Yifan Wang

(+49) 15227932317 | yifwang@lst.uni-saarland.de | ewanwong.github.io

EDUCATION

Saarland University   Saarbrücken, Germany

- *M.Sc.* in Language Science and Technology | **GPA: 1.1/1.0**   2021.10-expected 2023.12
- **Master Thesis:** Mitigating Stereotypical Bias in Generative Language Models via Prefix-Tuning (Supervisor: Prof. Dr. Vera Demberg & Dongqi Pu)
- **Relevant Coursework:** Computational Linguistics, Statistics with R, Neural Network: Implementation and Application, Machine Translation, Pre-trained LMs in Computational Semantics, Learning from Explanations, Cross-lingual Learning etc.

Shanghai Jiaotong University   Shanghai, China

- *B.A.* in Germanistics | **GPA: 89/100**   2016.9-2020.6
- **Bachelor Thesis:** Comparison of the Perspectives of the German and Chinese Media on COVID-19 on the Basis of a Quantitative Analysis of the Corpus (Supervisor: Prof. Likun Fan)
- **Relevant Coursework:** Introduction to Linguistics, German Lexicology, Calculus, Linear Algebra, Discrete Mathematics, C++, Natural Language Processing, Data Mining, Data Structure and Algorithms etc.

Ruprecht-Karls-University Heidelberg   Heidelberg, Germany

- *Exchange student* in Germanistics in the Cultural Comparison   2018.10-2019.3

PUBLICATIONS

- Dongqi Pu, **Yifan Wang,** and Vera Demberg." Incorporating Distributions of Discourse Structure for Long Document Abstractive Summarization." In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL 2023).*

RESEARCH EXPERIENCE

Saarland University   Saarbrücken, Germany

Research Assistant   2022.11-present

- Incorporated discourse structure to automatic summarization, faithfulness evaluation and relevant domains. Achieved improved performance over SOTA across diverse datasets.
- Implemented benchmark methods, performed model evaluation and analysis experiments.
- Conducted large scale literature reviews in various fields including automatic summarization, faithfulness evaluation, controllable text generation, style transfer and parameter-efficient fine-tuning.

Saarland University, Software Project   Saarbrücken, Germany

Student Researcher   2022.4-2022.10

- Collected the first public Cantonese monolingual dataset at the scale of one million sentences.
- Trained an transformer-based unsupervised Mandarin-Cantonese machine translation system.
- Exploited the similarity and divergence between dialects by information sharing on the word embedding level. Studied the effect of different model architectures on the performance in the unsupervised machine translation setting.

Shanghai Jiaotong University, Student Research Project   Shanghai, China
Research Assistant   2018.4-2018-7
- Studied the sequence of German grammatical structure acquisition of adult Chinese native speakers. Demonstrated the development of second language follows Pienemann's Processability Theory.
- Invited participants and conducted experiments.

SKILLS

Programming: Python, R, C++, familiar with NLP tools including Huggingface Transformers, PyTorch, NLTK, SpaCy, etc.
Languages: Mandarin Chinese (Native), English (Fluent), German (Intermediate).
Tools: LaTex, HTML, Markdown, Linux, Git, etc.