

王晓智

✉ wangxz20@mails.tsinghua.edu.cn

🏠 <https://bakser.github.io>

教育经历

2020 – 至今 📌 博士研究生. 计算机科学与技术系, 清华大学.

导师: 李涓子 教授

2016 – 2020 📌 工学学士. 计算机科学与技术系, 清华大学.

工作经历

2017 – 2020 📌 研究助理 清华大学自然语言处理实验室, 受刘知远副教授指导。

2019.7 – 2019.9 📌 访问学生 蒙特利尔学习算法研究所 (Mila), 受 Prof. Jian Tang 指导。

研究经历

研究兴趣

主要集中在自然语言处理领域的两个话题: (1) **大模型机理分析**, 即通过探测和观察手段尝试理解大模型的工作原理, 从而从基础意义上改进大模型; (2) **知识获取**, 即如何从自然语言中理解并抽取复杂结构化知识, 从而丰富知识图谱等多元知识库, 支撑广泛下游应用。

大模型机理分析: 主要论文

- 1 Peng, H., **Wang, X.**, Hu, S., Jin, H., Hou, L., Li, J., ... Liu, Q. (2022). COPEN: Probing conceptual knowledge in pre-trained language models. In *Proceedings of EMNLP*.
- 2 Su, Y., **Wang, X.**, Qin, Y., Chan, C.-M., Lin, Y., Wang, H., ..., Li, J. et al. (2022). On transferability of prompt tuning for natural language processing. In *Proceedings of NAACL*.
- 3 **Wang, X.**, Wen, K., Zhang, Z., Hou, L., Liu, Z. & Li, J. (2022). Finding skill neurons in pre-trained transformer-based language models. In *Proceedings of EMNLP*.
- 4 Qin, Y., **Wang, X.**, Su, Y., Lin, Y., Ding, N., Yi, J., ..., Hou, L. et al. (2021). Exploring universal intrinsic task subspace via prompt tuning. In *Arxiv e-prints*.

知识获取: 主要论文

- 1 Peng, H., **Wang, X.**, Yao, F., Zeng, K., Hou, L., Li, J., ... Shen, W. (2023). The devil is in the details: On the pitfalls of event extraction evaluation. *Findings of ACL*.
- 2 **Wang, X.**, Chen, Y., Ding, N., Peng, H., Wang, Z., Lin, Y., ..., Liu, Z. et al. (2022). MAVEN-ERE: A unified large-scale dataset for event coreference, temporal, causal, and subevent relation extraction. *Proceedings of EMNLP*.

- 3 Wang, X., Gao, T., Zhu, Z., Zhang, Z., Liu, Z., Li, J. & Tang, J. (2021). KEPLER: A unified model for knowledge embedding and pre-trained language representation. *Transactions of the Association for Computational Linguistics*.
- 4 Wang, Z., Wang, X., Han, X., Lin, Y., Hou, L., Liu, Z., ... Zhou, J. (2021). CLEVE: Contrastive Pre-training for Event Extraction. *Proceedings of ACL*.
- 5 Wang, X., Jia, S., Han, X., Liu, Z., Li, J., Li, P. & Zhou, J. (2020). Neural Gibbs Sampling for Joint Event Argument Extraction. *Proceedings of AACL*.
- 6 Wang, X., Wang, Z., Han, X., Jiang, W., Han, R., Liu, Z., ... Zhou, J. (2020). MAVEN: A Massive General Domain Event Detection Dataset. *Proceedings of EMNLP*.
- 7 Wang, X., Han, X., Liu, Z., Sun, M. & Li, P. (2019). Adversarial training for weakly supervised event detection. *Proceedings of NAACL*, 998–1008.
- 8 Wang, X., Wang, Z., Han, X., Liu, Z., Li, J., Li, P., ... Ren, X. (2019). HMEAE: Hierarchical modular event argument extraction. *Proceedings of EMNLP*.
- 9 Wang, X., Han, X., Lin, Y., Liu, Z. & Sun, M. (2018). Adversarial multi-lingual neural relation extraction. *Proceedings of COLING*.

仅列出（共同）一作论文，完整论文列表见scholar.google.com/citations?user=DjpXXZkAAAAJ






学术服务

Program Committee Member/Reviewer: AAAI/IJCAI/COLING 2020, AAAI/ACL/EMNLP 2021, AAAI/COLING/SIGIR/CCKS/EMNLP 2022, AAAI/ACL/EMNLP/NeurIPS 2023, ACL Rolling Review.

重要荣誉

- 2020
- 清华大学优良毕业生
 - 清华大学计算机系优秀毕业生
 - 清华大学优秀本科毕业论文
- 2019
- 钟士模奖学金. 清华大学计算机系本科生最高荣誉，每年级 3 人。
 - “清华之友-蔡雄”奖学金. 清华大学本科生科技创新领域最高荣誉，每年 10 人。
 - 商汤奖学金. 商汤公司全国选拔 29 名 AI 领域优秀本科生给予奖励。
 - 腾讯“犀牛鸟”精英人才培养计划成员，旨在校企合作培养高素质研究人才。全国大三以上（及研究生）学生中选拔 56 人。结题评审获二等奖学金（前三名）。
 - 清华大学“未来学者”计划成员，全校本科生中选拔 10 名，给予 20 万元经费支持。
 - CCF 优秀大学生奖，由中国计算机学会（CCF）评选，全国高校本科生中 73 人获奖。

重要荣誉 (continued)

- 2018  **清华大学星火计划第十二期成员**, 旨在培养高水平科技创新人才, 全校 2016 级 3000 余名本科生中仅不足 50 人入选。
- 2017  **金奖**. ACM/ICPC 青岛区域赛
-  **金奖**. CCF 大学生计算机系统与程序设计竞赛 (CCF CCSP)
- 2016  **金奖**. 中国大学生程序设计竞赛 (CCPC) 长春区域赛
- 2015  **三等奖**. NOI 全国青少年信息学奥林匹克竞赛 (国家级)