Wenhao Yu

CONTACT INFORMATION	355 Fitzpatrick Hall Department of Computer Science and Engineering University of Notre Dame, IN 46556, US	Phone: (+1) (574) 292-6516 E-mail: wyu1@nd.edu Website: wyu97.github.io

EDUCATION EXPERIENCE

University of Notre Dame, IN, US   Aug. 2019 – *present*
Ph.D. in Computer Science and Engineering   GPA: 3.95/4.0
- Advisor: Dr. Meng Jiang `http://www.meng-jiang.com/`
- External Committee members: Dr. Heng Ji (UIUC), Dr. Scott 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 knowledge-intensive applications, such as open-domain question answering, commonsense reasoning and natural language generation.

INDUSTRY EXPERIENCE

Research Intern at Bloomberg CTO office, New York City, New York, US
Will work with Drs. Mark Dredze and David Rosenberg   Will start from June 2023

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

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
First Place at the 14th Notre Dame CSE Annual Poster Contest	2020
Notre Dame CSE Select 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

[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**, Chuxu Zhang, Zhihan

Zhang, Shuohang Wang, Zhuosheng Zhang, Yuwei Fang, 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

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 PC/Reviewer

- ACL 2021-2023
- NeurIPS 2022
- AAAI 2021-2023
- EMNLP 2020-2022
- ICML 2022-2023
- ICLR 2023
- NAACL 2021-2022
- IJCAI 2021-2023
- ARR 2021-2023

Journal Invited Reviewer:

- CSUR
- TKDE
- TNNLS
- WWWJ
- KBS