

## Xiao Xu

Research Center for Social Computing and Information Retrieval  
Harbin Institute of Technology  
Heilongjiang Province, China

WeChat: LooperXX27  
Email: looperxx27@gmail.com  
Homepage: <https://looperxx.github.io/>

### RESEARCH INTERESTS

Vision-Language Learning, Multimodal Large Language Model.

### EDUCATION

*Ph.D. Student*, Harbin Institute of Technology, China 2020.09 - Present  
Major: Computer Science. Advisor: Prof. Wanxiang Che

*B.E.*, Northeastern University, China 2016.09 - 2020.06  
Major: Software Engineering

### PUBLICATIONS

Exploring Multi-Grained Concept Annotations: A New Data Paradigm for Multimodal LLMs.  
**Xiao Xu**, Tianhao Niu, Yuxi Xie, Libo Qin, Wanxiang Che, Min-Yen Kan.  
ICLR 2025 Submission. Under Review.

ManagerTower: Aggregating the Insights of Uni-Modal Experts for Vision-Language Representation Learning.  
**Xiao Xu**, Bei Li, Chenfei Wu, Shao-Yen Tseng, Anahita Bhiwandiwala, Shachar Rosenman, Vasudev Lal, Wanxiang Che, Nan Duan.  
Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL). Oral. 2023.

BridgeTower: Building Bridges Between Encoders in Vision-Language Representation Learning.  
**Xiao Xu**, Chenfei Wu, Shachar Rosenman, Vasudev Lal, Wanxiang Che, Nan Duan.  
Proceedings of the AAAI Conference on Artificial Intelligence (AAAI). Oral. 2023.

Text is no more Enough! A Benchmark for Profile-based Spoken Language Understanding.  
**Xiao Xu**<sup>\*</sup>, Libo Qin<sup>\*</sup>, Kaiji Chen, Guoxing Wu, Linlin Li, Wanxiang Che.  
Proceedings of the AAAI Conference on Artificial Intelligence (AAAI). Oral. 2022.

Semantic-Guided Image Augmentation with Pre-trained Models.  
Bohan Li, **Xiao Xu**, Xinghao Wang, Yutai Hou, Yunlong Feng, Feng Wang, Xuanliang Zhang, Qingfu Zhu, Wanxiang Che.  
Proceedings of the AAAI Conference on Artificial Intelligence (AAAI). 2024.

Modularized Pre-training for End-to-end Task-oriented Dialogue.  
Libo Qin, **Xiao Xu**, Lehan Wang, Yue Zhang, Wanxiang Che.  
IEEE/ACM Transactions on Audio, Speech, and Language Processing (TASLP). 2023.

AGIF: An Adaptive Graph-Interactive Framework for Joint Multiple Intent Detection and Slot Filling.  
Libo Qin, **Xiao Xu**, Wanxiang Che, Ting Liu.  
Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: Findings (EMNLP-Findings). 2020.

Dynamic Fusion Network for Multi-Domain End-to-end Task-Oriented Dialog.  
Libo Qin, **Xiao Xu**, Wanxiang Che, Yue Zhang, Ting Liu.  
Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL). 2020.

V-DPO: Mitigating Hallucination in Large Vision Language Models via Vision-Guided Direct Preference Optimization.

Yuxi Xie, Guanzhen Li, **Xiao Xu**, Min-Yen Kan.  
Findings of the Association for Computational Linguistics: EMNLP 2024. 2024.

Pro-HAN: A Heterogeneous Graph Attention Network for Profile-Based Spoken Language Understanding.

Dechuan Teng, Chunlin Lu, **Xiao Xu**, Wanxiang Che, Libo Qin.  
IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2024.

OpenSLU: A Unified, Modularized, and Extensible Toolkit for Spoken Language Understanding.

Libo Qin, Qiguang Chen, **Xiao Xu**, Yunlong Feng, Wanxiang Che.  
Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL). Demo. 2023.

A Preliminary Evaluation of ChatGPT for Zero-shot Dialogue Understanding.

Wenbo Pan, Qiguang Chen, **Xiao Xu**, Wanxiang Che, Libo Qin.  
Arxiv Preprint. 2023.

A Two-Stage Framework with Self-Supervised Distillation For Cross-Domain Text Classification.

Yunlong Feng, Bohan Li, Libo Qin, **Xiao Xu**, Wanxiang Che.  
The 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING). 2024.

GL-GIN: Fast and Accurate Non-Autoregressive Model for Joint Multiple Intent Detection and Slot Filling.

Libo Qin, Fuxuan Wei, Tianbao Xie, **Xiao Xu**, Wanxiang Che, Ting Liu.  
Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics (ACL). Oral. 2021.

CORAL: Order-Agnostic Language Modeling for Efficient Iterative Refinement.

Yuxi Xie, Anirudh Goyal, Xiaobao Wu, Xunjian Yin, **Xiao Xu**, Min-Yen Kan, Liangming Pan, William Yang Wang.  
ICLR 2025 Submission. Under Review.

In-Context Transfer Learning: Demonstration Synthesis by Transferring Similar Tasks.

Dingzirui Wang, Xuanliang Zhang, Qiguang Chen, Longxu Dou, **Xiao Xu**, Rongyu Cao, Yingwei Ma, Qingfu Zhu, Wanxiang Che, Binhua Li, Fei Huang, Yongbin Li.  
ICLR 2025 Submission. Under Review.

Self-Constructed Context Decompilation with Fined-grained Alignment Enhancement.

Yunlong Feng, Yang Xu, Dechuan Teng, Honglin Mu, **Xiao Xu**, Libo Qin, Wanxiang Che, Qingfu Zhu.  
Findings of the Association for Computational Linguistics: EMNLP 2024. 2024.

M<sup>3</sup>CoT: A Novel Benchmark for Multi-Domain Multi-step Multi-modal Chain-of-Thought.  
Qiguang Chen, Libo Qin, Jin Zhang, Zhi Chen, **Xiao Xu**, Wanxiang Che. Proceedings of the 62th Annual Meeting of the Association for Computational Linguistics (ACL). Oral. 2024.

Don't be Contradicted with Anything! CI-ToD: Towards Benchmarking Consistency for Task-oriented Dialogue System.

Libo Qin, Tianbao Xie, Shijue Huang, Qiguang Chen, **Xiao Xu**, Wanxiang Che.  
Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP). Poster. 2021.

IPGAN: Generating Informative Item Pairs by Adversarial Sampling.

Guibing Guo, Huan Zhou, Bowei Chen, Zhirong Liu, **Xiao Xu**, Xu Chen, Zhenhua Dong, Xiuqiang He.

IEEE Transactions on Neural Networks and Learning Systems (TNNLS). 2020.

<b>EXPERIENCES</b>	Visiting Student in National University of Singapore (NUS), Singapore. 2023.09 - 2024.09 Topic: Vision-Language Learning, Multimodal Large Language Model. Advisor: Prof. Min-Yen Kan.
	Research Intern in MSRA NLC Group, Beijing, China. 2022.01 - 2023.01 Topic: Vision-Language Learning. Advisor: Dr. Nan Duan, Dr. Chenfei Wu.
<b>SERVICES</b>	Conference Reviewer: ICLR 2025, NeurIPS 2024, EMNLP 2024 & 2022, ACL 2023, AAAI 2023, NLPCC 2023 & 2022. Community: MLNLP-2022 Outstanding Organizer.
<b>SKILLS</b>	Programming: Python, C/C++, Java, Shell, PyTorch, TensorFlow.
<b>AWARDS</b>	Stars of Tomorrow Internship Award of Microsoft Research Asia 2023
	National Encouragement Scholarship (for undergraduate student) (4/293) 2019
	Mathematical Contest In Modeling, Meritorious Winner (7.09%) 2019
	National Scholarship (for undergraduate student) (5/293) 2018
	National Scholarship (for undergraduate student) (10/420) 2017
	Top Ten Campus Singers of Northeastern University (Hunnan) 2016