

WEIJIA ZHANG

✉ weijia.zhangplus@gmail.com [in linkedin.com/in/weijia](https://www.linkedin.com/in/weijia) github.com/wjzhang392

Research Interest

- Large Language Models
- Hallucination Evaluation
- Retrieval-Augmented Generation
- Query-Focused Summarization

Education

University of Amsterdam <i>Ph.D. in IRLab, Informatics Institute</i> Advisor: Evangelos Kanoulas	Oct. 2020 – Feb. 2025 (Expected) Amsterdam, Netherlands
Sun Yat-sen University <i>M.E. in Information and Communication Engineering</i> Advisor: Guoli Wang	Sep. 2016 – Jun. 2019 Guangzhou, China
Sun Yat-sen University <i>B.E. in Automation</i>	Sep. 2012 – Jun. 2016 Guangzhou, China

Research Experience

Multi-Table Reasoning Supervisor: Evangelos Kanoulas <ul style="list-style-type: none">• Adopted chain-of-table prompting method to multi-table scenarios.• Explored plan-then-execute reasoning prompting strategies for multi-table question answering.	Mar. 2024 – present
Query-Focused Multi-Table Summarization (QFMTS) Supervisor: Evangelos Kanoulas <ul style="list-style-type: none">• Introduced the new task QFMTS and designed the pipeline of automatic data annotation using LLMs for the task.• Proposed two novel prompting-based methods for closed-source LLMs such as GPT-3.5.• Fine-tuned open-source Llama-2-chat models with QLoRA algorithms to benchmark the task.	May 2023 – Feb. 2024
Citation Evaluation in LLMs Supervisor: Evangelos Kanoulas <ul style="list-style-type: none">• Assessed the effectiveness of faithfulness metrics in fine-grained citation support scenarios.• Proposed comprehensive evaluation framework to evaluate metrics regarding correlation, classification, and retrieval evaluation.• Provided practical suggestions and directions for effective trustworthy faithfulness metrics.	Mar. 2022 – May 2023
Knowledge-Intensive Query-Focused Summarization Supervisor: Evangelos Kanoulas <ul style="list-style-type: none">• Extended query-focused summarization to more realistic knowledge-intensive task setting.• Built the human-annotated test collection with three different knowledge corpora.• Benchmarked the task with various models and demonstrated the new challenges.	Oct. 2020 – Feb. 2022

Working Experience

Qihoo 360 Search Lab <i>Research Intern</i> <ul style="list-style-type: none">• Improved large-scale data processing based on Hadoop, leading to faster speed.• Designed new features for existing GBDT model and increased F1 score of offline validation set by 1.2%.• Reproduced SOTA models of deep text matching and applied them to search scenarios.	Mar. 2018 – Jul. 2018 Beijing, China
--	--

Selected Publications

- [1] **Weijia Zhang**, Jia-Hong Huang, Svitlana Vakulenko, Yumo Xu, Thilina Rajapakse, Evangelos Kanoulas. *Beyond Relevant Documents: A Knowledge-Intensive Approach for Query-Focused Summarization using Large Language Models*. International Conference on Pattern Recognition (ICPR), 2024.
- [2] **Weijia Zhang**, Vaishali Pal, Jia-Hong Huang, Evangelos Kanoulas, Maarten de Rijke. *QFMTS: Generating Query-Focused Summaries over Multi-Table Inputs*. European Conference on Artificial Intelligence (ECAI), 2024.
- [3] **Weijia Zhang**, Mohammad Aliannejadi, Yifei Yuan, Jiahuan Pei, Jia-Hong Huang, Evangelos Kanoulas. *Towards Fine-Grained Citation Evaluation in Generated Text: A Comparative Analysis of Faithfulness Metrics*. International Natural Language Generation Conference (INLG), 2024.

- [4] **Weijia Zhang**, Svitlana Vakulenko, Thilina Rajapakse, Yumo Xu, Evangelos Kanoulas. *Tackling Query-Focused Summarization as A Knowledge-Intensive Task: A Pilot Study*. The First Workshop on Generative Information Retrieval (Gen-IR) at SIGIR, 2023.
- [5] **Weijia Zhang**, Svitlana Vakulenko, Thilina Rajapakse, Evangelos Kanoulas. *Scaling Up Query-Focused Summarization to Meet Open-Domain Question Answering*. Arxiv, 2021.
- [6] Zenan Xu, Qinliang Su, Xiaojun Quan, **Weijia Zhang**. *A Deep Neural Information Fusion Architecture for Textual Network Embeddings*. Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2019.
- [7] **Weijia Zhang**, Le Qin, Wei Zhong, Xuemei Guo, and Guoli Wang. *A Framework of Sequence Chunking for Human Activity Recognition Using Wearables*. Proceedings of the 2019 International Conference on Image, Video and Signal Processing (IVSP), 2019.
- [8] **Weijia Zhang**, Zhichao Tan, Guoli Wang, Xuemei Guo. *Quantified Living Habits Using RTI Based Target Footprint Data*. Proceedings of 2016 Chinese Intelligent Systems Conference: Volume II, 2016.

Technical Skills

Deep Learning: Hugging Face, PyTorch, Keras and TensorFlow

Programming: Python, Shell, C++ and Java