Hao Liu

Applied Scientist — Agentic AI Systems — Post-Training for Tool-Using LLMs — Ph.D. in Computer Science

📞 314-203-9875 ✉ haoliu4221@gmail.com 🔗 linkedin.com/in/hao-liu-929016161 🌐 haoliu-cola.github.io

PROFESSIONAL SUMMARY

Applied Scientist with a Ph.D. in Computer Science building production agentic systems and post-training for tool-using LLMs. Experience spans autonomous coding and DevOps agents, including orchestration, context gathering, evaluation-driven system improvement, and SFT/RL-based training for tool use, multi-step reasoning, and environment interaction. Published at ICLR, NeurIPS, and WWW, with a track record of translating research into deployed systems.

WORK EXPERIENCE

Applied Scientist
Amazon Web Services

Seattle, WA — *July 2025 – Present*

Kiro Autonomous Agent

- Built and shipped core orchestration workflows for a launched autonomous coding agent, spanning request understanding, repository grounding, plan generation, and delegated execution, in collaboration with applied scientists and engineers.
- Owned a delegation framework that translated plans into structured task execution, analyzed sub-agent results, and adaptively revised next actions in the execution loop.
- Investigated quality degradation in planning-based agent workflows, traced a major failure mode to information loss between planner and executor, and improved task success from 50% to 54%.
- Developed a context-gathering code exploration subagent that retrieved and summarized relevant repository context to help the main agent generate better plans when initial information was insufficient.

AWS DevOps Agent

- Building post-training pipelines for a DevOps investigation agent, focused on improving tool use, reasoning, and multi-step decision making in operational workflows.
- Built an SFT training pipeline and real-trajectory-based data curation workflow for agent behaviors including tool selection, schema-constrained invocation, result interpretation, and next-step intent generation.
- Developing RL-based training workflows for end-to-end agent interaction with tools and environment feedback in log investigation scenarios.

EDUCATION

Washington University in St. Louis
St. Louis, MO

Ph.D. in Computer Science & Engineering — *Sep. 2019 – May 2025*

Beijing Normal University
Beijing, China

B.S. in Mathematics — *Sep. 2015 – June 2019*

SELECTED RESEARCH

GOFA: A Generative One-For-All Model for Joint Graph Language Modeling　ICLR 2025

- Designed a graph foundation model that integrates GNN layers with LLM layers, leveraging the generative strengths of LLMs for free-form output and the structural learning capabilities of GNNs.
- Pioneered a novel pretraining paradigm focused on graph-level next-token prediction, facilitating large-scale self-supervised learning on various text-attributed graphs.
- Achieved state-of-the-art zero-shot performance on six node/link/graph-level datasets after instruction fine-tuning on a small number of data, demonstrating its potential as a foundation model.

One for All: Training One Graph Model for All Classification Tasks　ICLR 2024 Spotlight

- Developed the first graph foundation model capable of handling multiple classification tasks across various datasets and domains in supervised, few-shot, and zero-shot scenarios.

- Employed Large Language Models to standardize and unify graph data representation, transforming features of diverse graphs into consistent natural language formats, enabling multi-dataset training in the graph domain.
- Innovated a graph prompting paradigm, advancing the model's in-context learning capabilities.

Unsupervised Meta-Learning for Transductive Few-shot Node Tasks   WWW 2024
- Designed the first unsupervised meta-learning method for transductive few-shot node classification.
- Innovated an unsupervised meta-task construction method by leveraging the strengths of contrastive learning, enabling comprehensive use of graph nodes in the learning process.
- Achieved state-of-the-art performance on seven datasets, with at least 11.18% and up to 20.56% absolute accuracy improvement over existing meta-learning baselines.

PUBLICATIONS

Conference:

- GOFA: A Generative One-For-All Model for Joint Graph Language Modeling
  L. Kong*, J. Feng*, **Hao Liu***, C. Huang, J. Huang, Y. Chen, M. Zhang
  The Thirteenth International Conference on Learning Representations (**ICLR 2025**)

- One for All: Towards Training One Graph Model for All Classification Tasks
  Hao Liu*, J. Feng*, L. Kong*, N. Liang, D. Tao, Y. Chen, M. Zhang
  The Twelfth International Conference on Learning Representations (**ICLR 2024 Spotlight**)

- Graph Contrastive Learning Meets Graph Meta Learning: A Unified Method for Few-shot Node Tasks
  Hao Liu, J. Feng, L. Kong, D. Tao, Y. Chen, M. Zhang
  The Web Conference 2024 (**WWW 2024**)

- TabContrast: A Local-Global Level Method for Tabular Contrastive Learning
  Hao Liu, Y. Chen, B. Fritz, C. King
  NeurIPS 2023 Second Table Representation Learning Workshop (**NeurIPS 2023 TRL**)

- MAG-GNN: Reinforcement Learning Boosted Graph Neural Network
  L. Kong, J. Feng, **Hao Liu**, D. Tao, Y. Chen, M. Zhang
  Thirty-seventh Conference on Neural Information Processing Systems (**NeurIPS 2023**)

- Extending the Design Space of Graph Neural Networks by Rethinking Folklore Weisfeiler-Lehman
  J. Feng, L. Kong, **Hao Liu**, D. Tao, F. Li, M. Zhang, Y. Chen
  Thirty-seventh Conference on Neural Information Processing Systems (**NeurIPS 2023**)

Preprint:
- TAGLAS: An Atlas of Text-attributed Graph Datasets in the Era of Large Graph and Language Models
  J. Feng, **Hao Liu***, L. Kong*, Y. Chen, M. Zhang

- Time Associated Meta Learning for Clinical Prediction
  Hao Liu, M. Zhang, Z. Dong, L. Kong, Y. Chen, B. Fritz, D. Tao, C. King

(* indicates equal contribution)

TECHNICAL SKILLS

Programming: Python, C, Matlab
ML / LLM Frameworks: PyTorch, PyTorch Lightning, vLLM, TRL, VeRL, Scikit-learn, PyG
Agentic AI & Post-Training: Agent orchestration, tool-use workflows, SFT, RL-based training workflows, evaluation
Data / Developer Tools: Pandas, NumPy, Docker, Git, Jupyter, VS Code, Google Cloud Platform