

Fan Zhou (周凡)

PhD student, Shanghai Jiao Tong University

E-mail: koala99.zf@gmail.com

Homepage: koalazf99.github.io

Github: koalazf99

Education

Shanghai Jiao Tong University Ph.D. in Computer Science Advisor: Prof. Pengfei Liu	2025 – Present
Shanghai Jiao Tong University M.S. in Computer Science Advisor: Prof. Hongfei Fu	2021 – 2024
Shanghai Jiao Tong University B.S. in Computer Science, IEEE Honor Class.	2017 – 2021

Research Internships

LLM360, MBZUAI Host: Hector Zhengzhong Liu	Jun. 2024 – Apr. 2025
XLing Lab, The University of Hong Kong Mentor: Prof. Tao Yu	Apr. 2023 – Dec. 2023
Microsoft Research Asia Mentor: Haoyu Dong	Oct. 2021 – Aug. 2022

Awards and Service

Shanghai Scholarship	2019
SJTU Academic Scholarship	2018 – 2021
MSRA Star of Tomorrow	2022
<i>Reviewer of ICLR, NeurIPS, COLM, ACL, IJCAI, COLING</i>	

Publications

- [1] OctoThinker: Revisiting Mid-Training In the Era of RL Scaling
Zengzhi Wang*, **Fan Zhou***, Xuefeng Li, Pengfei Liu.
Blog, 2025. | A mid-training ablation study in the era of RL scaling.
- [2] MegaMath: Pushing the Limits of Open Math Copora
Fan Zhou*, Zengzhi Wang*, Nikhil Ranjan, Zhoujun Cheng, Liping Tang, Hector Zhengzhong Liu,

Eric P. Xing.

In Submission, 2025. | **70K** dataset downloads

- [3] Sailor2: Sailing in South-East Asia with Inclusive Multilingual LLM
Longxu Dou*, Qian Liu*, **Fan Zhou***, Changyu Chen*, Zili Wang, Ziqi Jin, Zichen Liu, Tongyao Zhu, Cunxiao Du, Penghui Yang, Haonan Wang, Xin Mao, Xiachong Feng, Man Tsung Yeung, Xinyi Wan, Kunat Pipatanakul, Fajri Koto, Min Si Thu, Hynek Kydlíček, Zeyi Liu, Qunshu Lin, Sittipong Sripaisarnmongkol, Kridtapad Sae-Khow, Niratitsai Thongchim, Taechawat Konkaew, Narong Borjindagoon, Matichon Maneegard, Phakphum Artkaew, Quan Nguyen, Wannaphong Phatthiyaphibun, Zheng-Xin Yong, Anh Dao, Hoang H. Tran, Mike Zhang, Tianyu Pang, Chao Du, Wei Lu, Min Lin.
Tech Report, 2025.
- [4] Diving into Self-Evolve Training for Multimodal Reasoning
Wei Liu*, Junlong Li*, Xiwen Zhang, **Fan Zhou**, Yu Cheng, Junxian He.
ICML 2025.
- [5] Programming Every Example: Lifting Pre-training Data Quality like Experts at Scale
Fan Zhou*, Zengzhi Wang*, Qian Liu, Junlong Li, Pengfei Liu.
ICML 2025. | **15K** dataset downloads
- [6] OlympicArena: Benchmarking Multi-discipline Cognitive Reasoning for Superintelligent AI
Zhen Huang, Zengzhi Wang, Shijie Xia, ..., **Fan Zhou**, ..., Yuxiang Zheng, Shaoting Zhang, Dahua Lin, Yu Qiao, Pengfei Liu.
NeurIPS 2024.
- [7] Dissecting Human and LLM Preferences
Junlong Li, **Fan Zhou**, Shichao Sun, Yikai Zhang, Hai Zhao, Pengfei Liu.
ACL 2024.
- [8] OpenAgents: an Open Platform for Language Agent in the Wild
Tianbao Xie*, **Fan Zhou***, Zhoujun Cheng*, Peng Shi*, Luoxuan Weng*, Yitao Liu*, Toh Jing Hua, Junning Zhao, Qian Liu, Che Liu, Leo Z. Liu, Yiheng Xu, Hongjin Su, Dongchan Shin, Caiming Xiong, Tao Yu.
COLM 2024. | **6.5K** Users | **4.5k** Github Star
- [9] Lemur: Harmonizing Natural Language and Code for Language Agents
Yiheng Xu*, Hongjin Su*, Chen Xing*, Boyu Mi, Qian Liu, Weijia Shi, Binyuan Hui, **Fan Zhou**, Yitao Liu, Tianbao Xie, Zhoujun Cheng, Siheng Zhao, Lingpeng Kong, Bailin Wang, Caiming Xiong, Tao Yu.
ICLR 2024 (Spotlight, Top 5%).
- [10] From Zero to Hero: Examining the Power of Symbolic Tasks in Instruction Tuning
Qian Liu*, **Fan Zhou***, Longxu Dou, Zhengbao Jiang, Min Lin.
Tech Report, 2023.
- [11] Reflection of Thought: Inversely Eliciting Numerical Reasoning in Language Models vis Solving Linear Systems
Fan Zhou*, Haoyu Dong*, Qian Liu, Zhoujun Cheng, Shi Han, Dongmei Zhang.
Neurips 2022 MATH-AI Workshop.
- [12] Tacube: Pre-computing Data Cubes for Answering Numerical-Reasoning Questions over Tabular Data
Fan Zhou, Mengkang Hu, Haoyu Dong, Zhoujun Cheng, Fan Cheng, Shi Han, and Dongmei Zhang.
EMNLP 2022 (Oral, Top 20%) .

- [13] Table Pre-training: A Survey on Model Architectures, Pretraining Objectives, and Downstream Tasks
Haoyu Dong, Zhoujun Cheng, Xinyi He, Mengyu Zhou, Anda Zhou, **Fan Zhou**, Ao Liu, Shi Han,
Dongmei Zhang.
IJCAI 2022.

Below are Some Undergraduate Projects

- [14] Exploring Image Regions Not Well Encoded by an INN
Zenao Ling, **Fan Zhou**, Meng Wei, Quanshi Zhang.
AISTATS 2022.
- [15] Quantification and Analysis of Layer-wise and Pixel-wise Information Discarding
Haotian Ma, Hao Zhang, **Fan Zhou**, Yinqing Zhang, Quanshi Zhang.
ICML 2022.