

Rui Zhang

CONTACT	Room 333 17 Hillhouse Avenue New Haven, CT 06511	r.zhang@yale.edu https://ryanzhumich.github.io/ +1-734-741-3578
RESEARCH INTERESTS	My research interest lies in various natural language processing problems in understanding, generation, and grounding that requires effective understanding of contexts. To be specific, I propose models based on neural network to solve challenging problems in (1) End-to-End Neural Modeling for Entities, Sentences, Documents, and Multi-party Multi-turn Dialogues, (2) Text Summarization for Emails, News, and Scientific Articles, (3) Cross-lingual Information Retrieval for Low-Resource Languages, (4) Context-Dependent Text-to-SQL Semantic Parsing in Human-Computer Interaction.	
EDUCATION	Yale University	Aug 2017 - May 2020
	Ph.D., Computer Science Advised by Prof. Dragomir Radev	
	University of Michigan, Ann Arbor	Aug 2015 - May 2017
	Ph.D. student, Computer Science Advised by Prof. Dragomir Radev and Prof. Honglak Lee GPA 4.0/4.0	
	University of Michigan, Ann Arbor	Sept 2013 - Apr 2015
	B.S.E., Computer Engineering Summa Cum Laude GPA 3.9/4.0	
	Shanghai Jiao Tong University	Sept 2011 - Aug 2015
RESEARCH INTERNSHIPS	Google AI , Mountain View	June 2019 - Aug 2019
	Research Intern mentored by Dr. Yang Li Deep Learning and Human-Computer Interaction	
	Grammarly Research , New York	May 2018 - Aug 2018
	Research Intern mentored by Dr. Joel Tetreault Email Subject Line Generation (published in ACL 2019)	
PUBLICATIONS	IBM Thomas J. Watson Research Center , New York	May 2017 - Aug 2017
	Research Intern mentored by Dr. Cícero Nogueira dos Santos and Dr. Bing Xiang Neural Coreference Resolution (published in ACL 2018)	
	[16] 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 and Dragomir Radev. <i>In the 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2019</i>	
	[15] 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 and Dragomir Radev.

In the 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2019

[14] This Email Could Save Your Life: Introducing the Task of Email Subject Line Generation.

Rui Zhang, Joel Tetreault.

In the 57th Annual Meeting of the Association for Computational Linguistics (ACL), 2019

[13] Improving Low-Resource Cross-lingual Document Retrieval by Reranking with Deep Bilingual Representations.

Rui Zhang, Caitlin Westerfield, Sungrok Shim, Garrett Bingham, Alexander Fabbri, Neha Verma, William Hu, Dragomir Radev.

In the 57th Annual Meeting of the Association for Computational Linguistics (ACL), 2019

[12] 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 and Dragomir Radev.

In the 57th Annual Meeting of the Association for Computational Linguistics (ACL), 2019

[11] ScisummNet: A Large Annotated Corpus and Content-Impact Models for Scientific Paper Summarization with Citation Networks.

Michihiro Yasunaga, Jungo Kasai, **Rui Zhang**, Alexander Fabbri, Irene Li, Dan Friedman, Dragomir Radev.

In the Thirty-Third AAAI Conference on Artificial Intelligence (AAAI), 2019

[10] Surprise Languages: Rapid-Response Cross-Language IR.

Douglas Oard, Petra Galuscakova, Kathleen McKeown, Marine Carpuat, Mohamed Elbadrashiny, Ramy Eskander, Kenneth Heafield, Efsun Kayi, Chris Kedzie, Smaranda Muresan, Suraj Nair, Xing Niu, Dragomir Radev, Anton Ragni, Han-Chin Shing, Yan Virin, Weijia Xu, **Rui Zhang**, Elena Zotkina, Joseph Barrow, Mark Gales.

In the Ninth International Workshop on Evaluating Information Access (EVIA), 2019

[9] 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.

In the 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2018

[8] 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.

In the 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2018

[7] Neural Coreference Resolution with Deep Biaffine Attention by Joint Mention Detection and Mention Clustering.

Rui Zhang, Cícero Nogueira dos Santos, Michihiro Yasunaga, Bing Xiang, Dragomir Radev.

In the 56th Annual Meeting of the Association for Computational Linguistics (ACL), 2018

[6] Improving Text-to-SQL Evaluation Methodology.

Catherine Finegan-Dollak, Jonathan K. Kummerfeld, Li Zhang, Karthik Ramanathan, Sesh Sadasivam, **Rui Zhang**, Dragomir Radev.

In the 56th Annual Meeting of the Association for Computational Linguistics (ACL), 2018

[5] TypeSQL: Knowledge-based Type-Aware Neural Text-to-SQL Generation.

Tao Yu, Zifan Li, Zilin Zhang, **Rui Zhang**, Dragomir Radev.

In the 16th Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT), 2018

[4] Addressee and Response Selection in Multi-Party Conversations with Speaker Interaction RNNs.

Rui Zhang, Honglak Lee, Lazaros Polymenakos, Dragomir Radev.

In The Thirty-Second AAAI Conference on Artificial Intelligence (AAAI), 2018

[3] Graph-based Neural Multi-Document Summarization.

Michihiro Yasunaga, **Rui Zhang**, Kshitijh Meelu, Ayush Pareek, Krishnan Srinivasan, Dragomir Radev.

In the Conference on Computational Natural Language Learning (CoNLL), 2017

[2] Effects of Text Corpus Properties on Short Text Clustering Performance.

Catherine Finegan-Dollak, Reed Coke, **Rui Zhang**, Xiangyi Ye, Dragomir Radev.

In the 54th Annual Conference of the Association for Computational Linguistics (ACL), 2016

[1] Dependency Sensitive Convolutional Neural Networks for Modeling Sentences and Documents.

Rui Zhang, Honglak Lee, Dragomir Radev.

In the 15th Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT), 2016

PROFESSIONAL SERVICE

Organization Committee member, Int-Ex 2020

Program Committee member, the 13th Workshop on Graph-Based Natural Language Processing (TextGraphs 2019)

Program Committee member, the Workshop on e-Commerce and NLP (ECNLP 2019)

Program committee member, the Conference on Empirical Methods in Natural Language Processing (EMNLP), 2019

Program Committee member, the Annual Meeting of the Association for Computational Linguistics (ACL), 2019

Program Committee member, the Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), 2019

Program Committee member, the AAAI Conference on Artificial Intelligence (AAAI), 2019

Journal Reviewer, Transactions on Knowledge and Data Engineering (TKDE), 2019

Student Volunteer, the Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), 2016

MENTORING

Michihiro Yasunaga (2017). Undergraduate independent study.
Graph-based Neural Multi-Document Summarization.

Garrett Bingham (2018). Undergraduate independent study.
Predicting Cutoffs for Information Retrieval.

Caitlin Westerfield (2018). Undergraduate thesis project.
Learning to Rank for Low-Resource Cross-lingual Information Retrieval.

Sungrok Shim (2019). Undergraduate independent study.
Text-to-SQL with Graph Convolutional Neural Networks.

Eric Xue (2019). Undergraduate independent study.
Text-to-SQL using GANs for Robust Training.

TEACHING
ASSISTANT

Natural Language Processing, Spring 2018, Yale University	evaluation score: 4.4/5.0
Artificial Intelligence, Fall 2017, Yale University	evaluation score: 4.8/5.0
Natural Language Processing, Fall 2016, University of Michigan	
Natural Language Processing, Coursera	
Artificial Intelligence, Winter 2015, University of Michigan	
Intro to Programming, Fall 2014, University of Michigan	
East Asian Study, Summer 2013, Shanghai Jiao Tong University	

HONORS AND
AWARDS

Yale Conference Travel Fellowship	Yale University, 2017, 2018
Rackham Conference Travel Grant	University of Michigan, March 2016
CSE Departmental Fellowship	University of Michigan, 2015-2016
Outstanding Undergraduate Research Award	University of Michigan, March 2015
James B. Angell Scholar	University of Michigan, 2015
Wang Chu Chien-Wen Research Scholarship	University of Michigan, May 2014
Academic Excellence Scholarship (top 1%)	University of Michigan, Oct 2012, Oct 2013
Bao Steel Excellence Scholarship	Shanghai Jiao Tong University, Oct 2012
University Merit Student	Shanghai Jiao Tong University, Sept 2012
Dean's List	Every Semester

PRESENTATIONS

“Interleaving Thoughts of Interlocutors via RNNs for Response Selection in Multi-Turn Dialog Systems”.

Invited poster at the 5th IBM Research Cognitive Colloquium, the IBM T.J. Watson Research Center, New York, September 2016

“This Email Could Save Your Life: Introducing the Task of Email Subject Line Generation”.

Invited poster at the Microsoft Research AI Breakthroughs Workshop, Redmond, September 2019

“Editing-based SQL Query Generation for Cross-Domain Context-Dependent Questions”.

Natural Language, Dialog and Speech (NDS) Symposium, The New York Academy of Sciences, New York, November 2019