Wenhao Yu

CONTACTS E-mail: wenhaoyu97@gmail.com Websites: (1) Homepage (2) Google Scholar

WORKING EXPERIENCE
Tencent AI Lab, Seattle, Washington, US Aug. 2023 – Now
Senior Research Scientist in NLP

EDUCATION EXPERIENCE
University of Notre Dame, IN, US Aug. 2019 – Jul. 2023
Ph.D. in Computer Science and Engineering GPA: 3.95/4.0
- Advisor: Dr. Meng Jiang `http://www.meng-jiang.com/`
- Committee members: Dr. Heng Ji (UIUC), Dr. Scott Wen-tau Yih (FAIR),

Sichuan University, Chengdu, China Aug. 2015 – Jul. 2019
B.S. in Computer Science and Technology GPA: 3.84/4.0
- Graduated with honors

RESEARCH INTERESTS
I am interested in **Language Model + Knowledge** (i.e., knowledge-enhanced methods) for solving various real-world applications, such as open-domain question answering, commonsense reasoning and natural language generation. Besides, I am aslo actively contributing to open-community project, such as StarCoder.

INDUSTRY EXPERIENCE
Research Intern at Allen Institute of AI (AI2), Seattle, Washington, US
Worked with Drs. Ashish Sabharwal and Peter Clark Sept. 2022 - May 2023

Visiting Student Researcher at Microsoft Research, Redmond, Washington, US
Worked with Drs. Chenguang Zhu and Michael Zeng Sept. 2021 - Sept. 2022

Research Intern at Microsoft Research, Redmond, Washington, US
Worked with Drs. Chenguang Zhu and Michael Zeng June 2021 - Sept. 2021

HONORS AND AWARDS
Bloomberg Ph.D. Fellowship 2022
Best Paper Award at SoCalNLP 2022
NAACL 2022 Travel Award 2022
Notre Dame Outstanding Student Fellowship 2019
National Scholarship of China 2018

WORKSHOP ORGANIZATION
[W1] Chenguang Zhu, Shuohang Wang, Meng Jiang **Wenhao Yu**, Lu Wang, Huan Sun. "The First Workshop on Knowledge Augmented Methods for Natural Language Processing", in AAAI 2023. [Website] [Blog] [Twitter]
- With more than 70 people (virtual & in-person) attending the workshop

CONFERENCE TUTORIAL
[T2] Chenguang Zhu, Yichong Xu, Xiang Ren, Yuchen Lin, Meng Jiang, **Wenhao Yu**. "Knowledge-Augmented Methods for Natural Language Processing", in ACL 2022.
- With more than 200 people (virtual & in-person) attending the tutorial
- With more than 200 people following our Github repository [Link]

[T1] **Wenhao Yu**, Meng Jiang, Zhiting Hu, Qingyun Wang, Heng Ji, Nazneen Fatema Rajani. "Knowledge-Enriched Natural Language Generation", in EMNLP 2021.
- With more than 120 people (virtual & in-person) attending the tutorial
- With more than 500 people following our Github repository [Link]

PEER-REVIEWED PUBLICATIONS
[P14] [Language Model + Open-domain QA] Zhihan Zhang, **Wenhao Yu**, Zheng Ning, Mingxuan Ju, Meng Jiang. "Exploring Contrast Consistency of Open-Domain Question Answering Systems on Minimally Edited Questions", in TACL 2023.

[P13] [Large Language Model + Open-domain QA] Noah Ziems, **Wenhao Yu**, Zhihan Zhang, Meng Jiang. "Large Language Models are Built-in Autoregressive Search Engines", in ACL 2023.

[P12] [Large Language Model + Open-domain QA] **Wenhao Yu**, Dan Iter, Shuohang Wang, Yichong Xu, Mingxuan Ju, Soumya Sanyal, Chenguang Zhu, Michael Zeng, Meng Jiang. "Generate rather than Retrieve: Large Language Models are Strong Context Generators", in ICLR 2023.

[P11] [Document Retrieval, Commonsense QA] **Wenhao Yu**, Chenguang Zhu, Zhihan Zhang, Shuohang Wang, Zhuosheng Zhang, Meng Jiang. "Retrieval Augmentation for Commonsense Reasoning: A Unified Approach", in EMNLP 2022.

[P10] [Document Retrieval, Knowledge Graph] **Wenhao Yu**, Mingxuan Ju (equal contribution), Tong Zhao, Chuxu Zhang, Yanfang Ye. "Knowledge Graph Enhanced Passage Reader for Open-domain Question Answering", in EMNLP 2022.

[P9] [Entity Memory, Open-domain QA] Zhihan Zhang, **Wenhao Yu**, Chenguang Zhu, Meng Jiang. "A Unified Encoder-Decoder Framework with Entity Memory", in EMNLP 2022. (selected as oral presentation, 166 out of 1392 accepted papers)

[P8] [Document Retrieval + Knowledge Graph] Ziniu Hu, Yichong Xu, **Wenhao Yu**, Shuohang Wang, Ziyi Yang, Chenguang Zhu, Kai-Wei Chiang, Yizhou Sun. "Empowering Language Models with Knowledge Graph Reasoning for Open-Domain Question Answering", in EMNLP 2022.

[P7] [Multi-task Learning] Zhuosheng Zhang, Shuohang Wang, Yichong Xu, Yuwei Fang, **Wenhao Yu**, Yang Liu, Hai Zhao, Chenguang Zhu, Michael Zeng. "Task Compass: Scaling Multi-task Pre-training with Task Prefix", in EMNLP 2022.

[P6] [Language Model Pre-training + English Dictionary] **Wenhao Yu**, Chenguang Zhu, Yuwei Fang, Donghan Yu, Shuohang Wang, Yichong Xu, Meng Jiang. "Dict-BERT: Enhancing Language Model Pre-training with Dictionary", in ACL 2022.

[P5] [Document Retrieval + Knowledge graph] Donghan Yu, Chenguang Zhu, Yuwei Fang, **Wenhao Yu**, Shuohang Wang, Yichong Xu, Xiang Ren, Yiming Yang, Michael Zeng. "KG-FiD: Infusing Knowledge Graph in Fusion-in-Decoder for Open-Domain Question Answering", in ACL 2022.

[P4] [Survey, Knowledge Graph, Text Retrieval] **Wenhao Yu**, Chenguang Zhu, Zaitang, Li, Zhiting Hu, Qingyun Wang, Heng Ji, Meng Jiang. "A Survey of Knowledge-Enhanced Text Generation", in ACM Computing Survey 2022 (CSUR, IF:10.28).

[P3] [Language Generation, Entity Knowledge] Xiangyu Dong, **Wenhao Yu**, Chenguang Zhu, Meng Jiang. "Injecting Entity Types into Entity-Guided Text Generation", in EMNLP 2021 (selected as oral presentation).

[P2] [Document Retrieval, Transfer Learning] **Wenhao Yu**, Lingfei Wu, Yu Deng, Qingkai Zeng, Ruchi Mahindru, Sinem Guven, Meng Jiang. "Technical Question Answering across Tasks and Domains", in NAACL 2021.

[P1] [Joint Retrieval and Generation] **Wenhao Yu**, Lingfei Wu, Qingkai Zeng, Shu Tao, Yu Deng, Meng Jiang. "Crossing Variational Autoencoders for Answer Retrieval", in ACL 2020 (selected as oral presentation).

PH.D THESIS **Wenhao Yu**. "Retrieval-augmented Generation across Heterogeneous Knowledge", accepted to NAACL Ph.D thesis proposal track 2022.

MENTORED STUDENTS **Tong Chen**, Tecent AI intern, Ph.D student at the University of Washington
Zhihan Zhang, on [P9], now Ph.D student at the University of Notre Dame
Zaitang Li, on [P4], now Ph.D student at the Chinese University of Hong Kong
Xiangyu Dong, on [P3], now software engineer at the Bloomberg LP

PROFESSIONAL SERVICES Conference Reviewer and Journal Invited Reviewer:

- ICLR
- ICML
- Neurips
- AAAI
- IJCAI
- ACL
- EMNLP
- NAACL
- CSUR
- TACL