# Yuansheng Ni

Email: yuansheng.ni@uwaterloo.ca          Personal Website: https://yuanshengni.github.io

## EDUCATION

**University of Waterloo**                                                                                          Waterloo, CA
MS in Computer Science                                                                                  Sept 2024-Present
● Advisor: Wenhu Chen

**Zhejiang University**                                                                                        Zhejiang, China
B.E. in Computer Science and Technology                                                        Sept 2018-Jun 2024
**University of Waterloo**                                                                                          Waterloo, CA
Mathematics, Non Degree, Exchange                                                               Sept 2023-Dec 2023

## RESEARCH EXPERIENCE

**TIGER Lab, University of Waterloo**                                                                    Waterloo, CA
Research Assistant; Advisor: Wenhu Chen                                                         Sept 2023-Present

**ZJU NLP Group, Zhejiang University**                                                          Zhejiang, China
Research Assistant; Advisor: Ninyu Zhang                                                        Jul 2023-June 2024

## PUBLICATIONS

(* indicates equal contribution.)

1. **II-Bench: An Image Implication Understanding Benchmark for Multimodal Large Language Models**
   Ziqiang Liu, Feiteng Fang, Xi Feng, Xinrun Du, Chenhao Zhang, Zekun Wang, Yuelin Bai, Qixuan Zhao, Liyang Fan, Chengguang Gan, Hongquan Lin, Jiaming Li, Yuansheng Ni, Haihong Wu, Yaswanth Narsupalli, Zhigang Zheng, Chengming Li, Xiping Hu, Ruifeng Xu, Xiaojun Chen, Min Yang, Jiaheng Liu, Ruibo Liu, Wenhao Huang, Ge Zhang, Shiwen Ni
   In NeurIPS D&B 2024.

2. **GenAI Arena: An Open Evaluation Platform for Generative Models**
   Dongfu Jiang, Max Ku, Tianle Li, Yuansheng Ni, Shizhuo Sun, Rongqi Fan, Wenhu Chen
   In NeurIPS D&B 2024.

3. **MMLU-Pro: A More Robust and Challenging Multi-Task Language Understanding Benchmark**
   Yubo Wang, Xueguang Ma, Ge Zhang, Yuansheng Ni, Abhranil Chandra, Shiguang Guo, Weiming Ren, Aaran Arulraj, Xuan He, Ziyan Jiang, Tianle Li, Max Ku, Kai Wang, Alex Zhuang, Rongqi Fan, Xiang Yue, Wenhu Chen.
   In NeurIPS D&B(**Spotlight**) 2024.

4. **VideoScore: Building Automatic Metrics to Simulate Fine-grained Human Feedback for Video Generation**
   Xuan He, Dongfu Jiang, Ge Zhang, Max Ku, Achint Soni, Sherman Siu, Haonan Chen, Abhranil Chandra, Ziyan Jiang, Aaran Arulraj, Kai Wang, Quy Duc Do, Yuansheng Ni, Bohan Lyu, Yaswanth Narsupalli, Rongqi Fan, Zhiheng Lyu, Yuchen Lin, Wenhu Chen
   In EMNLP Main, 2024.

5. **A Massive Multi-discipline Multimodal Understanding and Reasoning Benchmark for Expert AGI**
   Xiang Yue*, Yuansheng Ni*, Kai Zhang*, Tianyu Zheng*, Ruoqi Liu, Ge Zhang, Samuel Stevens, Dongfu Jiang, Weiming Ren, Yuxuan Sun, Cong Wei, Botao Yu, Ruibin Yuan, Renliang Sun, Ming Yin, Boyuan Zheng, Zhenzhu Yang, Yibo Liu, Wenhao Huang, Huan Sun, Yu Su, and Wenhu Chen
   In CVPR(**Best Paper Finalist**), 2024.

6. **EasyEdit: An Easy-to-use Knowledge Editing Framework for Large Language Models**
   Peng Wang, Ningyu Zhang, Bozhong Tian, Zekun Xi, Yunzhi Yao, Ziwen Xu, Mengru Wang, Shengyu Mao,

Xiaohan Wang, Siyuan Cheng, Kangwei Liu, Yuansheng Ni, Guozhou Zheng, Huajun Chen
In ACL SDT 2024.

## PREPRINTS

1. **MEGA-Bench: Scaling Multimodal Evaluation to over 500 Real-World Tasks**
   Jiacheng Chen, Tianhao Liang, Sherman Siu, Zhengqing Wang, Kai Wang, Yubo Wang, Yuansheng Ni, Wang Zhu, Ziyan Jiang, Bohan Lyu, Dongfu Jiang, Xuan He, Yuan Liu, Hexiang Hu, Xiang Yue, Wenhu Chen
   arXiv:2410.10563, 2024.

2. **MMMU-Pro: A More Robust Multi-discipline Multimodal Understanding Benchmark**
   Xiang Yue*, Tianyu Zheng*, Yuansheng Ni*, Yubo Wang, Kai Zhang, Shengbang Tong, Yuxuan Sun, Ming Yin, Botao Yu, Ge Zhang, Huan Sun, Yu Su, Wenhu Chen, Graham Neubig
   arXiv:2409.02813, 2024.

3. **A Comprehensive Study of Knowledge Editing for Large Language Models**
   Ningyu Zhang, Yunzhi Yao, Bozhong Tian, Peng Wang, Shumin Deng, Mengru Wang, Zekun Xi, Shengyu Mao, Jintian Zhang, Yuansheng Ni, Siyuan Cheng, Ziwen Xu, Xin Xu, Jia-Chen Gu, Yong Jiang, Pengjun Xie, Fei Huang, Lei Liang, Zhiqiang Zhang, Xiaowei Zhu, Jun Zhou, Huajun Chen
   arXiv:2401.01286, 2024.

## SELECTED HONORS & AWARDS

| | |
|---|---|
| CVPR Best Paper Finalist | 2024 |
| Meritorious Winner, MCM/ICM, Consortium for Mathematics and Its Application | 2022 |
| Second Prize, Mathematical Modeling Competition, School of Mathematical Science, Zhejiang University | 2022 |
| Third Prize, CUMCM, Industrial and Applied Mathematics of P.R.China | 2018 |
| Second Prize, Collegiate Mathematics Competition, Department of Education of Zhejiang | 2018 |
| Second Prize, Mathematical Modeling Competition, School of Mathematical Science, Zhejiang University | 2018 |

## SKILLS

**Programming Languages**: Python; Java; C/C++; Matlab; LaTeX
**Deep Learning Frameworks**: PyTorch; TensorFlow