

# Tao Yu

---

CONTACT INFORMATION	<a href="https://taoyds.github.io/">https://taoyds.github.io/</a> 420 Temple Street, New Haven, CT 06511	tao.yu@yale.edu 801-493-9166
RESEARCH INTERESTS	Natural language processing, deep learning, and machine learning. More specifically, I am currently interested in building conversational natural language interfaces for information access.	
EDUCATION	<b>Yale University</b> , New Haven, CT Ph.D. Computer Science - Natural Language Processing 08/2017 - Present <ul style="list-style-type: none"><li>Advisors: Dragomir R. Radev</li></ul> <b>Columbia University</b> , SEAS, New York, NY M.S. Data Science GPA: 3.82/4.0 09/2015 - 05/2017 <ul style="list-style-type: none"><li>Advisors: Owen Rambow and Kathleen McKeown</li></ul> <b>University of Utah</b> , Salt Lake City, UT B.S. Mathematics and Economics GPA: 3.86/4.0 01/2012 - 05/2015 <ul style="list-style-type: none"><li>Advisors: Jingyi Zhu and Peter Philips</li></ul>	
RESEARCH & INDUSTRY EXPERIENCE	<b>Yale University</b> , New York, NY Graduate Student Research – NLP 08/2017 - Present <ul style="list-style-type: none"><li>Working on dialog-based robust and reliable text-to-SQL systems.</li><li>Open Text-to-SQL Challenge Series: Spider, SPaC, CoSQL.</li></ul> <b>Salesforce Research</b> , Palo Alto, CA Research Intern mentored by Victoria Lin and Caiming Xiong – NLP 06/2019 - 08/2019 <b>Samsung Research America</b> , Mountain View, CA Research Intern – NLP 06/2018 - 08/2018 <b>Columbia CCLS &amp; NLP Group</b> , New York, NY Graduate Research Assistant 05/2016 - 10/2016 <ul style="list-style-type: none"><li>Developed systems for DARPA Text Analysis Conference (TAC) 2016 Belief and Sentiment Evaluation, and achieved the best performances on all 12 English and Spanish tasks.</li></ul> <b>AIG Science</b> , New York, NY App Development Intern 01/2016 - 03/2016 <ul style="list-style-type: none"><li>Developed a Hybrid Mobile App using React Native and built backend using Flask.</li></ul>	
PUBLICATIONS & PAPERS IN PREPARATION	1. <b>T. Yu</b> , R. Zhang, H. Er, S. Li, E. Xue, B. Pang, V. Lin, Y. Tan, T. Shi, Z. Li, Y. Jiang, M. Yasunaga, S. Shim, T. Chen, A. Fabbri, Z. Li, L. Chen, Y. Zhang, S. Dixit, V. Zhang, C. Xiong, R. Socher, W. Lasecki, D. Radev. “CoSQL: A Conversational Text-to-SQL Challenge Towards Cross-Domain Natural Language Interfaces to Databases”. <i>In the 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP)</i> , 2019. <i>Long Paper</i>	

2. R. Zhang, **T. Yu**, H. Er, S. Shim, E. Xue, V. Lin, T. Shi, C. Xiong, R. Socher, W. Lasecki, D. Radev. "Editing-based SQL Query Generation for Cross-Domain Context-Dependent Questions". *In the 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2019. *Long Paper*
3. **T. Yu**, R. Zhang, M. Yasunaga, Y. Tan, V. Lin, S. Li, H. Er, I. Li, B. Pang, T. Chen, E. Ji, S. Dixit, D. Proctor, S. Shim, J. Kraft, V. Zhang, C. Xiong, R. Socher, D. Radev. "SParC: Cross-Domain Semantic Parsing in Context". *In the 57th Annual Meeting of the Association for Computational Linguistics (ACL)*, 2019. *Long Paper*
4. R. Plunz, Y. Zhou, M. Vintimilla, K. Mckeown, **T. Yu**, L. Uguccioni, M. Sutto. "Twitter Sentiment in New York City Parks as Measure of Well-being". *Landscape and Urban Planning*, 2019. *Long Paper*
5. **T. Yu**, R. Zhang, K. Yang, M. Yasunaga, D. Wang, Z. Li, J. Ma, I. Li, Q. Yao, S. Roman, Z. Zhang, D. Radev. "Spider: A Large-Scale Human-Labeled Dataset for Complex and Cross-Domain Semantic Parsing and Text-to-SQL Task". *In the 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2018. *Long Paper*
6. **T. Yu**, M. Yasunaga, K. Yang, R. Zhang, D. Wang, Z. Li, D. Radev. "SyntaxSQLNet: Syntax Tree Networks for Complex and Cross Domain Text-to-SQL Task". *In the 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2018. *Long Paper*
7. **T. Yu**, Z. Li, Z. Zhang, R. Zhang, D. Radev. "TypeSQL: Knowledge-based Type-Aware Neural Text-to-SQL Generation". *In the 16th Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT)*, 2018. *Short Paper*
8. M. S. Rasooli, N. Farra, A. Radeva, **T. Yu** and K. McKeown. "Cross-Lingual Sentiment Transfer with Limited Resources". *Machine Translation*, 2017. *Long Paper*
9. O. Rambow, **T. Yu**, A. Radeva, S. Hamidian, A. Fabbri, C. Hidey, K. McKeown et. al. "The Columbia-GWU System at the 2016 TAC KBP BeSt Evaluation". *NIST TAC KBP Workshop*, 2016. *Long Paper*
10. **T. Yu**, C. Hidey, O. Rambow, K. McKeown. "Leveraging Sparse and Dense Feature Combinations for Sentiment Classification". *arXiv:1708.03940 [cs.CL]*, 2017. *Short Paper*

PROFESSIONAL SERVICES	Reviewer/PC: NAACL 2019, EMNLP 2019	
AWARDS	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
PROGRAMMING & OTHER SKILLS	Proficient in Python, PyTorch, TensorFlow, R, Java, Matlab, SQL. Familiar with C, C++, JavaScript, Theano, Hadoop, Spark. FAA Student Pilot.	