

CURRICULUM VITAE

PERSONAL INFORMATION Ruisheng Cao
🏠 June 23rd, 1996
🌐 [rhythmcao](#)
✉ ruishengcao@gmail.com
☎ +86 18217583036
🌐 [rhythmcao.github.io](#)



EDUCATION

- PhD candidate in Computer Technology, Shanghai Jiao Tong University 2021-present
- M.Eng. in Computer Technology, Shanghai Jiao Tong University 2018-2021
 - GPA : 3.94/4.0 Rankings : 3
- Zhiyuan Honors Program of Engineering (an elite program for **top 5%** students) 2014-2018
- Summer Program at Department of Computer Science, Cornell University Jun-Jul 2017
- Humanities Program at Hertford College part of the University of Oxford Aug 2015
- B.Eng. in Computer Science and Technology, Shanghai Jiao Tong University 2014-2018
 - GPA : 3.87/4.3 Rankings : top 10%

HONORS AND AWARDS

- **The 3rd Wu Wenjun AI Honorary Doctoral Program** (only **15** students in SJTU) 2021
- **Outstanding Graduate of Shanghai Jiao Tong University** 2021
- **National Scholarship for Graduate Students** 2018-2019, 2019-2020
- **Zhiyuan Honor Degree of Bachelor of Science** (only **48** undergraduates in SJTU) 2018
- **Zhiyuan Honors Scholarship (top 5%)** 2014-2015, 2015-2016
- **Academic Excellence Scholarship (First-Class, top 4)** 2014-2015

RESEARCH INTERESTS My research interests mainly lie in natural language understanding, especially semantic parsing, text-to-SQL, spoken language understanding, and other structured knowledge grounding problems. The already published work focuses on 1) dedicated model architecture which integrates structured knowledge into neural-symbolic processing, and 2) pseudo data generation or augmentation which leverages the dual model through cycle learning and iterative training.

PUBLICATIONS

- **Ruisheng Cao**, Lu Chen, Jieyu Li, Hanchong Zhang, Hongshen Xu, Wangyou Zhang and Kai Yu. *A Heterogeneous Graph to Abstract Syntax Tree Framework for Text-to-SQL*. In IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2023 (01) : 1-16.
- **Ruisheng Cao**, Lu Chen, Zhi Chen, Yanbin Zhao, Su Zhu and Kai Yu. *LGESQL : Line Graph Enhanced Text-to-SQL Model with Mixed Local and Non-Local Relations (Long Oral)*. In 59th Annual Meeting of the Association for Computational Linguistics (ACL) and the 11th International Joint Conference on Natural Language Processing, pages 2541-2555, 2021.
- **Ruisheng Cao**, Su Zhu, Chenyu Yang, Chen Liu, Rao Ma, Yanbin Zhao, Lu Chen and Kai Yu. *Unsupervised Dual Paraphrasing for Two-stage Semantic Parsing (Long Online Oral)*. In 58th Annual Meeting of the Association for Computational Linguistics (ACL), Seattle, USA, pages 6806-6817, 2020.
- Su Zhu*, **Ruisheng Cao*** (equal contribution, responsible for codes) and Kai Yu. *Dual Learning for Semi-Supervised Natural Language Understanding*. IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 28, pp. 1936-1947, 2020.
- **Ruisheng Cao**, Su Zhu, Chen Liu, Jieyu Li and Kai Yu. *Semantic Parsing with Dual Learning (Long Oral)*. In 57th Annual Meeting of the Association for Computational Linguistics (ACL), Florence, Italy, pages 51-64, 2019.

Currently, 3 conference papers (ACL) and 2 journal articles (TPAMI and TASLP) as first or co-first author. Other co-authored publications are listed below.

PUBLICATIONS

- Jieyu Li, Lu Chen, **Ruisheng Cao**, Su Zhu, Hongshen Xu, Zhi Chen, Hanchong Zhang and Kai Yu. *Exploring Schema Generalizability of Text-to-SQL*. In Findings of the Association for Computational Linguistics (ACL), Toronto, Canada, pages 1344-1360, 2023.
- Hanchong Zhang, Jieyu Li, Lu Chen, **Ruisheng Cao**, Yunyan Zhang, Yu Huang, Yefeng Zheng, and Kai Yu. *CSS : A Large-scale Cross-schema Chinese Text-to-SQL Medical Dataset*. In Findings of the Association for Computational Linguistics (ACL), Toronto, Canada, pages 6970-6983, 2023.
- Sheng Jiang, Su Zhu, **Ruisheng Cao**, Qingliang Miao and Kai Yu. *SPM : A Split-Parsing Method for Joint Multi-Intent Detection and Slot Filling*. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL), Volume 5 Industry Track, Toronto, Canada, pages 668-675, 2023.
- Jieyu Li, Zhi Chen, Lu Chen, Zichen Zhu, Hanqi Li, **Ruisheng Cao** and Kai Yu. *DIR : A Large-Scale Dialogue Rewrite Dataset for Cross-Domain Conversational Text-to-SQL*. Applied Sciences 13, no. 4 : 2262, 2023.
- Zihan Zhao, Lu Chen, **Ruisheng Cao**, Hongshen Xu, Xingyu Chen and Kai Yu. *TIE : Topological Information Enhanced Structural Reading Comprehension on Web Pages*. Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics : Human Language Technologies (NAACL). Seattle, United States, pages 1808-1821, 2022.
- Da Ma, Xingyu Chen, **Ruisheng Cao**, Zhi Chen, Lu Chen, and Kai Yu. *Relation-Aware Graph Transformer for SQL-to-Text Generation*. Applied Sciences 12, no. 1 : 369, 2022.
- Su Zhu, Lu Chen, **Ruisheng Cao**, Zhi Chen, Qingliang Miao and Kai Yu. *Few-Shot NLU with Vector Projection Distance and Abstract Triangular CRF*. Natural Language Processing and Chinese Computing : 10th CCF International Conference, NLPC 2021, Qingdao, China, October 13-17, 2021, Proceedings, Part I 10. Springer International Publishing, 2021 : 505-516.
- Zhi Chen, Lu Chen, Hanqi Li, **Ruisheng Cao**, Da Ma, Mengyue Wu and Kai Yu. *Decoupled Dialogue Modeling and Semantic Parsing for Multi-Turn Text-to-SQL*. Findings of the Association for Computational Linguistics : ACL-IJCNLP, pages 3063-3074, 2021.
- Zhi Chen, Lu Chen, Yanbin Zhao, **Ruisheng Cao**, Zihan Xu, Su Zhu and Kai Yu. *ShadowGNN : Graph Projection Neural Network for Text-to-SQL Parser*. Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics : Human Language Technologies (NAACL), online, pages 5567-5577, 2021.
- Yanbin Zhao, Lu Chen, Zhi Chen, **Ruisheng Cao**, Su Zhu and Kai Yu. *Line Graph Enhanced AMR-to-Text Generation with Mix-Order Graph Attention Networks*. In 58th Annual Meeting of the Association for Computational Linguistics (ACL), Seattle, USA, pages 732-741, 2020.
- Chen Liu, Su Zhu, Zijian Zhao, **Ruisheng Cao**, Lu Chen and Kai Yu. *Jointly Encoding Word Confusion Network and Dialogue Context with BERT for Spoken Language Understanding*. In 21st Annual Conference of the International Speech Communication Association (InterSpeech), Shanghai, China, pages 871-875, 2020.

The complete checklist is available on [Google scholar](#).

CHALLENGES AND COMPE- TITIONS

- Rank 1** [DuSQL](#) : A Large-Scale and Pragmatic Chinese Text-to-SQL Dataset. 2021/10-present
- Rank 1** [Spider](#) : Yale Semantic Parsing and Text-to-SQL Challenge. 2021/06-2022/09
- Rank 1** [CSpider](#) : The Chinese Semantic Parsing and Text-to-SQL Challenge. 2022/05-2023/06
- Rank 1** [NL2SQL](#) : A Large-Scale Chinese Single-Table Text-to-SQL Dataset. 2022/04-present
- Rank 2** The 9th China National Conference on Social Media Processing (SMP 2020). The Evaluation of Chinese Human-Computer Dialogue Technology, ECDT : Few Shot Spoken Language Understanding. 2020