

Zehao Liu

State College, PA 16801 | zml5418@psu.edu | (814)699-2543 | [linkedin.com/in/zehao-liu-4631a5259](https://www.linkedin.com/in/zehao-liu-4631a5259)

Biography

Zehao Liu is presently a Ph.D candidate in the College of Information Sciences and Technology at Pennsylvania State University, advised by Professor Vasant Honavar. He earned his M.S. and B.Eng. degrees from Southwest Jiaotong University.

His current research interest include post-training for large language models and multi-modal language models, reinforcement learning, on-policy distillation, AI for healthcare.

Skills

Programming Languages: Python, C++, C, MATLAB

Machine Learning: Deep Learning, Time Series Modeling, Clustering, Representation Learning, Unsupervised Learning

LLM & Multimodal Models: LLM/VLM Post-training and Alignment, SFT, Reinforcement Learning, GRPO, LoRA

Frameworks: PyTorch, Hugging Face Transformers, DeepSpeed, TRL, VERL, vLLM

Systems & Tools: Git, Docker, Slurm

Education

Penn State University

Ph.D in informatics, advised by Professor Vasant Honavar

State College, PA, USA

Aug.2023 - now

Southwest Jiaotong University

M.S in Computer Science and Engineering, advised by Professor Hongjun Wang

Chengdu, Sichuan, China

Sep.2020 - Jun.2023

Southwest Jiaotong University

B.E in Biomedical Engineering

Chengdu, Sichuan, China

Sep.2016 - Jun.2020

Projects

Skin-R1: Grounded Clinical Reasoning for Dermatology VLMs (<https://github.com/l593191569/Skin-R1>):

Post-trained a **Qwen2.5-VL-7B** vision-language model with **SFT + GRPO reinforcement learning**. Constructed a textbook-grounded reasoning dataset with hierarchical taxonomy and differential-diagnosis graphs and designed structure-aware RL rewards for LoRA training, improving reasoning accuracy and generalization on dermatology benchmarks.

SC-SDPO: Difficulty-Aware Self-Distillation for LLM Reasoning: Proposed SC-SDPO, a pass-rate-weighted variant of Self-Distillation Policy Optimization (SDPO) for LLM post-training. Derived a theoretical justification showing GRPO's advantage normalization absorbs the $p(1-p)$ learnability factor, motivating a per-question loss weight $w = [\hat{p}(1-\hat{p})]^{1/2}$ that concentrates gradient budget on intermediate-difficulty questions. Implemented batch-adaptive normalization as a zero-cost byproduct of on-policy rollouts, forming an implicit curriculum.

Auditing Thinking-Pattern Leakage in Hybrid Reasoning Models (Ongoing): Recent post-training methods report large math-benchmark gains (AIME, HMMT) under Qwen3's `enable_thinking=False` mode, but since both modes share one set of weights, these “non-thinking” gains may reflect **leaked reasoning ability** rather than genuine direct-inference improvement. We extract **hidden state vector directions** from base-model thinking traces and measure how far a post-trained model's non-thinking outputs **drift** along them, anchoring this geometric signal to text-level reasoning markers. We find leakage concentrated specifically along the **exploration axis**, and propose a **causal-mediation decomposition** that ablates the leaked direction to separate borrowed reasoning from real capability gain.

Self-Evolving Knowledge Graph Framework for Automated Research (Ongoing): Designed and built a general **self-evolving framework** for **automated research** — turning a corpus of papers into a structured, queryable **knowledge graph** that supports learning, literature review, and survey writing with minimal manual curation. Its core novelty is a schema that **grows itself**: the taxonomy, node fields, and relation types expand

adaptively at checkpoints as new papers arrive, rather than being fixed upfront, so the framework scales to any fast-moving research area where manual review cannot keep pace with publication velocity. Papers are auto-ingested (arXiv API, Semantic Scholar citation graphs), parsed into motivation → method → results → limitations nodes via LLM extraction, and linked by typed cross-paper edges (*extends*, *contradicts*, *mitigates*, *generalizes*, ...) that capture non-trivial relationships beyond citation — negative findings, mitigation chains, and shared limitations tracked as global entities for longitudinal analysis. A programmable CLI (outline, export-section, gaps) then drives survey generation and coverage auditing directly from the graph. Already deployed on **LLM self-distillation** research — mapping 205 papers across 8 years, 17 sub-domains, and 11 method categories into 318 typed relationship edges — where it surfaced cross-paper problem threads (e.g., diversity collapse, math degradation) and open gaps that directly informed my own work in this direction.

Publications and Preprints

- Honavar, V., Keshri, S.K., Ashtekar, N., & **Liu, Z.** (2026). PAC-Bayes Beyond Parameter Space: Behavioral Equivalence, Z-Information, and Exact Complexity Decomposition. arXiv preprint arXiv:2608.11465.
- **Liu, Z.**, Ren, W., Zhang, J., Zhao, T., Zhu, J., Li, X., & Honavar, V. G. (2025). Skin-R1: Clinical Knowledge-Guided Dermatological Diagnosis Using Vision-Language Models. arXiv preprint arXiv:2511.14900 (**Accepted by ECCV 2026**).
- **Liu, Z.**, Cao, Y., Chen, J., & Honavar, V. G. (2026). Restoring the Sweet Spot: Pass-Rate Weighted Self-Distillation for LLM Reasoning. arXiv preprint arXiv:2605.27765.
- Ren, W., Zhu, J., **Liu, Z.**, Zhao, T., & Honavar, V. (2025). A comprehensive survey of electronic health record modeling: From deep learning approaches to large language models. arXiv preprint arXiv:2507.12774.
- **Liu, Z.**, Wang, H., Chen, W., Wang, L., & Li, T. (2022). Bilateral discriminative autoencoder model orienting co-representation learning. *Knowledge-Based Systems*, 245, 108653.
- Yang, W., Wang, H., Zhang, Y., **Liu, Z.**, & Li, T. (2022). Self-supervised discriminative representation learning by fuzzy autoencoder. *ACM Transactions on Intelligent Systems and Technology*, 14(1), 1-18.

Professional Service

Journal reviewer: Information Science, Journal of Big Data

Awards, Scholarships & Grants

- National Postgraduate Scholarship of China
- Outstanding graduate student of Sichuan Province

Teaching Experience

- Teaching Assistant, “DS 330: Visualization & Visual Analytics”, Fall 2025, Spring 2026, Penn State University.
- Teaching Assistant, “DS 320: Data Integration & Fusion”, Spring 2025, Penn State University.
- Teaching Assistant, “Fundamentals of Computer Programming”, Spring 2021, Southwest Jiaotong University.