

Tao Yu

CONTACT INFORMATION	https://taoyds.github.io/ Room CB-318 Chow Yei Ching Building, Pokfulam, Hong Kong	tao.yu@yale.edu +1-801-493-9166
RESEARCH INTERESTS	Natural Language Processing and Deep Learning. The goal of my research is to design and build conversational natural language interfaces that can help humans explore and reason over data in any application (e.g., relational databases and mobile apps) in a robust and trusted manner. It involves studies in semantic parsing, dialogue systems, pretraining, human computer interaction, and beyond text (speech, vision, robotics, instruction following, etc.).	
ACADEMIC EMPLOYMENT	The University of Hong Kong Assistant Professor in Computer Science. Co-director, HKU NLP Lab 08/2021 - Present	
	University of Washington , Seattle, WA Postdoc in UW NLP Group, Computer Science & Engineering est. 09/2021 - 07/2022 <ul style="list-style-type: none">Working with Noah Smith and Mari Ostendorf	
EDUCATION	Yale University , New Haven, CT Ph.D. Computer Science - Natural Language Processing 08/2017 - 05/2021 <ul style="list-style-type: none">Advisors: Dragomir R. RadevCommittee: Kathleen McKeown, Luke Zettlemoyer, Robert Frank, Marynel Vázquez Columbia University , New York, NY M.S. Data Science GPA: 3.82/4.0 09/2015 - 05/2017 <ul style="list-style-type: none">Advisors: Owen Rambow and Kathleen McKeown University of Utah , Salt Lake City, UT B.S. Mathematics and Economics GPA: 3.86/4.0 01/2012 - 05/2015 <ul style="list-style-type: none">Advisors: Jingyi Zhu and Peter Philips	
RESEARCH INTERNSHIPS	Microsoft Research , Redmond, WA NLP Research Intern 06/2020 - 08/2020 <ul style="list-style-type: none">Context representation in conversational semantic parsing via pre-training Mentors: Ahmed Hassan Awadallah , Oleksandr Polozov , and Chris Meek	
	Salesforce Research , Palo Alto, CA NLP Research Intern 06/2019 - 08/2019 <ul style="list-style-type: none">Language model pre-training for table semantic parsing Mentors: Victoria Lin , Caiming Xiong , and Richard Socher	
	Samsung Research America , Mountain View, CA NLP Research Intern 06/2018 - 08/2018	
	Columbia CCLS & NLP Group , New York, NY Graduate Research Assistant 05/2016 - 10/2016 <ul style="list-style-type: none">Developed systems for Text Analysis Conference (TAC) 2016 Belief and Sentiment Evaluation, and achieved the best performances on all 12 English and Spanish tasks. Mentors: Owen Rambow and Kathleen McKeown	

1. Logic-Consistency Text Generation from Semantic Parses
Chang Shu, Yusen Zhang, Xiangyu Dong, Peng Shi, **Tao Yu**, Rui Zhang
Findings of ACL 2021
2. SCoRe: Pre-Training for Context Representation in Conversational Semantic Parsing
Tao Yu, Rui Zhang, Oleksandr Polozov, Chris Meek, Ahmed Hassan Awadallah
ICLR 2021
3. GraPPa: Grammar-Augmented Pre-Training for Table Semantic Parsing
Tao Yu, Chien-Sheng Wu, Xi Victoria Lin, Bailin Wang, Yi Chern Tan, Xinyi Yang, Dragomir Radev, Richard Socher, Caiming Xiong.
ICLR 2021
4. QMSum: A New Benchmark for Query-based Multi-domain Meeting Summarization
Ming Zhong, Da Yin, **Tao Yu**, Ahmad Zaidi, Mutethia Mutuma, Rahul Jha, Ahmed Hassan Awadallah, Asli Celikyilmaz, Yang Liu, Xipeng Qiu, Dragomir Radev
NAACL 2021
5. DART: Open-Domain Structured Data Record to Text Generation
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
NAACL 2021
6. Semantic Evaluation for Text-to-SQL with Distilled Test Suites
Ruiqi Zhong, **Tao Yu**, Dan Klein
EMNLP 2020
7. Did You Ask a Good Question? A Cross-Domain Question Intention Classification Benchmark for Text-to-SQL
Yusen Zhang, Xiangyu Dong, Shuaichen Chang, **Tao Yu**, Peng Shi, Rui Zhang
EMNLP 2020 Workshop on Interactive and Executable Semantic Parsing
8. CoSQL: A Conversational Text-to-SQL Challenge Towards Cross-Domain Natural Language Interfaces to Databases
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
EMNLP 2019
9. Editing-based SQL Query Generation for Cross-Domain Context-Dependent Questions
Rui Zhang, **Tao Yu**, He Yang Er, Sungrok Shim, Eric Xue, Xi Victoria Lin, Tianze Shi, Caiming Xiong, Richard Socher, Dragomir Radev
EMNLP 2019
10. SPaC: Cross-Domain Semantic Parsing in Context
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
ACL 2019
11. Twitter Sentiment in New York City Parks as Measure of Well-being
Richard A Plunz, Yijia Zhou, Maria Isabel Carrasco Vintimilla, Kathleen Mckeown, **Tao Yu**, Laura Ugucioni, Maria Paola Sutto
Landscape and Urban Planning 2019

12. Spider: A Large-Scale Human-Labeled Dataset for Complex and Cross-Domain Semantic Parsing and Text-to-SQL Task
Tao Yu, Rui Zhang, Kai Yang, Michihiro Yasunaga, Dongxu Wang, Zifan Li, James Ma, Irene Li, Qingning Yao, Shanelle Roman, Zilin Zhang, Dragomir Radev
EMNLP 2018
13. SyntaxSQLNet: Syntax Tree Networks for Complex and Cross Domain Text-to-SQL Task
Tao Yu, Michihiro Yasunaga, Kai Yang, Rui Zhang, Dongxu Wang, Zifan Li, Dragomir Radev
EMNLP 2018
14. TypeSQL: Knowledge-based Type-Aware Neural Text-to-SQL Generation
Tao Yu, Zifan Li, Zilin Zhang, Rui Zhang, Dragomir Radev
NAACL 2018
15. Cross-Lingual Sentiment Transfer with Limited Resources
 Mohammad Sadegh Rasooli, Noura Farra, Axinia Radeva, Tao Yu, and Kathleen McKeown
Machine Translation 2017
16. The Columbia-GWU System at the 2016 TAC KBP BeSt Evaluation
 Owen Rambow, Tao Yu, Axinia Radeva, Sardar Hamidian, Alexander R. Fabbri, Debanjan Ghosh, Christopher Hidey, Tianrui Peng, Mona Diab, Kathleen McKeown, Smaranda Muresan
NIST TAC KBP Workshop 2016

PROFESSIONAL
SERVICES

Organizing Committee Member

The Workshop on Interactive and Executable Semantic Parsing at EMNLP 2020

Program Committee Member

ACL Rolling Review

ACL: 2020, 2021

EMNLP: 2019, 2020, 2021

NAACL: 2019, 2021

ICLR: 2022

COLING: 2020

AAACL-IJCNLP: 2020

Workshop on Natural Language Processing for Programming (**NLP4Prog**) at ACL 2021

AWARDS

Yale Nathan Hale Associates Scholarship Award	2020
Yale Conference Travel Fellowship	2018
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

TEACHING
EXPERIENCE

Natural Language Processing (CPSC 577), Yale University	2019
Database Systems (CPSC 537), Yale University	2019
Introduction to Natural Language Processing (CS 4705), Columbia University	2017
Natural Language Processing (Coursera)	2016

TALKS & PRESENTATIONS	Invited Talk, NLP Reading Group, Google Research	Apr 2021
	Invited Talk, CS, The University of Hong Kong	Mar 2021
	Invited Talk, CS, The National University of Singapore	Mar 2021
	Invited Talk, CS, The University of Wisconsin-Madison	Mar 2021
	Invited Talk, CS, Simon Fraser University	Mar 2021
	Invited Talk, CSE, The University of Minnesota Twin Cities	Mar 2021
	Invited Talk, CS, City University of Hong Kong	Feb 2021
	Invited Talk, CS, Fudan University	Feb 2021
	Invited Talk, CS, Southern University of Science and Technology	Jan 2021
	Invited Poster, Microsoft Research AI Breakthroughs Workshop, Redmond	Sep. 2019

STUDENTS	Ming Zhong (Summer Intern 2020, Fudan MS → UIUC PhD)
	Da Yin (Summer Intern 2020, PKU BS → UCLA PhD)
	Yusen Zhang (Summer Intern 2020, Emory MS → PSU PhD)
	Michihiro Yasunaga (Project Student, 2018-19, Yale BS → Stanford PhD)
	Suyi Li (Project Student, 2019, Yale MS → HKUST PhD)
	Kai Yang (Project Student, 2018, Yale MS → Google SE)
	James Ma (Project Student, 2018, Yale BS → LinkedIn DS)

REFERENCES	Dragomir Radev Professor, Yale University Email: dragomir.radev@yale.edu Cell: 646-420-1367
	Kathleen McKeown Professor, Columbia University Email: kathy@cs.columbia.edu Cell: 212-939-7118
	Luke Zettlemoyer Professor, University of Washington Email: lsz@cs.washington.edu Cell: 206-685-1227
	Owen Rambow Professor, Stony Brook University Email: owen.rambow@gmail.com Cell: 917-679-5265
	Caiming Xiong Managing Director, Salesforce Research Email: cxiong@salesforce.com Cell: 716-435-9316
	Ahmed Hassan Awadallah Principal Research Manager, Microsoft Research AI Email: hassanam@microsoft.com