

JIACHENG YE

✉ carsonye@connect.hku.hk 🌐 <https://jiacheng-ye.github.io>

EDUCATION

The University of Hong Kong 2022.09 – 2026.06 (Expected)

- Ph.D. in Computer Science, co-advised by [Lingpeng Kong](#) and [Tao Yu](#) at [HKUNLP Group](#).

Fudan University 2019.09 – 2021.12

- M.S. in Computer Science, advised by [Qi Zhang](#) at [FudanNLP Group](#). Ranking: 2/145

Sun Yat-sen University 2015.09 – 2019.06

- B.E. in Information System.

RESEARCH INTERESTS

I am passionate about the research on below topics:

- Diffusion Language Models ([Diffusion-of-Thoughts](#), [DiffuSearch](#), [Beyond-Autoregression](#), [Dream 7B](#), [DreamCoder](#), [DreamOn](#)).
- In-context Learning ([SAIL](#), [CEIL](#), [OpenICL](#));
- Data Generation ([ZeroGen](#), [ProGen](#), [SunGen](#), [SymGen](#), [G-LLaVA](#));

SELECTED PUBLICATIONS

Check the full list on my [Google Scholar](#).

- Zirui Wu*, Lin Zheng*, Zhihui Xie, **Jiacheng Ye**, Jiahui Gao, Yansong Feng, Zhenguo Li, Victoria W., Guorui Zhou, Lingpeng Kong. *DreamOn: Diffusion Language Models For Code Infilling Beyond Fixed-Size Canvas*. [Code]
- Zhihui Xie*, **Jiacheng Ye***, Lin Zheng*, Jiahui Gao*, Jingwei Dong, Zirui Wu, Xueliang Zhao, Shansan Gong, Xin Jiang, Zhenguo Li, Lingpeng Kong. *Dream-Coder 7B*. [Code]
- **Jiacheng Ye***, Zhihui Xie*, Lin Zheng*, Jiahui Gao*, Zirui Wu, Xin Jiang, Zhenguo Li, Lingpeng Kong. *Dream 7B: Diffusion Large Language Models*. [Code]
- **Jiacheng Ye**, Zhenyu Wu, Jiahui Gao, Xin Jiang, Zhenguo Li, Lingpeng Kong. *Implicit Search via Discrete Diffusion: A Study on Chess*. ICLR 2025. [Code] [chess-agent]
- **Jiacheng Ye**, Jiahui Gao, Shansan Gong, Lin Zheng, Xin Jiang, Zhenguo Li, Lingpeng Kong. *Beyond Autoregression: Discrete Diffusion for Complex Reasoning and Planning*. ICLR 2025. [Code]
- Shansan Gong, Shivam Agarwal, Yizhe Zhang, **Jiacheng Ye**, Lin Zheng, Mukai Li, Chenxin An, Peilin Zhao, Wei Bi, Jiawei Han, Hao Peng, Lingpeng Kong. *Scaling Diffusion Language Models via Adaptation from Autoregressive Models*. ICLR 2025.
- Jiahui Gao*, Renjie Pi*, Jipeng Zhang, **Jiacheng Ye**, Wanjun Zhong, Yufei Wang, Lanqing Hong, Jianhua Han, Hang Xu, Zhenguo Li, Lingpeng Kong. *G-LLaVA: Solving Geometric Problem with Multi-Modal Large Language Model*. ICLR 2025. [Code]
- **Jiacheng Ye***, Shansan Gong*, Liheng Chen*, Lin Zheng, Jiahui Gao, Han Shi, Chuan Wu, Xin Jiang, Zhenguo Li, Wei Bi, Lingpeng Kong. *Diffusion of Thoughts: Chain-of-Thought Reasoning in Diffusion Language Models*. NeurIPS 2024. [Code]
- **Jiacheng Ye**, Xijia Tao, Lingpeng Kong. *Language Versatilists vs. Specialists: An Empirical Revisiting on Multilingual Transfer Ability*. 2024. [Code]
- **Jiacheng Ye**, Chengzu Li, Lingpeng Kong, Tao Yu. *Generating Data for Symbolic Language with Large Language Models*. EMNLP 2023. [Code]

- Zhiyong Wu*, Yaoxiang Wang*, **Jiacheng Ye***, Lingpeng Kong. *Self-adaptive In-context Learning*. ACL 2023. [Code]
- Zhenyu Wu*, Yaoxiang Wang*, **Jiacheng Ye***, Jiangtao Feng, Jingjing Xu, Yu Qiao, Zhiyong Wu. *OpenICL: An Open-Source Framework for In-context Learning*. ACL 2023. [Code]
- **Jiacheng Ye**, Zhiyong Wu, Jiangtao Feng, Tao Yu, Lingpeng Kong. *Compositional Exemplars for In-context Learning*. ICML 2023. [Code]
- Jiahui Gao*, Renjie Pi*, Yong Lin, Hang Xu, **Jiacheng Ye**, Zhiyong Wu, Xiaodan Liang, Zhenguo Li, Lingpeng Kong. *Self-Guided High-Quality Data Generation in Efficient Zero-Shot Learning*. ICLR 2023, spotlight.
- **Jiacheng Ye**, Jiahui Gao, Zhiyong Wu, Jiangtao Feng, Tao Yu, and Lingpeng Kong. *ProGen: Progressive Zero-shot Dataset Generation via In-context Feedback*. EMNLP-Findings 2022. [Code]
- **Jiacheng Ye***, Jiahui Gao*, Qintong Li, Hang Xu, Jiangtao Feng, Zhiyong Wu, Tao Yu, Lingpeng Kong. *ZeroGen: Efficient Zero-shot Learning via Dataset Generation*. EMNLP 2022. [Code] [PaperDigest Most Influential Papers]
- **Jiacheng Ye***, Ruijian Cai*, Tao Gui, Qi Zhang. *Heterogeneous Graph Neural Networks for Keyphrase Generation*. EMNLP 2021. [Code]
- Yichao Luo*, Yige Xu*, **Jiacheng Ye**, Xipeng Qiu and Qi Zhang. *Keyphrase Generation with Fine-Grained Evaluation-Guided Reinforcement Learning*. EMNLP-Findings 2021.
- **Jiacheng Ye**, Xiang Zhou, Xiaoqing Zheng, Tao Gui, Qi Zhang. *Uncertainty-Aware Sequence Labeling*. TASLP 2021.
- **Jiacheng Ye**, Tao Gui, Yichao Luo, Yige Xu, Qi Zhang. *ONE2SET: Generating Diverse Keyphrases as a Set*. ACL 2021. [Code] [Blog]
- FudanNLP Group. *TextFlint: Unified Multilingual Robustness Evaluation Toolkit for Natural Language Processing*. ACL 2021. [Site] [Code] [Blog]
- Tao Gui*, **Jiacheng Ye***, Qi Zhang, Zhengyan Li, Zichu Fei, Yeyun Gong, Xuanjing Huang. *Uncertainty-Aware Label Refinement for Sequence Labeling*. EMNLP 2020. [Code]
- Tao Gui*, **Jiacheng Ye***, Qi Zhang, Yaqian Zhou, Yeyun Gong, Xuanjing Huang. *Leveraging Document-Level Label Consistency for Named Entity Recognition*. IJCAI 2020. [Code]

INTERNSHIPS

Huawei Noah's Ark Lab - AI Theory Lab (Hong Kong)	2024.01 - Present
• Research about Diffusion Language Model with Jiahui Gao and Zhenguo Li . Awarded HK Research Star.	
Shanghai Artificial Intelligence Laboratory	2021.11 - 2023.07
• Research about LLMs with Zhiyong Wu and Lingpeng Kong .	
Tencent	2021.06 - 2021.11
• Worked on the ranking algorithm in the recommender system.	
NetEase	2018.08 - 2018.12
• Worked on data engineering with Hadoop, Spark, Hive, etc.	

HONORS & AWARDS

- Outstanding graduate of Shanghai (1%), Fudan University, 2022.
- National Scholarship (1%), Ministry of Education of China, 2021.
- Glarun Scholarship of CETC-NRIET (5%), Fudan University, 2020.
- The Second-grade Scholarship (10%), Sun Yat-sen University, 2016/2017/2018.