Weixuan Wang

✉ weixuan.wang@ed.ac.uk | ⌂ https://weixuan-wang123.github.io/

⚥ https://scholar.google.com/citations?user=qAduuoUAAAAJ&hl=en&oi=ao

Career Objectives

Seeking Positions in Research of Multilingual Large language Models and Commercial-viable NLP Products and Services

Capability

- **7 years** solid experience in the research and development in **Multilingual NLP** with **9** papers published in peer-reviewed conferences and journals and **2** patents
- Hands-on experience in developing end-to-end multilingual NLP commercial systems, which has been successfully deployed in AI service

Education

Informatics, University of Edinburgh

3-year doctoral program at the Institute for Language, Cognition and Computation (ILCC)

Supervised by Dr. Alexandra Birch and Dr. Barry Haddow   *Sep. 2023 – Now*

School of Software Technology, Dalian University of Technology

Master in Neural Language Processing,

GPA:3.42 average score:85.17   *Sep. 2017 – June. 2020*

School of Software Technology, Dalian University of Technology

Bachelor of Engineering in Software Engineering

GPA:3.52 average score:85.2   *Sep. 2013 – June. 2017*

Working Experience

Tutor, University of Edinburgh

Worked as a tutor for course Natural Language Understanding, Generation, and Machine Translation (2023-2024)   *Feb. 2024 – May. 2024*

AI Researcher, IT Innovation and Research Center (IIRC), Huawei Technologies Co., LTD

Worked on simultaneous speech translation, and large commercial machine translation platform, and multilingual LLM.   *Apr. 2023 – Sep. 2023*

AI Researcher, AI Application Research Center (AARC), Huawei Technologies Co., LTD

Worked on Knowledge-enhanced NMT, domain-specific NMT, and Machine Translation Data Processing Pipeline.   *June 2020 – Apr. 2023*

Projects

- ***Large Commercial Neural Machine Translation***   *Jan. 2023 – Sep. 2023*

  Responsibility: developing commercial NMT with LLM, designing the multilingual training data proportion

  Achievements: Performance based on PanGu LLM on par with existing end-to-end NMT model in 10+ languages

- ***End-to-end Simultaneous Speech Translation*** *July 2022 - May 2023*

  Responsibility: designing and evaluating the first end-to-end simultaneous translation system for Huawei meeting

  Achievements: The end-to-end simultaneous translation system has lower latency and higher performance than the cascaded translation system

- **Cascade Simultaneous Speech Translation** *June 2021 – July 2022*

  Responsibility: designing and developing cascade simultaneous translation system for Huawei meeting

  Achievements: SST end-to-end BLEU scores (EN<-> ZH) on par with existing commercial systems; The service successfully supported 110 + online conference sessions and 30 + onsite SST translation service for 2021/2022 Huawei STW event

- **Knowledge-enhanced NMT** *Sep. 2020 – May 2021*

  Responsibility: designing, implementing, and evaluating knowledge-enhance NMT algorithms based on topic modelling and homographic representation learning (HDR)

  Achievements: BLEU enhancement on the vanilla transformer Topic-enhanced model (+1.57 EN->DE), HDR (+2.3 in EN->RU) with algorithm deployed in related Huawei machine translation service; 2 peer-reviewed conference papers

- **Domain-specific NMT** *June 2020 – Feb. 2021*

  Responsibility: developing ICT domain-specific machine translation model leveraging domain dictionaries

  Achievements: BLEU increase +0.7 on SOTA EN->ZH working model

- **Machine Translation Data Processing Pipeline** *July 2019 – July 2020*

  Responsibility: lead developing corpus processing pipeline consisting of corpus cleaning, tokenize, data enhancement, and other functions like visualization

  Achievements: Huawei machine translation data processing pipeline V1.0

Selected Publications

- **Wang, W.**, Haddow, B., & Birch, A. (2024). Retrieval-augmented Multilingual Knowledge Editing. In Proceedings of 62nd Annual Meeting of the Association for Computational Linguistics, (ACL), 2024.

- **Wang, W.**, Haddow, B., Birch, A., & Peng, W. (2023). Assessing the reliability of large language model knowledge. In Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, (NAACL, oral).

- Cao, H., Zhao, T., **Wang, W.**, & Peng, W. (2023). Bilingual word embedding fusion for robust unsupervised bilingual lexicon induction. Information Fusion, 97, 101818.

- Cheng, H., Zhang, M., **Wang, W.**, Li, L., Liu, Q., & Zhang, Z. (2023). Evaluating the Efficacy of Length-Controllable Machine Translation. arXiv preprint arXiv:2305.02300.

- **Wang, W.**, Peng, W., & Liu, Q. (2023). Learning Homographic Disambiguation Representation for Neural Machine Translation. arXiv preprint arXiv:2304.05860.

- **Wang, W.**, Lee, C. M., Liu, J., Colakoglu, T., & Peng, W. (2023). An empirical study of cyclical learning rate on neural machine translation. Natural Language Engineering, 29(2), 316-336.

- **Wang, W.**, Meng, X., Yan, S., Tian, Y., & Peng, W. (2022, December). Huawei BabelTar NMT at WMT22 Biomedical Translation Task: How we further improve domain-specific NMT. In Proceedings of the Seventh Conference on Machine Translation (WMT) (pp. 930-935).

- Feng, Z., Cao, H., Zhao, T., **Wang, W.**, Peng, W. (2022). Cross-lingual Feature Extraction from Monnolingual Corpora for Low-resource Unsupervised Bilingual Lexicon Induction. In Proceedings of the 29th International Conference on Computational Linguistics.

- **Wang, W.**, Peng, W., Huang, C. H., & Wang, H. (2022). Positively transitioned sentiment dialogue corpus for developing emotion-affective open-domain chatbots. arXiv preprint arXiv:2208.04565.

- **Wang, W.**, Peng, W., Zhang, M., & Liu, Q. (2021, November). Neural Machine Translation with Heterogeneous Topic Knowledge Embeddings. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (pp. 3197-3202)*.

- **Wang, W.**, Peng, W., Meng, X., & Liu, Q. 2021a. Huawei AARC's submissions to the WMT21 Biomedical Translation Task: Domain Adaption from a Practical Perspective. In *Proceedings of the Sixth Conference on Machine Translation:*

Shared Task Papers

➢ **Wang, W.**, Cai, X., Huang, C. H., Wang, H., Lu, H., Liu, X., & Peng, W. (2021). Emily: Developing An Emotion-affective Open-Domain Chatbot with Knowledge Graph-based Persona. arXiv preprint arXiv:2109.08875.

➢ Liu, X., **Wang, W.**, Liang, W., & Li, Y. (2020). Speed up the training of neural machine translation. *Neural Processing Letters, 51(1), 231-249.*

➢ Liu, X., & **Wang, W.** (2019, December). Multiway Attention for Neural Machine Translation. In *Proceedings of the 2019 2nd International Conference on Algorithms, Computing and Artificial Intelligence (pp. 265-270).*

Patents

➢ 梁文新、**王伟璇**、刘馨月、张宪超：《基于双向循环神经网络的神经机器翻译方法和系统》，2018，专利申请号：CN201811358954.4

➢ **王伟璇**、WEI PENG：《一种利用外部同形异义信息解决机器翻译歧义问题的方法》，2022，专利申请号：TBA

Challenge

● **WMT 2022 biomedical shared tasks**

Responsibility: lead model training and finetuning for WMT 22 biomedical machine translation shared task

Achievements: four 1st place for 5 language directions (English<->Italian, English<->Portuguese, Chinese->English), 1 shared-task publication

● **WMT 2021 biomedical shared tasks**

Responsibility: lead model training and finetuning for WMT 21 biomedical machine translation shared task

Achievements: four 1st place for 4 language directions (English->French, English<->Italian, Chinese->English) and three runner-up for English->German, English->Chinese, French->English, 1 shared-task publication

● **2016 Mathematical Contest In Modeling (MCM)**

Responsibility: select the clustering model to solve "Refugee migration" problem and design the methods to improve the generalization ability and scalability of the model.

Achievements: first place in the contest.

Professional Awards

Excellent Team Award for Simultaneous Speech Translation, AARC, Huawei Technologies, Co., Ltd. (2021)

Technology Breakthrough Award for Machine Translation, AARC, Huawei Technologies, Co., Ltd. (2021)

Future Star Award for outstanding achievements for neural machine translation initiatives, Huawei Technologies, Co., Ltd. (2021)

National Scholarship for Postgraduate students, Dalian University of Technology (2017-2020)

National Encouragement scholarship, Dalian University of Technology (2016)

The Second Prize Scholarship, Dalian University of Technology (2016)

National Encouragement scholarship, Dalian University of Technology (2015)

The Second Prize Scholarship, Dalian University of Technology (2015)