

Tao Yu

CONTACT INFORMATION	272 Elm Street New Haven, CT 06511	801-493-9166 tao.yu@yale.edu
RESEARCH INTERESTS	Natural language processing, deep learning, and machine learning. More specifically, I am currently interested in sequence to SQL, question answering, knowledge graph, and text generation.	
EDUCATION	Yale University , New Haven, CT Ph.D. Computer Science - Natural Language Processing 08/2017 - Present • Advisors: Dragomir R. Radev Columbia University , SEAS, New York, NY M.S. Data Science GPA: 3.82/4.0 09/2015 - 05/2017 • 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 • Advisors: Jingyi Zhu and Peter Philips	
RESEARCH & INDUSTRY EXPERIENCE	Yale University , New York, NY Graduate Student Research – NLP 08/2017 - Present • Working on text to SQL and sequential question answering projects. Samsung Research America , Mountain View, CA Research Intern – NLP 06/2018 - 08/2018 Columbia CCLS & NLP Group , New York, NY Graduate Research Assistant 05/2016 - 10/2016 • 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. AIG Science , New York, NY App Development Intern 01/2016 - 03/2016 • Developed a Hybrid Mobile App using React Native and built backend using Flask.	
PUBLICATIONS & PAPERS IN PREPARATION	<ol style="list-style-type: none">1. 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 “. <i>In the 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP)</i>, 2018. <i>Long Paper</i>2. 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“. <i>In the 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP)</i>, 2018. <i>Long Paper</i>3. T. Yu, Z. Li, Z. Zhang, R. Zhang, D. Radev. “TypeSQL: Knowledge-based Type-Aware Neural Text-to-SQL Generation“. <i>In the 16th Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT)</i>, 2018. <i>Short Paper</i>4. M. S. Rasooli, N. Farra, A. Radeva, T. Yu and K. McKeown. “Cross-Lingual Sentiment Transfer with Limited Resources”. <i>Machine Translation</i>, 2017. <i>Long Paper</i>	

5. 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*
6. **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*

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

PROGRAMMING & OTHER SKILLS Proficient in Python, PyTorch, TensorFlow, R, Java, Matlab, SQL. Familiar with C, C++, JavaScript, Theano, Hadoop, Spark. FAA Student Pilot.