

李克钰 Keaney Li

✉ chlorophyll@sjtu.edu.cn

☎ 13370856505

🌐 keyuli.space

🔍 Google Scholar

🐙 Github

🤗 Huggingface



个人优势

- 研究方向: *LLM Continue Pre-Training, Mid-Training; Agentic RL; Multi-Agent System*
- 以第一作者身份发表两篇 CCF A 类论文在 ICLR 2026, ACL 2026 Main Conference
- Claude Code, Codex 深度使用者, 打通数据筛选, 模型训练, 模型评测迭代管线
- 多机多卡 (256+) 模型训练经验, 熟悉 ray, megatron, sglang, slime, llama-factory, verl 等框架

教育经历

- 2024 – 2029 **博士, 上海交通大学 & 上海创智学院** 计算机科学与技术
导师: 王德泉老师, 刘鹏飞老师
- 2020 – 2024 **本科, 上海交通大学** 数学与应用数学
上海市优秀毕业生, 上海交大郭晨晨奖学金, 本科生奖学金, 优秀团员, 三好学生
学积分 92.6/100, GPA 4.03/4.3, 排名 4/45

实习经历

- Tencent Youtu Lab** 研究实习生 2026.05–now
Supervisor: Xiaoyu Tan
- Data / CPT: 数据筛选与能力诊断。**主导 Coding Agent 的 CPT/Midtrain 数据闭环; 基于 **Tencent WorkBuddy** 线上真实回流的 **674.9 万**条候选轨迹, 设计 DCP 任务分类、DDP 难度分层与 PPL-delta 能力缺口指标, 将 **673 万**条有效候选收敛为 **299.8 万**训练池和 **150 万**条可追溯 SFT 数据
 - Midtrain / SFT: 模型训练与数据配比。**在 Tencent Youtu 自研 **30B-A3B** 模型上开展 CPT 与 SFT; 依据筛选阶段各数据桶的模型能力表现动态调整 SFT 数据配比, 并在 **128/256 卡**上完成多机模型训练
 - Iteration: 评测、迭代与落地。**与评测团队在 SWE-bench-verified、Tencent WorkBuddy Bench 上完成评测并据此调优; 打通“发现-筛选-评测-配比-训练-回流”闭环, 最终在 wb-bench 上提升 **10%**, 显著增强模型的 coding、planning、web 能力, 供下游业务团队部署使用
 - Agentic RL: 智能体强化学习课题研究。**主导面向智能体与推理模型的在线强化学习及 On-Policy Distillation 研究, 围绕长轨迹稀疏奖励下的信用分配、过程级知识迁移与 Teacher Signal 偏差, 推进 ReSCC、Process-Tangent、FDD 三个方向。**ReSCC:** 基于 veRL/GRPO 构建带自蒸馏与轨迹级信用重加权的 Agent RL 框架, 在 WebShop、Search-QA、ALFWorld 及 3B/7B 模型上完成多组实验; 多个设置的单种子 Best-Val 达到或超过对应 SOTA 参考结果; **Process-Tangent OPD:** 探索过程表征变化信息辅助在线蒸馏, 在 8×H200、500-step、三种子评测中取得 58.66 Macro, 较 SOTA 提升 4.16; **FDD-OPD:** 研究多视角 Teacher Signal 的去偏与筛选机制, 完成数学、代码、科学混合数据上的 8×H200 训练闭环, Stage-o 信号 AUC 较 plain 方法提升 6.7%

实习经历（续）

- **Agentic RL** 三个方向均处于 *Work in Progress*，持续进行多随机种子、*matched baseline* 及下游 Agent 任务验证；积累了完整的 RL 训练、奖励设计、消融评测、*Checkpoint* 恢复与多卡实验经验
- **上海创智学院 · GAIR 实验室** 2025.05–2026.05
Supervisor: Pengfei Liu
- **AgencyBench**: 真实场景 Agent 评估，GitHub Star **100+**、Hugging Face 下载量 **10K+**
LIMI: 提出“少即是多”的 Agent 训练范式，以极少量高质量数据激发智能体能力
- **CoPR**: 基于真实代码 PR 演进序列的数据合成框架，显著提升 Agent 长程性能
DatasetResearch: 首个评估 Agent 自主发现与合成数据集能力的基准

研究成果

Conference Proceedings

- 1 K. Li, J. Gao, and D. Wang, “Aligned agents, biased swarm: Measuring bias amplification in multi-agent systems,” in *The Fourteenth International Conference on Learning Representations*.
- 2 K. Li, J. Shi, Y. Xiao, M. Jiang, J. Sun, Y. Wu, D. Fu, S. Xia, X. Cai, T. Xu, et al., “Agencybench: Benchmarking the frontiers of autonomous agents in 1m-token real-world contexts,” in *Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2026, pp. 7422–7440.
- 3 J. Gao, J. Zhao, K. Li, and D. Wang, “Kan-mixer: Kolmogorov-arnold networks for gene expression prediction in plant species,” in *European Conference on Computer Vision*, Springer, 2024, pp. 135–150.

Articles

- 1 D. Fu, Y. Wu, X. Cai, L. Ye, S. Xia, Z. Huang, W. Si, T. Xu, J. Sun, K. Li, et al., “Argo: Asynchronous rollout with human guidance for research agent optimization,”
- 2 M. Jiang, D. Fu, J. Shi, J. Zeng, W. Si, K. Li, X. Li, Y. Xiao, W. Li, D. Wang, et al., “Davinci-agency: Unlocking long-horizon agency data-efficiently,” *arXiv preprint arXiv:2602.02619*, 2026.
- 3 D. Fu, Y. Wu, X. Cai, L. Ye, S. Xia, Z. Huang, W. Si, T. Xu, J. Sun, K. Li, et al., “Interaction as intelligence part ii: Asynchronous human-agent rollout for long-horizon task training,” *arXiv preprint arXiv:2510.27630*, 2025.
- 4 K. Li, M. Jiang, D. Fu, Y. Wu, X. Hu, D. Wang, and P. Liu, “Datasetresearch: Benchmarking agent systems for demand-driven dataset discovery,” *arXiv preprint arXiv:2508.06960*, 2025.
- 5 Y. Wu, D. Fu, W. Si, Z. Huang, M. Jiang, K. Li, S. Xia, J. Sun, T. Xu, X. Hu, et al., “Innovatorbench: Evaluating agents’ ability to conduct innovative llm research,” *arXiv preprint arXiv:2510.27598*, 2025.
- 6 Y. Xiao, M. Jiang, J. Sun, K. Li, J. Lin, Y. Zhuang, J. Zeng, S. Xia, Q. Hua, X. Li, et al., “Limi: Less is more for agency,” *arXiv preprint arXiv:2509.17567*, 2025.