inspired representations for downstream NLP tasks leveraging structural similarities with knowledge graphs.
- **IntentionQA (Led by Ding Wenxuan)**: Built a benchmark to evaluate LMs' understanding of purchase intentions in E-commerce.
- Preprocessed data, aligned products with intentions via ASER, and generated negative distractors.
- Evaluated 19 LMs, identifying reasoning limitations in predicting user intent and handling real-world E-commerce scenarios.

Washington University in St. Louis *Sep 2024 – Dec 2024* *Advisor: Marion Neumann*

- Developed an **inductive recommendation system** for new e-commerce products using the Amazon Co-Purchasing Network.
- Constructed a co-purchasing graph with 519K nodes and 964K edges, encoding product features, categories, and structural metrics.
- Applied a **modified GraphSAGE** for link prediction to enable recommendations for isolated nodes with limited data.
- Designed scalable, real-time updates for adaptive recommendations on dynamic product catalogs.

RenLab, HKUST *Feb 2025 – June 2025* *Advisor: Yi R. (May) Fung, Qingyun Wang*

- Researched **text-guided medical image editing** and developed evaluation frameworks for multimodal models.
- Contributed to **MedEBench**, a benchmark of 1,182 clinical image-prompt triplets across 70 tasks and 13 anatomical regions.
- Designed clinically grounded metrics for **Editing Accuracy, Contextual Preservation, and Visual Quality** using ROI-based assessments and attention-grounding analysis.
- Evaluated seven state-of-the-art models, identifying common failure patterns in medically meaningful edits.

Final Year Project *May 2025 – Oct 2025* *Advisors: Ruida Wang[†], Tong Zhang, Yi R. (May) Fung*

- Initiated **PHYSlib**, a Lean4 library for formalizing university-level physics concepts and statements on top of a **UnitSystem**, enabling rigorous machine-verified reasoning.
- Built **LeanPhysBench**, a benchmark of 200 textbook and competition-level formalized physics statements.
- Conducted systematic evaluations on 5 open-source and 3 closed-source models, revealing poor knowledge transfer in Lean-expert models.

RenLab, HKUST *Sep 2025 – Ongoing*

Advisor: Yi R. (May) Fung

- Leading the development of an **LLM-agent simulation framework** for public policy deliberation and policy-maker training.
- Designing **multi-agent debate environments** that capture real-world legislative dynamics, including negotiation, value conflict, and consensus formation.

- Building datasets from **official policy and legislative records** across domains such as education, environment, and social welfare, distinguishing passed vs. failed proposals.
- Reformulating policy objectives into **debatable issues** and orchestrating **multi-round simulations** where LLM-based agents represent diverse stakeholders.
- Investigating how deliberative reasoning and social interaction among LLM agents can model **decision change points** and **policy trajectory outcomes**.
- Aiming to construct a **dialectical policymaking model** capable of explaining not only outcomes but also the reasoning processes behind public decisions.

Standardized Tests

- IELTS: 7.0

Awards & Scholarships

- First Prize – 37th Chinese Physics Olympiad (Provincial Level) 2020
- First Prize – 38th Chinese Physics Olympiad (Provincial Level) 2021
- First Prize – Chinese Mathematical Olympiad in Senior (Provincial Level) 2021
- Talent Development Scholarship – HKSAR Government Scholarship Fund 2023
- University's Scholarship Scheme for Continuing Undergraduate Students 2023–24
- HKUST Alumni Endowment Fund High Flyers Program Scholarship 2023–24
- HKSAR Government Scholarship Fund – Reaching Out Award 2024–25
- Dean's List 2022–25

Extracurricular Activities

- Mechanical Engineer – HKUST RoboMaster Team ENTERPRIZE Sep 2022 – Feb 2023