



Haowei Zhang

Date of birth: 04/03/1999 **Gender:** Male **Phone:** (+86) 18516760840

Email: haowei.zhang123@gmail.com

Website: <https://haowei-freesky.github.io/>

WORK EXPERIENCE

Yale NLP Lab – remote

City: remote | Website: <https://yale-nlp.github.io/>

Research Assistant

[10/2024 – 02/2025]

- video understanding and reasoning of multimodal LLMs
- one paper accepted at CVPR as co-first author

LMU Research Group – Munich, Germany

City: Munich | Country: Germany | Website: <https://sites.google.com/view/zhenhan/home>

Research Assistant

[10/2023 – 08/2024]

- improving reasoning and calibration capabilities of (multimodal) LLMs
- collaborating with Tresp Lab: <https://tresp-lab.github.io/>
- one paper accepted at EMNLP 2024 Findings as first author

Yale NLP Lab – remote

City: remote | Website: <https://yale-nlp.github.io/>

Research Assistant

[03/2023 – 10/2023]

- table-to-text generation capabilities of LLMs
- one paper accepted at EMNLP 2023 Industry Track as co-first author
- one paper accepted at EMNLP 2024 Demo Track (Oral) as first author

EDUCATION AND TRAINING

Master

Technical University of Munich (TUM) [30/09/2021 – 12/11/2024]

City: Munich | Country: Germany | Field(s) of study: Robotics, Cognition, Intelligence

- Core courses: Machine Learning, Deep Learning, Computer Vision, Techniques in AI, Guided Research in NLP
- Final grade: 1.7/1.0 (German system, equivalent to 91/100)

Exchange semesters in the bachelor program

Harz University of Applied Sciences [01/10/2020 – 30/09/2021]

City: Wernigerode | Country: Germany | Field(s) of study: Smart Automation | Final grade: The same as the grade at Tongji University (double-degree program)

Bachelor

Tongji University [01/09/2017 – 30/09/2021]

City: Shanghai | Country: China | Field(s) of study: Mechatronics | Final grade: 4.54/5.0 (equivalent to 91/100)

PROJECTS

[10/2024 – 02/2025]

MMVU: A Comprehensive Benchmark for Expert-Level Multi-Disciplinary Video Understanding and Reasoning

- comprehensive expert-level, multi-discipline benchmark for evaluating foundation models in video understanding

- attracting great attention from the community (download >10k on HF), used by Qwen2.5-VL, Kimi-VL, Seed1.5-VL, etc.
- accepted at CVPR 2025 as co-first author (stay tuned)

Links: <https://arxiv.org/abs/2501.12380> | <https://mmvu-benchmark.github.io/>

[10/2023 – 03/2024]

Visual Question Decomposition on Multimodal Large Language Models

- development of a novel evaluation framework to assess the quality of generated sub-questions
- proposing a finetuning pipeline to enhance the capabilities of multimodal LLMs for visual question decomposition
- instructing multimodal LLMs to perform selective question decomposition
- accepted at EMNLP 2024 Findings as first author

Links: <https://aclanthology.org/2024.findings-emnlp.107/> | <https://vqd-emnlp2024.github.io/>

[07/2023 – 09/2023]

OpenT2T: An Open-Source Toolkit for Table-to-Text Generation

- first open-source toolkit for table-to-text generation, designed to reproduce performance comparison and to expedite the development of new models
- accepted at EMNLP 2024 Demo Track (Oral) as first author

Link: <https://aclanthology.org/2024.emnlp-demo.27/>

[03/2023 – 07/2023]

Investigating Table-to-Text Generation Capabilities of Large Language Models in Real-World Information Seeking Scenarios

- investigation of the table-to-text capabilities of different LLMs within real-world information seeking scenarios
- demonstrating that high-performing LLMs can effectively serve as table-to-text generators, evaluators, and feedback generators
- accepted at EMNLP 2023 Industry Track as co-first author

Links: <https://aclanthology.org/2023.emnlp-industry.17/> | <https://github.com/yale-nlp/LLM-T2T>

COMMUNITY SERVICE

Organizer / Reviewer

Organizer of Knowledge-Intensive Multimodal Reasoning @ ICCV 2025

Reviewer of Reasoning and Planning for LLMs @ ICLR 2025

Reviewer of ACL Rolling Review 2025

HONOURS AND AWARDS

[04/2021] Tongji University

Outstanding Graduate of Tongji University

[11/2020] Deutscher Akademischer Austauschdienst (DAAD)

DAAD-Scholarship (German national level scholarship)

LANGUAGE SKILLS

Mother tongue(s): Chinese

Other language(s): English - C1, IELTS 7.0 | German - B2

SKILLS

Python / Natural Language Processing / (Multimodal) Large Language Models / Uncertainty Quantification / Social Computing / Web Mining