

# Rui Wang

Tenure-track Researcher  
National Institute of Information  
and Communications Technology (NICT)  
Kyoto, Japan.

Phone: (+81) 0774-98-6986  
Email: wangrui@nict.go.jp  
Homepage: <https://wangruinlp.github.io>

## Experience

2019- Tenure-track Researcher, Advanced Translation Technology Laboratory, NICT, Kyoto, Japan.

2016-2019: Postdoctoral Researcher, Advanced Translation Technology Laboratory, NICT, Kyoto, Japan.

2012-2016: Ph.D. in Computer Science, Shanghai Jiao Tong University, working with Bao-Liang Lu and Hai Zhao, Shanghai, China.

—2014: Joint Ph.D., CNRS-University Lyon1, working with Sabine Ploux, Lyon, France.

—2013: Internship Research Fellow, NICT, working with Masao Utiyama, Isao Goto, Andrew Finch, and Eiichiro Sumita, Kyoto, Japan.

2009-2012: Master in Computer Science, University of Chinese Academy of Sciences, working with Xiao li and Xi Zhou, Beijing & Urumqi, China.

2005-2009: Bachelor in Computer Science, Harbin Institute of Technology, Harbin, China.

## Grants and Fundings

2019-2021: PI of NICT tenure-track startup fund: "Toward Intelligent MT" (8,000,000JPY).

2019-2022 (19K20354): PI of JSPS grant-in-aid for early-career scientists (national fund of Japan): "Un-supervised NMT in Universal Scenarios" (4,180,000JPY).

## Academic Services

Area co-chair of CCL-2019 and CCL-2018.

Organization co-chair of PACLIC-29 and YCCL-2012.

PC member of ACL, EMNLP, NAACL, COLING, IJCNLP, AACL, IJCNN, AACL, IJCAI, etc.

Reviewer of JSPS national fund, Computational Linguistics (standing reviewer), Transactions of the Association for Computational Linguistics (standing reviewer), IEEE/ACM TASLP, TALLIP, Computer Speech and Language (outstanding reviewer), Neurocomputing, Neural Processing Letters, Language Resources and Evaluation, etc.

## Internship Hosting and Supervising

2020 @NICT: Sufeng Duan (Ph.D. student of SJTU) and Xinlu Zhang (Ph.D. student of CAS).

2019 @NICT: Zhuosheng Zhang (master student of SJTU), Zuchao Li (Ph.D. student of SJTU), Chaoqun Duan (Ph.D. student of HIT), and Fengshun Xiao (master student of SJTU).

2018 @NICT: Haipeng Sun (Ph.D. student of HIT), Shu Jiang (Ph.D. student of SJTU), and Mingming Yang (Ph.D. student of HIT/Soochow University).

2017 @NICT: Kehai Chen (Ph.D. of HIT->researcher of NICT) and Zhisong Zhang (master of SJTU->CMU).

## Selected Publications ("\*" denotes the corresponding author)

### *Shared Tasks*

CONLL-2019 (Corresponding author): 1st in the DM sub-task and the 2nd overall.

WMT-2019 (Corresponding author): 1st in the only unsupervised MT task (German-Czech) by BLEU and human evaluation

WMT-2018 (Second author): 1st places in four tasks (English<->Estonian and English<->Finnish) and tied 1st in other three tasks (Chinese->English and English<->Turkish).

WAT-2018 (First author): 1st places in Myanmar (Burmese) <- English by BLEU.

CWMT-2018 (First author): Tied 1st places in English<->Chinese by significance test.

### *Journal Articles*

*Unsupervised Neural Machine Translation with Cross-lingual Language Representation Agreement*

Haipeng Sun, **Rui Wang** (as the supervisor), Kehai Chen, Masao Utiyama, Eiichiro Sumita, and Tiejun Zhao

IEEE/ACM Transactions on Audio, Speech, and Language Processing (TASLP), accepted

*Neural Machine Translation with Sentence-level Topic Context*

Kehai Chen, **Rui Wang** (as the supervisor), Masao Utiyama, Eiichiro Sumita, and Tiejun Zhao.

IEEE/ACM Transactions on Audio, Speech, and Language Processing (TASLP), Volume 27, Issue 12, 2019.

*Sentence Selection and Weighting for Neural Machine Translation Domain Adaptation.*

**Rui Wang**, Masao Utiyama, Andrew Finch, Lemao Liu, Kehai Chen, and Eiichiro Sumita.

IEEE/ACM Transactions on Audio, Speech, and Language Processing (TASLP), Volume 26, Issue 10, 2018.

*Graph-based Bilingual Word Embedding for Statistical Machine Translation.*

**Rui Wang**, Hai Zhao\*, Bao-Liang Lu, Sabine Ploux\*, Masao Utiyama, and Eiichiro Sumita.

ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), Volume 17, Issue 4, 2018.

*A Neural Approach to Source Dependency-Based Context Model for Statistical Machine Translation.*

Kehai Chen, Tiejun Zhao, Muyun Yang, Lemao Liu\*, Akihiro Tamura, **Rui Wang**, Masao Utiyama, and Eiichiro Sumita.

IEEE/ACM Transactions on Audio, Speech, and Language Processing (TASLP), Volume 26, Issue 2, 2017.

*Converting Continuous-Space Language Models into N-gram Language Models with Efficient Bilingual Pruning for Statistical Machine Translation.*

**Rui Wang**, Masao Utiyama\*, Isao Goto, Eiichiro Sumita, Hai Zhao\*, and Bao-Liang Lu.

ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), Volume 15, Issue 3, 2016.

*Bilingual Continuous-Space Language Model Growing for Statistical Machine Translation.*

**Rui Wang**, Hai Zhao\*, Bao-Liang Lu, Masao Utiyama, and Eiichiro Sumita.

IEEE/ACM Transactions on Audio, Speech, and Language Processing (TASLP), Volume:23, Issue:7, 2015.

## Conference Proceedings

*Content Word Aware Neural Machine Translation*

Kehai Chen, **Rui Wang\***, Masao Utiyama, and Eiichiro Sumita

The 58th Annual Meeting of the Association for Computational Linguistics (ACL-2020), Seattle, USA

*Knowledge Distillation for Multilingual Unsupervised Neural Machine Translation*

Haipeng Sun, **Rui Wang\***, Kehai Chen, Masao Utiyama, Eiichiro Sumita, and Tiejun Zhao\*

The 58th Annual Meeting of the Association for Computational Linguistics (ACL-2020), Seattle, USA

*Regularized Context Gates on Transformer for Machine Translation*

Xintong Li, Lemao Liu, **Rui Wang**, Guoping Huang, and Max Meng

The 58th Annual Meeting of the Association for Computational Linguistics (ACL-2020), Seattle, USA

*Data-dependent Gaussian Prior Objective for Language Generation*

Zuchao Li, **Rui Wang\***, Kehai Chen, Masao Utiyama, Eiichiro Sumita, Zhuosheng Zhang, and Hai Zhao\*

International Conference on Learning Representations (ICLR-2020), Addis Ababa, Ethiopia

**Note: this is a full-score paper and a long-time talk presentation**

*Neural Machine Translation with Universal Visual Representation*

Zhuosheng Zhang, Kehai Chen, **Rui Wang\***, Masao Utiyama, Eiichiro Sumita, Zuchao Li, and Hai Zhao\*

International Conference on Learning Representations (ICLR-2020), Addis Ababa, Ethiopia

*Explicit Sentence Compression for Neural Machine Translation*

Zuchao Li, **Rui Wang\***, Kehai Chen, Masao Utiyama, Eiichiro Sumita, Zhuosheng Zhang, and Hai Zhao\*

Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI-2020), New York, USA

*SG-Net: Syntax-Guided Machine Reading Comprehension*

Zhuosheng Zhang, Yuwei Wu, Junru Zhou, Sufeng Duan, Hai Zhao\*, and **Rui Wang\***

Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI-2020), New York, USA

*Recurrent Positional Embedding for Neural Machine Translation*

Kehai Chen, **Rui Wang\***, Masao Utiyama, and Eiichiro Sumita

2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP2019), Hong Kong, China

*Unsupervised Bilingual Word Embedding Agreement for Unsupervised Neural Machine Translation*

Haipeng Sun, **Rui Wang\***, Kehai Chen, Masao Utiyama, Eiichiro Sumita, and Tiejun Zhao\*

The 57th Annual Meeting of the Association for Computational Linguistics (ACL-2019), Florence, Italy

*Neural Machine Translation with Reordering Embeddings*

Kehai Chen, **Rui Wang\***, Masao Utiyama, and Eiichiro Sumita

The 57th Annual Meeting of the Association for Computational Linguistics (ACL-2019), Florence, Italy

*Sentence-Level Agreement for Neural Machine Translation*

Mingming Yang, **Rui Wang\***, Kehai Chen, Masao Utiyama, Eiichiro Sumita, Min Zhang\*, and Tiejun Zhao

The 57th Annual Meeting of the Association for Computational Linguistics (ACL-2019), Florence, Italy

*Lattice-Based Transformer Encoder for Neural Machine Translation*Fengshun Xiao, Jiangtong Li, Hai Zhao, **Rui Wang**, and Kehai Chen

The 57th Annual Meeting of the Association for Computational Linguistics (ACL-2019), Florence, Italy

*A Survey of Domain Adaptation for Neural Machine Translation.*Chenhui Chu and **Rui Wang**

The 27th International Conference on Computational Linguistics (COLING-2018), Santa Fe, USA.

*Exploring Recombination for Efficient Decoding of Neural Machine Translation.*Zhisong Zhang, **Rui Wang\***, Masao Utiyama, Eiichiro Sumita, and Hai Zhao\*

2018 Conference on Empirical Methods in Natural Language Processing (EMNLP-2018), Brussels, Belgium.

*Dynamic Sentence Sampling for Efficient Training of Neural Machine Translation.***Rui Wang**, Masao Utiyama, and Eiichiro Sumita.

The 56th Annual Meeting of the Association for Computational Linguistics (ACL-2018), Melbourne, Australia.

*Syntax-Directed Attention for Neural Machine Translation*Kehai Chen, **Rui Wang\***, Masao Utiyama, Eiichiro Sumita, and Tiejun Zhao.

The Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-2018), New Orleans, USA.

*Instance Weighting for Neural Machine Translation Domain Adaptation.***Rui Wang**, Masao Utiyama, Lemao Liu, Kehai Chen, and Eiichiro Sumita.

Conference on Empirical Methods in Natural Language Processing (EMNLP-2017), Copenhagen, Denmark.

*Neural Machine Translation with Source Dependency Representation.*Kehai Chen, **Rui Wang\***, Masao Utiyama, Lemao Liu, Akihiro Tamura, Eiichiro Sumita, and Tiejun Zhao.

Conference on Empirical Methods in Natural Language Processing (EMNLP-2017), Copenhagen, Denmark.

*Sentence Embedding for Neural Machine Translation Domain Adaptation.***Rui Wang**, Andrew Finch, Masao Utiyama, and Eiichiro Sumita.

The 55th annual meeting of the Association for Computational Linguistics (ACL-2017), Vancouver, Canada.

*Connecting Phrase based Statistical Machine Translation Adaptation.***Rui Wang**, Hai Zhao\*, Bao-Liang Lu, Masao Utiyama\*, and Eiichiro Sumita.

The 26th International Conference on Computational Linguistics (COLING-2016), Osaka, Japan

*A Bilingual Graph-based Semantic Model for Statistical Machine Translation.***Rui Wang**, Hai Zhao\*, Sabine Ploux\*, Bao-Liang Lu, and Masao Utiyama.

25th International Joint Conference on Artificial Intelligence (IJCAI-16), New York, USA.

*Neural Network Based Bilingual Language Model Growing for Statistical Machine Translation.***Rui Wang**, Hai Zhao\*, Bao-Liang Lu, Masao Utiyama, and Eiichiro Sumita.

Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP-2014), Doha, Qatar.

*Converting Continuous-Space Language Models into N-gram Language Models for Statistical Machine Translation.***Rui Wang**, Masao Utiyama, Isao Goto, Eiichiro Sumita, Hai Zhao, and Bao-Liang Lu.

Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing (EMNLP-2013), Seattle, USA.

## ***Under Review***

*Towards More Diverse Input Representation for Neural Machine Translation*

Kehai Chen, **Rui Wang**<sup>\*</sup>, Masao Utiyama, Eiichiro Sumita, Tiejun Zhao, Muyun Yang, and Hai Zhao.  
IEEE/ACM Transactions on Audio, Speech, and Language Processing (TASLP), Major Revision

*Unsupervised Neural Machine Translation for Similar and Distant Language Pairs: An Empirical Study*

Haipeng Sun, **Rui Wang** (as the supervisor), Kehai Chen, Masao Utiyama, Eiichiro Sumita, and Tiejun Zhao

ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), Major Revision.

## ***Patent***

*Multi-domain Neural Machine Translation.*

Japanese Patent

**Rui Wang** and Masao Utiyama.

*Reordering Mechanism for Neural Machine Translation.*

Japanese Patent

Kehai Chen, **Rui Wang**, and Masao Utiyama.

Last updated: April 23, 2020  
<https://wangruinlp.github.io/>