





YU ZHAO

✉ yu.zhao@ed.ac.uk  Google Scholar  GitHub  Twitter  LinkedIn

EDUCATION

University of Edinburgh Ph.D. candidate, supervised by Prof. Pasquale Minervini and Prof. Mirella Lapata	2023/09 - now
Harbin Institute of Technology, Shenzhen M.Sc. Computer Science, supervised by Prof. Xiaolong Wang and Prof. Baotian Hu	2020/09 - 2022/12
Hefei University of Technology B.Eng. Computer Science and Technology	2016/09 - 2020/06

EXPERIENCE

Microsoft Research Cambridge, United Kingdom <i>Research Intern</i> – GUI Grounding and Multimodal Reasoning.	2025/05 - 2025/08
Xiaomi AI Lab, Beijing China <i>Full-Time NLP Engineer</i> – LLMs Pre-Training and Evaluation.	2023/03 - 2023/08

PUBLICATIONS

- MMLongBench: Benchmarking Long-Context Vision-Language Models Effectively and Thoroughly
Zhaowei Wang, Wenhao Yu, Xiyu Ren, Jipeng Zhang, **Yu Zhao**, Rohit Saxena, Liang Cheng, Ginny Wong, Simon See, Pasquale Minervini, Yangqiu Song, Mark Steedman; **NeurIPS 2025**, **Spotlight**
- Noiser: Bounded Input Perturbations for Attributing Large Language Models
Mohammad Reza Ghasemi Madani, Aryo Pradipta Gema, Gabriele Sarti, **Yu Zhao**, Pasquale Minervini, Andrea Passerini; **COLM 2025**
- Steering Knowledge Selection Behaviours in LLMs via SAE-Based Representation Engineering
Yu Zhao, Alessio Devoto, Giwon Hong, Xiaotang Du, Aryo Pradipta Gema, Hongru Wang, Xuanli He, Kam-Fai Wong, Pasquale Minervini; **NAACL 2025** main conference, **Oral**
- Are We Done with MMLU?
Aryo Pradipta Gema, Joshua Ong Jun Leang, Giwon Hong, Alessio Devoto, Alberto Carlo Maria Mancino, Rohit Saxena, Xuanli He, **Yu Zhao**, Xiaotang Du, Mohammad Reza Ghasemi Madani, Claire Barale, Robert McHardy, Joshua Harris, Jean Kaddour, Emile van Krieken, Pasquale Minervini; **NAACL 2025** main
- Structured Packing in LLM Training Improves Long Context Utilization
Konrad Staniszewski*, Szymon Tworkowski*, Sebastian Jaszczur, **Yu Zhao**, Henryk Michalewski, Łukasz Kuciński, Piotr Miłoś; **AAAI 2025**, **Oral**
- A Simple and Effective L_2 Norm-Based Strategy for KV Cache Compression
Alessio Devoto*, **Yu Zhao***, Simone Scardapane, Pasquale Minervini (* denotes equal contribution)
EMNLP 2024 main conference, **Oral**
- Analysing The Impact of Sequence Composition on Language Model Pre-Training
Yu Zhao, Yuanbin Qu, Konrad Staniszewski, Szymon Tworkowski, Wei Liu, Piotr Miłoś, Yuxiang Wu, Pasquale Minervini; **ACL 2024** main conference, **Oral**
- Analysing the Residual Stream of Language Models Under Knowledge Conflicts
Yu Zhao, Xiaotang Du, Giwon Hong, Aryo Pradipta Gema, Alessio Devoto, Hongru Wang, Xuanli He, Kam-Fai Wong, Pasquale Minervini; **MINT @ NeurIPS 2024**
- An Efficient Memory-Augmented Transformer for Knowledge-Intensive NLP Tasks
Yuxiang Wu, **Yu Zhao**, Baotian Hu, Pasquale Minervini, Pontus Stenetorp, Sebastian Riedel
EMNLP 2022 main conference, **Best Poster Award @ NeurIPS ENLSP**

10. Medical Dialogue Response Generation with Pivotal Information Recalling
Yu Zhao*, Yunxin Li*, Yuxiang Wu, Baotian Hu, Qingcai Chen, Xiaolong Wang, Yuxin Ding, Min Zhang
KDD 2022
11. MSDF: A General Open-Domain Multi-Skill Dialog Framework
Yu Zhao*, Xinshuo Hu*, Yunxin Li, Baotian Hu, Dongfang Li, Sichao Chen, Xiaolong Wang; **NLPCC 2021**

PREPRINT

1. Learning GUI Grounding with Spatial Reasoning from Visual Feedback
Yu Zhao, Wei-Ning Chen, Huseyin Inan, Samuel Kessler, Lu Wang, Lukas Wutschitz, Fangkai Yang, Chaoyun Zhang, Pasquale Minervini, Saravan Rajmohan, Robert Sim, *Preprint 2025*
2. Q-Filters: Leveraging QK Geometry for Efficient KV Cache Compression
Nathan Godey, Alessio Devoto, **Yu Zhao**, Simone Scardapane, Pasquale Minervini, Éric de la Clergerie, Benoît Sagot, *Preprint 2025*
3. The Hallucinations Leaderboard – An Open Effort to Measure Hallucinations in Large Language Models
Giwon Hong*, Aryo Pradipta Gema*, Rohit Saxena*, Xiaotang Du*, Ping Nie*, **Yu Zhao***, Laura Perez-Beltrachini, Max Ryabinin, Xuanli He, Pasquale Minervini*, *Preprint 2024*

AWARDS AND SCHOLARSHIPS

The University of Edinburgh, CDT in NLP, 4-year PhD Scholarship	2023
Best Poster Award @ NeurIPS ENLSP 2022. Award: 750.00 USD	2022
Baidu Inc. Expert of Language Understanding and Generation Evaluation	2021
The First Prize Scholarship, Harbin Institute of Technology	2021
The First Prize Scholarship, Hefei University of Technology	2017

COMPETITIONS

ICASSP 2023 SPGC Shared Task: Title Generation with Limited Resource Won the 1st prize (1/47)	2023
2022 Language and Intelligence Challenge: Knowledge Grounded Dialogue Won the 3rd prize (4/610) in human evaluation and 1st place in automatic evaluation	2022
2021 Language and Intelligence Challenge: Multi-Skilled Dialogue Won the 3rd prize (6/750) in human evaluation and 8th place in automatic evaluation.	2021

SKILLS AND TOOLS

Programming Languages	Python, C/C++, Bash, SQL
Deep Learning Frameworks	PyTorch, JAX, DeepSpeed, verl
Development Tools	Docker, Scrapy, Django, AzureML, Kubernetes
Natural Languages	Mandarin, Eastern Min, English

SERVICES

Program Committee Member/Conference Reviewer EMNLP 2022, 2023; ACL 2023; CoNLL 2023, 2024, 2025; ACL SRW 2024; ICLR 2025, 2026; ARR (2024-now); AACL 2025	
Teaching Assistant Foundations of Natural Language Processing @ University of Edinburgh	2025
Machine Learning Practical @ University of Edinburgh	2025