Tao Yu

CONTACT INFORMATION	https://taoyds.github.io/	tyu@cs.hku.hk
	Room 204, Chow Yei Ching Building, Pokfulam, Hong Kong	+852-5392-4676

RESEARCH INTERESTS

My research aims to develop **embodied AI agents** that ground language and perception into code and actions executable in digital and physical environments, helping people perform data science, control computers, and collaborate with robots to carry out real-world tasks.

ACADEMIC POSITION

University of Hong Kong, Hong Kong   08/2021 - *Present*

Assistant Professor in Computer Science
Director, XLANG Lab, co-director, HKU NLP Group

University of Washington, Seattle, WA   09/2021 - 08/2022

Postdoctoral Researcher, UW NLP Group

Working with Noah Smith, Mari Ostendorf and Luke Zettlemoyer

EDUCATION

Yale University, New Haven, CT   08/2017 - 05/2021

Ph.D. Computer Science

Advisors: The late Dragomir Radev (d. 2023)

Columbia University, New York, NY   09/2015 - 05/2017

M.S. Data Science

Advisors: Owen Rambow and Kathleen McKeown

University of Utah, Salt Lake City, UT   01/2012 - 05/2015

B.S. Mathematics and Economics

HONORS & AWARDS

Google Research Scholar Award	2023
Amazon Research Award	2022
Yale PhD Fellowship	2017
Connie and Robert White Honor Roll Scholarship	2014
Mathematics Department Scholarship	2014
International Student and Scholar Services Scholarship	2014
NSF Research Experience for Undergraduates (REU) Fellowship	2014

INVITED TALKS

Keynote at LLM Agents Workshop at ACL 2025	2025
Keynote at Table Representation Learning Workshop at ACL 2025	2025
Stanford NLP Seminar	2025
Keynote at Open-World Agents Workshop at NeurIPS 2024	2024
Tutorial on Language Agents: Foundations, Prospects, and Risks at EMNLP 2024	2024
Panelist, Workshop on LLM Agents at ICLR 2024	2024
Instacart Distinguished Speaker Series	2024
Keynote at Table Representation Learning Workshop at NeurIPS 2023	2023
ML4Code Seminar at MILA	2023
Morgan Stanley ML Speaker Seminar	2023
Keynote at Databases and Large Language Models Workshop at VLDB 2023	2023
Tutorial on Complex Reasoning over Natural Language at ACL 2023	2023

Apple KP Tech Talks	2023
Columbia NLP Seminar	2023
Cornell DB Seminar	2023
Microsoft Research Asia	2023
Amazon AWS	2022
AllState Tech Talks	2022
ServiceNow Research	2022
Google Research	2021

PUBLICATIONS

1. Yiheng Xu, Dunjie Lu, Zhennan Shen, Junli Wang, Zekun Wang, Yuchen Mao, Caiming Xiong, **Tao Yu**
   AgentTrek: Agent Trajectory Synthesis via Guiding Replay with Web Tutorials
   Under review

2. Hongjin Su, Ruoxi Sun, Jinsung Yoon, Pengcheng Yin, **Tao Yu**, Sercan O Arik
   Learn-by-interact: A Data-Centric Framework For Self-Adaptive Agents in Realistic Environments
   Under review

3. Yiheng Xu, Zekun Wang, Junli Wang, Dunjie Lu, Tianbao Xie, Amrita Saha, Doyen Sahoo, **Tao Yu**, Caiming Xiong
   Aguvis: Unified Pure Vision Agents for Autonomous GUI Interaction
   Under review

4. Fangyu Lei, Jixuan Chen, Yuxiao Ye, Ruisheng Cao, Dongchan Shin, Hongjin Su, Zhaoqing Suo, Hongcheng Gao, Wenjing Hu, Pengcheng Yin, Victor Zhong, Caiming Xiong, Ruoxi Sun, Qian Liu, Sida Wang, **Tao Yu**
   Spider 2.0: Evaluating Language Models on Real-World Enterprise Text-to-SQL Workflows
   Preprint 2024

5. Yanzhe Zhang, **Tao Yu**, Diyi Yang
   Attacking Vision-Language Computer Agents via Pop-ups
   Preprint 2024

6. Hongjin Su, Howard Yen, Mengzhou Xia, Weijia Shi, Niklas Muennighoff, Han-yu Wang, Haisu Liu, Quan Shi, Zachary S. Siegel, Michael Tang, Ruoxi Sun, Jinsung Yoon, Sercan O. Arik, Danqi Chen, **Tao Yu**
   BRIGHT: A Realistic and Challenging Benchmark for Reasoning-Intensive Retrieval
   Preprint 2024

7. Niklas Muennighoff, Hongjin Su, Liang Wang, Nan Yang, Furu Wei, **Tao Yu**, Amanpreet Singh, Douwe Kiela
   Generative Representational Instruction Tuning
   ICLR 2024 AGI Workshop, *Best Paper Award*

8. Tianbao Xie, Danyang Zhang, Jixuan Chen, Xiaochuan Li, Siheng Zhao, Ruisheng Cao, Toh Jing Hua, Zhoujun Cheng, Dongchan Shin, Fangyu Lei, Yitao Liu, Yiheng Xu, Shuyan Zhou, Silvio Savarese, Caiming Xiong, Victor Zhong, **Tao Yu**
   OSWorld: Benchmarking Multimodal Agents for Open-Ended Tasks in Real Computer Environments
   NeurIPS 2024, *∼1.5k GitHub stars*

9. Ruisheng Cao, Fangyu Lei, Haoyuan Wu, Jixuan Chen, Yeqiao Fu, Hongcheng Gao, Xinzhuang Xiong, Hanchong Zhang, Yuchen Mao, Wenjing Hu, Tianbao Xie, Hongshen Xu, Danyang Zhang, Sida Wang, Ruoxi Sun, Pengcheng Yin, Caiming Xiong, Ansong Ni, Qian Liu, Victor Zhong, Lu Chen, Kai Yu, **Tao Yu**
   Spider2-V: How Far Are Multimodal Agents From Automating Data Science and Engineering Workflows?
   NeurIPS 2024, *Spotlight*

10. Hongjin Su, Shuyang Jiang, Yuhang Lai, Haoyuan Wu, Boao Shi, Che Liu, Qian Liu, **Tao Yu**
EvoR: Evolving Retrieval for Code Generation
EMNLP Findings 2024

11. 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**
OpenAgents: An Open Platform for Language Agents in the Wild
COLM 2024, *~4k GitHub stars*

12. Bo-Ru Lu, Nikita Haduong, Chia-Hsuan Lee, Zeqiu Wu, Hao Cheng, Paul Koester, Jean Utke, **Tao Yu**, Noah A. Smith, Mari Ostendorf
Does Collaborative Human-LM Dialogue Generation Help Information Extraction from Human Dialogues?
COLM 2024

13. 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**
Lemur: Harmonizing Natural Language and Code for Language Agents
ICLR 2023, *Spotlight*

14. Tianbao Xie*, Siheng Zhao*, Chen Henry Wu, Yitao Liu, Qian Luo, Victor Zhong, Yanchao Yang, **Tao Yu**
Text2Reward: Reward Shaping with Language Models for Reinforcement Learning
ICLR 2023, *Spotlight*

15. Hongjin Su*, Weijia Shi*, Jungo Kasai, Yizhong Wang, Yushi Hu, Mari Ostendorf, Wen-tau Yih, Noah A Smith, Luke Zettlemoyer,**Tao Yu**
One Embedder, Any Task: Instruction-Finetuned Text Embeddings
ACL Findings 2023, *~4.5M downloads on HuggingFace, ~2k GitHub stars*

16. Yuhang Lai*, Chengxi Li*, Yiming Wang*, Tianyi Zhang*, Ruiqi Zhong*, Luke Zettlemoyer, Scott Wen-tau Yih, Daniel Fried, Sida Wang, **Tao Yu**
DS-1000: A Natural and Reliable Benchmark for Data Science Code Generation
ICML 2023

17. Tianyi Zhang, **Tao Yu**, Tatsunori B Hashimoto, Mike Lewis, Wen-tau Yih, Daniel Fried, Sida I Wang
Coder Reviewer Reranking for Code Generation
ICML 2023

18. Jiacheng Ye, Zhiyong Wu, Jiangtao Feng, **Tao Yu**, Lingpeng Kong
Compositional Exemplars for In-context Learning
ICML 2023

19. Jiacheng Ye, Chengzu Li, **Tao Yu**, Lingpeng Kong
Generating Data for Symbolic Language with Large Language Models
EMNLP 2023

20. Zhoujun Cheng, Jungo Kasai, **Tao Yu**
Batch Prompting: Efficient Inference with Large Language Model APIs
EMNLP 2023 Industry Track

21. Lianghao Xia, Chao Huang, Chunzhen Huang, Kangyi Lin, **Tao Yu**, Ben Kao
Automated Self-Supervised Learning for Recommendation
WWW 2023

22. Zhoujun Cheng*, Tianbao Xie*, Peng Shi, Chengzu Li, Rahul Nadkarni, Yushi Hu, Caiming Xiong, Dragomir Radev, Mari Ostendorf, Luke Zettlemoyer, Noah A Smith, **Tao Yu**
Binding Language Models in Symbolic Languages
ICLR 2023, *Spotlight*

23. Hongjin Su, Jungo Kasai, Chen Henry Wu, Weijia Shi, Tianlu Wang, Jiayi Xin, Rui Zhang, Mari Ostendorf, Luke Zettlemoyer, Noah A. Smith, **Tao Yu**
Selective Annotation Makes Language Models Better Few-Shot Learners
ICLR 2023

24. Tianbao Xie*, Chen Henry Wu*, Peng Shi, Ruiqi Zhong, Torsten Scholak, Michihiro Yasunaga, Chien-Sheng Wu, Ming Zhong, Pengcheng Yin, Sida I. Wang, Victor Zhong, Bailin Wang, Chengzu Li, Connor Boyle, Ansong Ni, Ziyu Yao, Dragomir Radev, Caiming Xiong, Lingpeng Kong, Rui Zhang, Noah A. Smith, Luke Zettlemoyer, **Tao Yu**
UnifiedSKG: Unifying and Multi-Tasking Structured Knowledge Grounding with Text-to-Text Language Models
EMNLP 2022, *Oral*

25. 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

26. Yushi Hu, Chia-Hsuan Lee, Tianbao Xie, **Tao Yu**, Noah A. Smith, Mari Ostendorf
In-Context Learning for Few-Shot Dialogue State Tracking
EMNLP Findings 2022

27. Jiacheng Ye, Jiahui Gao, Zhiyong Wu, Jiangtao Feng, **Tao Yu**, Lingpeng Kong
ProGen: Progressive Zero-shot Dataset Generation via In-context Feedback
EMNLP Findings 2022

28. Qi Liu, Zihuiwen Ye, **Tao Yu**, Phil Blunsom, Linfeng Song
Augmenting Multi-Turn Text-to-SQL Datasets with Self-Play
EMNLP Findings 2022

29. Yiru Chen, Ryan Li, Austin Mac, Tianbao Xie, **Tao Yu**, Eugene Wu
NL2INTERFACE: Interactive Visualization Interface Generation from Natural Language Queries
IEEE Visualization Conference NLVIZ Workshop, 2022

30. Ziming Mao*, Chen Henry Wu*, Ansong Ni, Yusen Zhang, Rui Zhang, **Tao Yu**, Budhaditya Deb, Chenguang Zhu, Ahmed H Awadallah, Dragomir Radev
DYLE: Dynamic Latent Extraction for Abstractive Long-Input Summarization
ACL 2022

31. Yusen Zhang, Ansong Ni, **Tao Yu**, Rui Zhang, Chenguang Zhu, Budhaditya Deb, Asli Celikyilmaz, Ahmed Hassan Awadallah, Dragomir Radev
An Exploratory Study on Long Dialogue Summarization: What Works and What's Next
EMNLP Findings 2021

32. Ansong Ni, Zhangir Azerbayev, Mutethia Mutuma, Troy Feng, Yusen Zhang, **Tao Yu**, Ahmed Hassan Awadallah, Dragomir Radev
SummerTime: Text Summarization Toolkit for Non-experts
EMNLP 2021 Demo Track

33. Heather Lent, Semih Yavuz, **Tao Yu**, Tong Niu, Yingbo Zhou, Dragomir Radev, Xi Victoria Lin
Testing Cross-Database Semantic Parsers Using Canonical Utterances
Evaluation of NLP Systems Workshop at EMNLP 2021, *Best Paper Award*

34. Chang Shu, Yusen Zhang, Xiangyu Dong, Peng Shi, **Tao Yu**, Rui Zhang
Logic-Consistency Text Generation from Semantic Parses
ACL Findings 2021

35. **Tao Yu**, Rui Zhang, Oleksandr Polozov, Chris Meek, Ahmed Hassan Awadallah
SCoRe: Pre-Training for Context Representation in Conversational Semantic Parsing
ICLR 2021

36. **Tao Yu**, Chien-Sheng Wu, Xi Victoria Lin, Bailin Wang, Yi Chern Tan, Xinyi Yang, Dragomir Radev, Richard Socher, Caiming Xiong
    GraPPa: Grammar-Augmented Pre-Training for Table Semantic Parsing
    ICLR 2021

37. Ming Zhong, Da Yin, **Tao Yu**, Ahmad Zaidi, Mutethia Mutuma, Rahul Jha, Ahmed Hassan Awadallah, Asli Celikyilmaz, Yang Liu, Xipeng Qiu, Dragomir Radev
    QMSum: A New Benchmark for Query-based Multi-domain Meeting Summarization
    NAACL 2021

38. Linyong Nan, Dragomir Radev, Rui Zhang, Amrit Rau, Abhinand Sivaprasad, Chiachun Hsieh, Nazneen Fatema Rajani, Xiangru Tang, Aadit Vyas, Neha Verma, Pranav Krishna, Yangxiaokang Liu, Nadia Irwanto, Jessica Pan, Faiaz Rahman, Ahmad Zaidi, Murori Mutuma, Yasin Tarabar, Ankit Gupta, **Tao Yu**, Yi Chern Tan, Xi Victoria Lin, Caiming Xiong, Richard Socher, Nazneen Fatema Rajani
    DART: Open-Domain Structured Data Record to Text Generation
    NAACL 2021

39. Ruiqi Zhong, **Tao Yu**, Dan Klein
    Semantic Evaluation for Text-to-SQL with Distilled Test Suites
    EMNLP 2020

40. **Tao Yu**, Rui Zhang, He Yang Er, Suyi Li, Eric Xue, Bo Pang, Xi Victoria Lin, Yi Chern Tan, Tianze Shi, Zihan Li, Youxuan Jiang, Michihiro Yasunaga, Sungrok Shim, Tao Chen, Alexander Fabbri, Zifan Li, Luyao Chen, Yuwen Zhang, Shreya Dixit, Vincent Zhang, Caiming Xiong, Richard Socher, Walter Lasecki, Dragomir Radev
    CoSQL: A Conversational Text-to-SQL Challenge Towards Cross-Domain Natural Language Interfaces to Databases
    EMNLP 2019

41. Rui Zhang, **Tao Yu**, He Yang Er, Sungrok Shim, Eric Xue, Xi Victoria Lin, Tianze Shi, Caiming Xiong, Richard Socher, Dragomir Radev
    Editing-based SQL Query Generation for Cross-Domain Context-Dependent Questions
    EMNLP 2019

42. **Tao Yu**, Rui Zhang, Michihiro Yasunaga, Yi Chern Tan, Xi Victoria Lin, Suyi Li, Heyang Er, Irene Li, Bo Pang, Tao Chen, Emily Ji, Shreya Dixit, David Proctor, Sungrok Shim, Jonathan Kraft, Vincent Zhang, Caiming Xiong, Richard Socher, Dragomir Radev
    SParC: Cross-Domain Semantic Parsing in Context
    ACL 2019

43. Richard A Plunz, Yijia Zhou, Maria Isabel Carrasco Vintimilla, Kathleen Mckeown, **Tao Yu**, Laura Uguccioni, Maria Paola Sutto
    Twitter Sentiment in New York City Parks as Measure of Well-being
    Landscape and Urban Planning 2019

44. **Tao Yu**, Rui Zhang, Kai Yang, Michihiro Yasunaga, Dongxu Wang, Zifan Li, James Ma, Irene Li, Qingning Yao, Shanelle Roman, Zilin Zhang, Dragomir Radev
    Spider: A Large-Scale Human-Labeled Dataset for Complex and Cross-Domain Semantic Parsing and Text-to-SQL Task
    EMNLP 2018, *∼300 submissions, ∼1k GitHub stars*

45. **Tao Yu**, Michihiro Yasunaga, Kai Yang, Rui Zhang, Dongxu Wang, Zifan Li, Dragomir Radev
    SyntaxSQLNet: Syntax Tree Networks for Complex and Cross Domain Text-to-SQL Task
    EMNLP 2018

46. **Tao Yu**, Zifan Li, Zilin Zhang, Rui Zhang, Dragomir Radev
    TypeSQL: Knowledge-based Type-Aware Neural Text-to-SQL Generation
    NAACL 2018

47. Mohammad Sadegh Rasooli, Noura Farra, Axinia Radeva, **Tao Yu**, and Kathleen McKeown
Cross-Lingual Sentiment Transfer with Limited Resources
Machine Translation 2017

48. Owen Rambow, **Tao Yu**, Axinia Radeva, Sardar Hamidian, Alexander R. Fabbri, Debanjan Ghosh, Christopher Hidey, Tianrui Peng, Mona Diab, Kathleen McKeown, Smaranda Muresan
The Columbia-GWU System at the 2016 TAC KBP BeSt Evaluation
NIST TAC KBP Workshop 2016

PROFESSIONAL SERVICES	**Organizing Committee Member**

Organizing Committee Member

Demonstration Chair at **ACL 2025**

Multi-Agent Workshop at **AAAI 2025**

Student Volunteer Chair at **ACL 2023**

SUKI: Workshop on Structured and Unstructured Knowledge Integration at **NAACL 2022**

The Workshop on Interactive and Executable Semantic Parsing at **EMNLP 2020**

Area Chair
ICLR 2024
COLM 2024
EMNLP 2023, 2024
ACL 2023, 2024

Program Committee Member/Referee
Nature, Nature Communications
TACL, ARR, ACL, EMNLP, NAACL
ICLR, ICML, NeurlPS

STUDENTS

Xinyuan Wang (HKU PhD, 2024-)

Bowen Wang (HKU PhD, 2024-)

Tianbao Xie (HKU PhD, 2022-)

Hongjin Su (HKU PhD, 2022-)

Yiheng Xu (HKU PhD, 2022-, co-advised with Lingpeng Kong)

Jiacheng Ye (HKU PhD, 2022-, co-advised with Lingpeng Kong)

Siheng Zhao (Intern 2023, NJU BS → USC PhD)

Yuhan Liu (Intern 2023, XJTU BS BS → NYU PhD)

Zhoujun Cheng (Intern 2022-23, SJTU BS/MS → UCSD PhD)

Fan Zhou (Intern 2023, SJTU BS/MS)

Leo Zeyu Liu (Intern 2023, UW BS/MS → UT Austin PhD)

Chen Wu (Intern 2022, THU BS → CMU PhD)

Ryan Li (Intern 2022, UW BS → Stanford MS)

Chengzu Li (Intern 2022, XJTU BS → Cambridge PhD)

Yiming Wang (Intern 2022, PKU BS → Harvard MS)

Ming Zhong (Intern 2020, Fudan MS → UIUC PhD)

Da Yin (Intern 2020, PKU BS → UCLA PhD)

TEACHING	*Instructor*: Advanced NLP (DATA 8005), University of Hong Kong	2024
	Instructor: Natural Language Processing (COMP 3361), University of Hong Kong	2024
	Instructor: Advanced NLP (DATA 8005), University of Hong Kong	2023
	Teaching Assistant: Natural Language Processing (CPSC 577), Yale University	2019
	Teaching Assistant: Database Systems (CPSC 537), Yale University	2019
	Teaching Assistant: Introduction to NLP (CS 4705), Columbia University	2017
	Teaching Assistant: Natural Language Processing (Coursera)	2016

GRANTS	Areas of Excellence, Research Grants Council of Hong Kong - Co-PI, my share: HKD$3.8M	2024
	Early Career Scheme, Research Grants Council of Hong Kong - PI, HKD$1,078,487	2023
	Google Research Scholar Award - PI, USD$35,000	2023
	Amazon Research Award - PI, USD$40,000 + $80,000 AWS Credits	2022

EXPERIENCE	**Microsoft Research**, Redmond, WA	06/2020 - 08/2020
	NLP Research Intern Mentors: Ahmed Hassan Awadallah, Oleksandr Polozov, and Chris Meek	
	Salesforce Research, Palo Alto, CA	06/2019 - 08/2019
	NLP Research Intern Mentors: Victoria Lin, Caiming Xiong, and Richard Socher	
	Samsung Research America, Mountain View, CA	06/2018 - 08/2018
	NLP Research Intern	
	Columbia CCLS & NLP Group, New York, NY	05/2016 - 10/2016
	Graduate Research Assistant Mentors: Owen Rambow and Kathleen McKeown	