

XIANG (TOMMY) YUE

Email: yue.149@osu.edu

Homepage: <https://xiangyue9607.github.io/>

EDUCATION

| | | |
|--------------|---|---------------------|
| Ph.D Student | Department of Computer Science and Engineering(CSE), The Ohio State University (OSU), USA Advisor: Prof. Huan Sun | <i>2018-present</i> |
| B.Eng. | School of Computer Science, Wuhan University (WHU), China GPA: 3.79/4.0, Rank: top 1%, Outstanding Graduates Advisor: Prof. Wen Zhang <i>Thesis: Drug-disease associations mining and study (in Chinese)</i> <i>(Excellent Graduation Thesis Award, top 5%)</i> | <i>2014-2018</i> |

RESEARCH INTERESTS

I have broad interests in Natural Language Processing with an emphasis on Question Answering (QA). Particularly, my research aims to build **trustworthy QA systems in the wild**, which must retain high standard aspects such as *generalization*, *privacy awareness*, and *transparency*.

Before doing NLP, I also had experience in Data Mining (Graph Embedding and Bioinformatics).

SELECTED PUBLICATIONS

([Google Scholar](#)) ([Semantic Scholar](#))

* indicates equal contributions

NLP (Question Answering, Clinical NLP, Privacy-preserving NLP):

- [1] (Ongoing project) “On the Generalizability of Open Retrieval QA”
- [2] (In submission) **Xiang Yue**, Ziyu Yao and Huan Sun, “Synthetic Question Value Estimation for Domain Adaptation of Question Answering”
- [3] Frederick Zhang, Heming Sun, **Xiang Yue**, Simon Lin and Huan Sun, “COUGH: A Challenge Dataset and Models for COVID-19 FAQ Retrieval”, **EMNLP 2021, Short Paper**
- [4] **Xiang Yue***, Minxin Du*, Tianhao Wang, Yaliang Li, Huan Sun and Sherman S. M. Chow, “Differential Privacy for Text Analytics via Natural Text Sanitization”, **ACL-IJCNLP 2021, Findings, Long Paper**
- [5] **Xiang Yue***, Xinliang (Frederick) Zhang*, Ziyu Yao, Simon Lin and Huan Sun, “CliniQG4QA: Generating Diverse Questions to Improve Clinical Reading Comprehension on New Contexts”, **NeurIPS 2020, Machine Learning for Health Workshop**
- [6] **Xiang Yue** and Shuang Zhou, “PHICON: Improving Generalization of Clinical Text De-identification Models via Data Augmentation”, **EMNLP 2020 Clinical NLP Workshop**
- [7] **Xiang Yue**, Bernal Jimenez Gutierrez and Huan Sun, “Clinical Reading Comprehension: A Thorough Analysis of the emrQA Dataset”, **ACL 2020, long paper**
- [8] Bernhard Kratzwald, **Xiang Yue**, Huan Sun and Stefan Feuerriegel, “Practical Annotation Strategies for Question Answering Datasets”, *arXiv preprint*

Data Mining (Graph Embedding and Bioinformatics):

- [9] **Xiang Yue**, Zhen Wang, Jingong Huang, Srinivasan Parthasarathy, Soheil Moosavinasab, Yungui Huang, Simon Lin, Wen Zhang, Ping Zhang and Huan Sun, "Graph Embedding on Biomedical Networks: Methods, Applications, and Evaluations", *Bioinformatics*, 2020 (Impact Factor: 4.531) **(ESI Highly Cited Paper: top 1% cited paper of its academic field)**
- [10] Zhen Wang, **Xiang Yue**, Soheil Moosavinasab, Yungui Huang, Simon Lin and Huan Sun, "Surf-Con: Synonym Discovery on Privacy-Aware Clinical Data", *The 25th ACM SIGKDD Conference on Knowledge Discovery and Data Mining 2019 (KDD 2019, research track, oral)*

INTERNSHIP

- Tencent AI Lab, *Bellevue, Washington* *2021.5-2021.8*
(Mentors: Dong Yu, Xiaoman Pan, Jianshu Chen, Wenlin Yao, Dian Yu)
 - Project: On the Generalizability of Open Retrieval QA
 - We comprehensively evaluate the generalizability of open retrieval QA systems and propose to extract pseudo-labelled Question-Passage pairs from the external knowledge source (e.g., Wikipedia) to improve the generalizability of the retriever (2%-4% absolute improvement on Accuracy@5/20).

ACADEMIC SERVICES

- Program Committee/Reviewer:
 - NLPCC'21, NAACL'20, KDD'19,20, AMIA 2020, BIBM 2018, 2020
 - IEEE Transactions on Knowledge and Data Engineering (TKDE)
 - IEEE Transactions on Neural Networks and Learning Systems
 - Bioinformatics
 - Nature Scientific Reports
 - ACM Transactions on Computing for Healthcare
 - BMC Medical Informatics and Decision Making

HONORS & AWARDS

- Best Student Research Poster 2020, OSU CSE
- KDD 2020, 2019 Student Travel Award
- Excellent Graduation Thesis Award of WHU (Scale: 5%) *June 2018*
- Outstanding Graduates of WHU (Scale: 10%) *May 2018*
- LEI JUN Scholarship (Scale: Top 1 Winner of National Scholarship, the highest prize for students in WHU) *2016-2017*
- First Class Scholarship (Scale: 5%), Three Times, WHU *2014-2017*
- Excellent Student (Scale: 5%), Three Times, WHU *2014-2017*
- National Scholarship (Scale: 1%), China *2014-2015*