

Boyuan (Steven) Zheng

279 Dreese Laboratories, 2015 Neil Ave, Columbus, OH 43210, USA

✉ zheng.2372@osu.edu 🎓 [Google Scholar](#) 🏠 [Homepage](#) 🐦 [Twitter](#) 🔗 [LinkedIn](#)

EDUCATION

The Ohio State University, Columbus

2023 - present

Ph.D. Candidate in Computer Science

Advisors: Prof. Yu Su

Johns Hopkins University, Baltimore

2021 - 2023

MS in Computer Science

Advisors: Prof. Benjamin Van Durme

Northeastern University, Shenyang

2017 - 2021

B.Tech (Honours) in Software Engineering

PROFESSIONAL EXPERIENCES

The Ohio State University, Research Assistant

Sep 2023 - Present

- Constructed a dataset for web agent development (**Mind2Web**), involving dumping webpages and action trajectories
- Developed an automated testing tool for evaluating web agents in real-time on live websites.
- Developed multimodal web agent leveraging LMMs and enhanced image detail grounding by integrating image markups
- Developed MMMU benchmarks for modality understanding and evaluated GPT-4V's capabilities on it

Amazon Alexa AI, Applied Scientist Intern

June 2023 - Sep 2023

- Trained query rewriting models based on FLAN-T5 with million-level datasets.
- Augmented model performance and user personalization with retrieval from user history queries.
- Optimized model efficiency by retrieval deduplication based on self-information and pairwise mutual information.

Johns Hopkins University, Research Assistant

Jan 2022 - May 2023

Supervisor: Prof. Benjamin Van Durme

- Built a large-scale coreference resolution benchmark in multiparty dialogue with rich third-person reference (**TACL'23**)
- Developed mention linking annotation interface for dialogue coreference resolution task
- Developed dataset projection method to extend English dataset to 62 languages with low cost
- Conducted an empirical study on 3 typical span finding methods across 5 Information Extraction tasks (**EMNLP'22**)

PUBLICATIONS

1. **SeeAct: Grounded GPT-4V(ision) is a Generalist Web Agent**
Boyuan Zheng, Boyu Gou, Jihyung Kil, Huan Sun, Yu Su
2. **Dual-View Visual Contextualization for Web Navigation**
Jihyung Kil, Chan Hee Song, Boyuan Zheng, Xiang Deng, Yu Su, Wei-Lun Chao
3. **MMMU: 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, Wenhua Chen
4. **Mind2Web: Towards a Generalist Agent for the Web (NeurIPS'23 Spotlight)**
Xiang Deng, Yu Gu, Boyuan Zheng, Shijie Chen, Samuel Stevens, Boshi Wang, Huan Sun, Yu Su
5. **Flatness-Aware Prompt Selection Improves Accuracy and Sample Efficiency (EMNLP'23)**
Lingfeng Shen, Weiting Tan, Boyuan Zheng, Daniel Khashabi
6. **Multilingual Coreference Resolution in Multiparty Dialogue (TACL'23)**
Boyuan Zheng, Patrick Xia, Mahsa Yarmohammadi, Benjamin Van Durme
7. **An Empirical Study on Finding Spans (EMNLP'22)**
Weiwei Gu*, Boyuan Zheng*, Yunmo Chen, Tongfei Chen, Benjamin Van Durme

8. **Learn to Remember: Transformer with Recurrent Memory for Document-level Machine Translation (NAACL'22)**
Yukun Feng, Feng Li, Ziang Song, **Boyuan Zheng**, Philipp Koehn,
9. **SemEval-2021 Task 4: Reading Comprehension of Abstract Meaning (SemEval'21)**
Boyuan Zheng, Zhenhua Ling, Quan Liu, Yuping Ruan, Si Wei, Xiaodan Zhu
10. **Exploring Generalization Ability of Pretrained Language Models on Arithmetic and Logical Reasoning (NLPCC'21)**
Cunxiang Wang*, **Boyuan Zheng***, Yuchen Niu and Yue Zhang

SKILLS

Deepspeed, Redis, Transformers, Faiss, Python, PyTorch, Tensorflow, Spark, Java, SQL, JavaScript

ACADEMIC SERVICE

Reviewer: EMNLP-2023, ACL-2023, AACL-2023, ACL Rolling Review, SemEval-2022, SemEval-2024

Volunteer: NAACL-2022, ICML-2022, EMNLP-2022

SELECTED AWARDS AND HONORS

- National Scholarship *2018*
- The First Prize Scholarship of Northeastern University *2018*
- Pacemaker to Merit Student *2018*
- Interdisciplinary Contest In Modeling - Meritorious Winner *2019*