Danni Liu

✉ liu.danni@outlook.com   🎓 Google Scholar   in LinkedIn: danni-liu

Education

02/2020 – now	**Ph.D. Computer Science** Karlsruhe Institute of Technology, Germany (07/2022 – now) Maastricht University, Netherlands (02/2020 – 06/2022) *Advisor*: Jan Niehues *Research focus*: Multilingual transfer learning for low-resource settings
02/2018 – 01/2020	**M.Sc. Artificial Intelligence** (cum laude) Maastricht University, Netherlands *Thesis*: Low-Latency End-to-End Speech Recognition with Enhanced Readability
09/2014 – 08/2017	**B.Sc. Data Science and Knowledge Engineering** (cum laude) Maastricht University, Netherlands *Thesis*: Short Text Topic Modeling Enhanced by Collocations and Word Embeddings
09/2011 – 08/2014	**B.Sc. International Business Administration** (cum laude) Erasmus University Rotterdam, Netherlands *Thesis*: Relationship Between Board Composition and R&D Intensity

Professional Experience

07/2021 – 01/2022	**Facebook / Meta AI**, *Research Intern*, Remote. For 10/2021 – 01/2022 employed as contingent worker via Blue Lynx *Project:* Low-latency TTS for speech-to-speech translation, published in Interspeech [6]
03/2019 – 08/2019	**Nuance Communications (now Microsoft)**, *Research Intern*, Aachen, Germany.
08/2015 – 01/2019	**Open Line**, *Trainee → Business Intelligence Consultant*, Maastricht-Airport, Netherlands.

Talks

09/2024	**University of Toronto & OntarioTech University**, "Retrieval-Augmented Approaches for Low-Resource Conditions in Speech Translation" [slides] and
07/2024	"Learning from Less: Machine Translation and Summarization in Low-Resource Settings" [slides]
12/2022	**IWSLT Lecture Series**, "Learning Shared Representations for Translation in Low-Resource Conditions" [video] [slides]
02/2022	**Meta AI**, "Studying Shared Representations in NMT through Zero-Shot Translation" [slides]

Teaching & Mentoring

2020 – now	**Supervision of Bachelor/Master's Theses & Internships** Leading to publications [14, 12, 10, 8, 5, 4]
2022 – now	**Tutor**, Karlsruhe Institute of Technology, Germany *Courses:* Natural Language Processing, Machine Translation (guest lecturer), multiple seminars
2019 – 2021	**Teaching Assistant**, Maastricht University, Netherlands *Courses:* Advanced Natural Language Processing, Probability and Statistics

Teaching & Mentoring (continued)

2012 – 2013   **Teaching Assistant**, Erasmus University Rotterdam, Netherlands
Courses: Skills 1: IT, mentoring course for first-year undergraduates

Publications

14. Siqi Li*, **Danni Liu***, and Jan Niehues. 2024. Optimizing Rare Word Accuracy in Direct Speech Translation with a Retrieval-and-Demonstration Approach.   *EMNLP*

13. Tu Anh Dinh, Carlos Mullov, Leonard Bärmann, Zhaolin Li, **Danni Liu**, Simon Reiß, Jueun Lee, Nathan Lerzer, Jianfeng Gao, Fabian Peller-Konrad, Alexander Waibel, Tamim Asfour, Michael Beigl, Rainer Stiefelhagen, Carsten Dachsbacher, Klemens Böhm, Jan Niehues. 2024. SciEx: Benchmarking Large Language Models on Scientific Exams with Human Expert Grading and Automatic Grading.   *EMNLP*

12. Lukas Hilgert, **Danni Liu**, and Jan Niehues. 2024. Evaluating and Training Long-Context Large Language Models for Question Answering on Scientific Papers.   *CustomNLP4U Workshop @ EMNLP*

11. **Danni Liu** and Jan Niehues. 2024. Recent Highlights in Multilingual and Multimodal Speech Translation. [slides]   *IWSLT*

10. Vladimir Solovyev, **Danni Liu**, and Jan Niehues. 2024. Language-Independent Representations Improve Zero-Shot Summarization.   *NAACL*

9. **Danni Liu** and Jan Niehues. 2024. How Transferable are Attribute Controllers on Pretrained Multilingual Translation Models?. [slides] [poster]   *EACL*

8. Yunus Demirag, **Danni Liu**, and Jan Niehues. 2024. Benchmarking Diffusion Models for Machine Translation.   *EACL Student Research Workshop*

7. **Danni Liu** and Jan Niehues. 2022. Learning an Artificial Language for Knowledge-Sharing in Multilingual Translation. [video] [slides]   *WMT*

6. **Danni Liu**, Changhan Wang, Hongyu Gong, Xutai Ma, Yun Tang, and Juan Pino. 2022. From Start to Finish: Latency Reduction Strategies for Incremental Speech Synthesis in Simultaneous Speech-to-Speech Translation. [poster]   *Interspeech*

5. Tu Anh Dinh, **Danni Liu**, and Jan Niehues. 2022. Tackling Data Scarcity in Speech Translation using Zero-shot Multilingual Machine Translation Techniques.   *ICASSP*

4. Sai Koneru, **Danni Liu**, and Jan Niehues. 2021. Unsupervised Machine Translation On Dravidian Languages.   *DravidianLangTech Workshop @ EACL*

3. **Danni Liu**, Jan Niehues, James Cross, Francisco Guzmán, and Xian Li. 2021. Improving Zero-shot Translation by Disentangling Positional Information. [video] [slides]   *ACL*

2. **Danni Liu**, Gerasimos Spanakis, and Jan Niehues. 2020. Low-Latency Sequence-to-Sequence Speech Recognition and Translation by Partial Hypothesis Selection. [video] [slides]   *Interspeech*

1. **Danni Liu**, Jan Niehues, and Gerasimos Spanakis. 2020. Adapting End-to-End Speech Recognition for Readable Subtitles. [video] [slides]   *IWSLT*

Shared Tasks and Demonstrations (Selected)

S8. Christian Huber, Tu Anh Dinh, Carlos Mullov, Ngoc-Quan Pham, Thai Binh Nguyen, Fabian Retkowski, Stefan Constantin, Enes Ugan, **Danni Liu**, Zhaolin Li, Sai Koneru, Jan Niehues, and Alexander Waibel. 2023. End-to-End Evaluation for Low-Latency Simultaneous Speech Translation.   *EMNLP Demo*

S7. **Danni Liu**, Thai Binh Nguyen, Sai Koneru, Enes Yavuz Ugan, Ngoc-Quan Pham, Tuan Nam Nguyen, Tu Anh Dinh, Carlos Mullov, Alexander Waibel, and Jan Niehues. 2023. KIT's Multilingual Speech Translation System for IWSLT 2023. [poster] 🏆 **Winning system in multilingual track (constrained)** *IWSLT*

S6. Peter Polák, **Danni Liu**, Ngoc-Quan Pham, Jan Niehues, Alexander Waibel, and Ondřej Bojar. 2023. Towards Efficient Simultaneous Speech Translation: CUNI-KIT System for Simultaneous Track at IWSLT 2023. *IWSLT*

S5. Abteen Ebrahimi, Manuel Mager, Adam Wiemerslage, Pavel Denisov, Arturo Oncevay, **Danni Liu**, Sai Koneru, Enes Yavuz Ugan, Zhaolin Li, Jan Niehues, Monica Romero, Ivan G Torre, Tanel Alumäe, Jiaming Kong, Sergey Polezhaev, Yury Belousov, Wei-Rui Chen, Peter Sullivan, Ife Adebara, Bashar Talafha, Alcides Alcoba Inciarte, Muhammad Abdul-Mageed, Luis Chiruzzo, Rolando Coto-Solano, Hilaria Cruz, Sofía Flores-Solórzano, Aldo Andrés Alvarez López, Ivan Meza-Ruiz, John E Ortega, Alexis Palmer, Rodolfo Joel Zevallos Salazar, Kristine Stenzel, Thang Vu, and Katharina Kann. 2022. Findings of the Second AmericasNLP Competition on Speech-to-Text Translation. 🏆 **Winning system in speech translation and machine translation tracks** *NeurIPS 2022 Competition Track*

S4. Ngoc-Quan Pham, Tuan Nam Nguyen, Thai-Binh Nguyen, **Danni Liu**, Carlos Mullov, Jan Niehues, and Alexander Waibel. 2022. Effective combination of pretrained models - KIT@IWSLT 2022 *IWSLT*

S3. Peter Polák, Ngoc-Quan Pham, Tuan Nam Nguyen, **Danni Liu**, Carlos Mullov, Jan Niehues, Ondřej Bojar, and Alexander Waibel. 2022. CUNI-KIT System for Simultaneous Speech Translation Task at IWSLT 2022. 🏆 **Winning system in simultaneous translation track** *IWSLT*

S2. **Danni Liu** and Jan Niehues. 2021. Maastricht University's Large-Scale Multilingual Machine Translation System for WMT 2021. [poster] 🏛 **Best performing system from university participants** *WMT*

S1. **Danni Liu** and Jan Niehues. 2021. Maastricht University's Multilingual Speech Translation System for IWSLT 2021. [video] [slides] *IWSLT*

Service

Reviewer	ACL Rolling Review (2024, 2023); ACL (2024), EMNLP (2024, 2023), EACL (2023), COLING (2024, 2025), WMT (2022), ICASSP (2024), Interspeech (2024), EACL-SRW (2024), WiNLP (2024), DravidianLangTech (2024).
Volunteer	ACL 2020