Jie He Ph.D.

j.he@ed.ac.uk
https://probe2.github.io/

Education

School of Informatics, University of Edinburgh — Edinburgh, UK
Ph.D. in ILCC — 2022 - 2025 *(expected)*

- Advisor: Prof. Jeff Pan
- Research area: RAG-based Large Language Model, Knolwedge Distillation, Commonsense Reasoning, Robustness.

Department of Computer Science, Tianjin University — Tianjin, China
M.S. in Computer Science — 2019 - 2022

- Advisor: Prof. Deyi Xiong
- Research area: Commonsense Reasoning.

Department of Software Engineering, Shandong University — Shandong, China
B.E. of Engineering, Software Engineering — 2015 - 2019

Research Interests

My primary research interests lie in developing Language Technology for All, focusing on natural language understanding and generation. I am particularly interested in knowledge retrieval-based NLP systems. Recently, I have expanded my focus to include the analysis and application of large language models (LLMs) [7, 8, 11, 12], with a specific emphasis on robustness and multi-modal capabilities. Commonsense involves both the construction of commonsense datasets [1, 13] and methods incorporating knowledge into NLP models [4, 5, 9]. Interpretability involves diagnosing and analyzing the errors of existing models [2] and giving the model the ability to explain its predictions. I am also interested in discourse phenomena [3].

Employments

MSR, Redmond, US | Research intern — 2024.07 - 2024.10

- Advisor: Pei Zhou, Longqi Yang, Jennifer Neville
- Tool learning.

Apple AI/ML, Seattle, US | Research intern — 2024.02 - 2024.05

- Advisor: Yiwen Sun, Benjamin Han
- Investigating to generate rationales for knowledge-based question answering without training.

Cardiff NLP, UK | Collaboration — 2022.10 - Present

- Advisor: Victor Gutierrez Basulto

Copenhagen NLP, Remote | Research intern — 2021.08 - 2022.08

- Advisor: Daniel Hershcovich
- Systematic analysis of the application of graph neural networks in neural machine translation systems. Used GCN to encode UD, AMR, and other semantic/syntax information to observe performance improvements in challenging machine translation test sets.

UCL, Remote | Research intern — 2021.09 - 2022.01

- Advisor: Pasquale Minervini
- Systematic research on the application of prompt learning in natural language explanation tasks.

Huawei Noah's Ark Lab, China | Collaboration — 2020.09 - 2021.07

- Advisor: Qun Liu
- Chinese text error correction. Investigated the error produced by GPT2's text generation and developed a set of annotation tools for annotation. Published a paper in ACL2021.

Scholarship & Awards

- **Apple Scholars** in AI/ML PhD fellowship 2024 - 2026
- **PhD Scholarship**, School of Informatics, University of Edinburgh 10/2022 - 09/2025

Publications

1. **Jie He***, Tao Wang*, Deyi Xiong, Qun Liu. The Box is in the Pen: Evaluating Commonsense Reasoning in Neural Machine Translation. *Findings of the Association for Computational Linguistics: EMNLP 2020*.
2. **Jie He***, Bo Peng*, Yi Liao, Qun Liu, Deyi Xiong. TGError: An Error-Annotated Dataset and Benchmark Tasks for Text Generation from Pretrained Language Models. *The 59th Annual Meeting of the Association for Computational Linguistics (ACL2021)*.
3. **Jie He**, Wanqiu Long, Deyi Xiong. Evaluate Discourse Cohesion in Pre-trained Language Models. *Proceedings of the 3rd Workshop on Computational Approaches to Discourse in COLING 2022*.
4. **Jie He**, Yu Fu. MetaXCR: Reinforcement-Based Meta-Transfer Learning for Cross-Lingual Commonsense Reasoning. *Proceedings of The 1st Transfer Learning for Natural Language Processing Workshop in Neurips 2022*.
5. **Jie He**, Simon Yu, Victor Gutierrez Basulto, Jeff Pan. BUCA: A Binary Classification Approach to Unsupervised Commonsense Question Answering. *The 61th Annual Meeting of the Association for Computational Linguistics (ACL 2023)*.
6. Simon Yu*, **Jie He***, Victor Gutierrez Basulto, Jeff Pan. Instances and Labels: Hierarchy-aware Joint Supervised Contrastive Learning for Hierarchical Multi-Label Text Classification. *Findings of the Association for Computational Linguistics: EMNLP 2023*.
7. Xiongtao Zhou*, **Jie He***, Yuhua Ke, Guangyao Zhu, Victor Gutierrez Basulto, Jeff Z. Pan. An Empirical Study on Parameter-Efficient Fine-Tuning for MultiModal Large Language Models. *Findings of the Association for Computational Linguistics: ACL 2024*.
8. Simon Yu*, **Jie He***, Pasquale Minervini, Jeff Pan. Evaluating the Adversarial Robustness of Retrieval-Based In-Context Learning for Large Language Models. *Conference On Language Modeling 2024 (COLM 2024)*.
9. **Jie He**, Simon Yu, Victor Gutierrez Basulto, Jeff Pan. kNN-XICL: Cross-lingual In-Context Learning with Nearest Neighbor Inference. *In submission*.
10. **Jie He***, Yijun Yang*, Wanqiu Long, Victor Gutierrez Basulto, Jeff Z. Pan. Exploring Knowledge Graph to Text Generation with Large Language Models: Techniques, Challenges, and Innovations. *In submission*.
11. **Jie He**, Wendi Zhou, Xiang Lorraine Li, Jeff Z. Pan. Similarity-Based Domain Adaptation with LLMs for Cross-Domain Sentiment Classification. *In submission*.
12. Xiongtao Zhou*, **Jie He***, Lanyu Chen, jingyu li, Haojing Chen, Victor Gutierrez Basulto, Jeff Z. Pan, Hanjie Chen MiCEval: Unveiling Multimodal Chain of Thought's Quality via Image Description and Reasoning Steps *In submission*.
13. Yijun YANG, **Jie He**, Pinzhen Chen, Victor Gutierrez Basulto, Jeff Z. Pan. UniArk: A Holistic Approach to Unbiased and Consistent Factual Knowledge Extraction. *The North American Chapter of the Association for Computational Linguistics (NAACL 2024)*.
14. Yu Fu*, **Jie He***, Yifang Yang, Qun Liu, Deyi Xiong. Meta-RTL: Reinforcement-Based Meta-Transfer Learning for Low-Resource Commonsense Reasoning. Preprint.

Teaching

Tutor: Knowledge Graph, University of Edinburgh (Fall 2023)

Academic Services

Reviewers: *AACL (2022, 2023)*, *SIGDIAL (2023)*, *EMNLP (2023)*, *EACL (2024)*, *NAACL (2024)*, *ACL (2024)*, *ARR (2023.10 - now)*, *COLM (2024)*, *ICLR (2025)*, *Coling (2025)*

Area Chair: *EMNLP (2024)*, *ARR (2024.06 - now)*

Skills

Programming: `Python`, `C/C++`, `Java`, `MATLAB`, `LaTeX`.

Languages: Chinese (native), English (fluent).

Toolkits and Frameworks: `PyTorch`, `Transformers`, `Numpy`, `Pandas`, `Huggingface`, `AllenNLP`.