WMT 2019

Fourth Conference on Machine Translation

Proceedings of the Conference

Volume 1: Research Papers

©2019 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-950737-27-7

Introduction

The Fourth Conference on Machine Translation (WMT 2019) took place on Thursday, August 1 and Friday, August 2, 2019 in Florence, Italy, immediately following the 57th Annual Meeting of the Association for Computational Linguistics (ACL 2019).

This is the fourth time WMT has been held as a conference. The first time WMT was held as a conference was at ACL 2016 in Berlin, Germany, the second time at EMNLP 2017 in Copenhagen, Denmark, and the third time at EMNLP 2028 in Brussels, Belgium. Prior to being a conference, WMT was held 10 times as a workshop. WMT was held for the first time at HLT-NAACL 2006 in New York City, USA. In the following years the Workshop on Statistical Machine Translation was held at ACL 2007 in Prague, Czech Republic, ACL 2008, Columbus, Ohio, USA, EACL 2009 in Athens, Greece, ACL 2010 in Uppsala, Sweden, EMNLP 2011 in Edinburgh, Scotland, NAACL 2012 in Montreal, Canada, ACL 2013 in Sofia, Bulgaria, ACL 2014 in Baltimore, USA, and EMNLP 2015 in Lisbon, Portugal.

The focus of our conference is to bring together researchers from the area of machine translation and invite selected research papers to be presented at the conference.

Prior to the conference, in addition to soliciting relevant papers for review and possible presentation, we conducted 8 shared tasks. This consisted of four translation tasks: Machine Translation of News, Biomedical Translation, Robust Machine Translation, and Similar Language Translation, two evaluation tasks: Metrics and Quality Estimation, as well as the Automatic Post-Editing and Parallel Corpus Filtering tasks.

The results of all shared tasks were announced at the conference, and these proceedings also include overview papers for the shared tasks, summarizing the results, as well as providing information about the data used and any procedures that were followed in conducting or scoring the tasks. In addition, there are short papers from each participating team that describe their underlying system in greater detail.

Like in previous years, we have received a far larger number of submissions than we could accept for presentation. WMT 2019 has received 48 full research paper submissions (not counting withdrawn submissions). In total, WMT 2019 featured 12 full research paper oral presentations and 102 shared task poster presentations.

The invited talk was given by Marine Carpuat from the University of Maryland, College Park, USA. It was titled “Semantic, Style & Other Data Divergences in Neural Machine Translation".

We would like to thank the members of the Program Committee for their timely reviews. We also would like to thank the participants of the shared task and all the other volunteers who helped with the evaluations.

Ondřej Bojar, Rajen Chatterjee, Christian Federmann, Mark Fishel, Yvette Graham, Barry Haddow, Matthias Huck, Antonio Jimeno Yepes, Philipp Koehn, André Martins, Christof Monz, Matteo Negri, Aurélie Névéol, Mariana Neves, Matt Post, Marco Turchi, and Karin Verspoor

Co-Organizers

Organizers:

Ondřej Bojar (Charles University in Prague)
Rajen Chatterjee (FBK)
Christian Federmann (MSR)
Mark Fishel (University of Tartu)
Yvette Graham (DCU)
Barry Haddow (University of Edinburgh)
Matthias Huck (LMU Munich)
Antonio Jimeno Yepes (IBM Research Australia)
Philipp Koehn (University of Edinburgh / Johns Hopkins University)
André Martins (Unbabel)
Christof Monz (University of Amsterdam)
Matteo Negri (FBK)
Aurélie Névéol (LIMSI, CNRS)
Mariana Neves (German Federal Institute for Risk Assessment)
Matt Post (Johns Hopkins University)
Marco Turchi (FBK)
Karin Verspoor (University of Melbourne)

Invited Speaker:

Marine Carpuat (University of Maryland, College Park)

Program Committee:

Tamer Alkhouli (RWTH Aachen University)
Antonios Anastasopoulos (Carnegie Mellon University)
Yuki Arase (Osaka University)
Mihael Arcan (INSIGHT, NUI Galway)
Duygu Ataman (Fondazione Bruno Kessler - University of Edinburgh)
Eleftherios Avramidis (German Research Center for Artificial Intelligence (DFKI))
Amittai Axelrod (Didi Chuxing)
Parnia Bahar (RWTH Aachen University)
Ankur Bapna (Google AI)
Petra Barancikova (Charles University in Prague, Faculty of Mathematics and Physics)
Joost Bastings (University of Amsterdam)
Rachel Bawden (University of Edinburgh)
Meriem Beloucif (University of Hamburg)
Graeme Blackwood (IBM Research AI)
Frédéric Blain (University of Sheffield)
Chris Brockett (Microsoft Research)
Bill Byrne (University of Cambridge)
Elena Cabrio (Université Côte d'Azur, Inria, CNRS, I3S, France)

Marine Carpuat (University of Maryland)
Francisco Casacuberta (Universitat Politècnica de València)
Sheila Castilho (Dublin City University)
Rajen Chatterjee (Apple Inc)
Boxing Chen (Alibaba)
Colin Cherry (Google)
Mara Chinea-Rios (Universitat Politècnica de València)
Chenhui Chu (Osaka University)
Ann Clifton (Spotify)
Marta R. Costa-jussà (Universitat Politècnica de Catalunya)
Josep Crego (SYSTRAN)
Raj Dabre (NICT)
Steve DeNeefe (SDL Research)
Michael Denkowski (Amazon)
Mattia A. Di Gangi (Fondazione Bruno Kessler)
Miguel Domingo (Universitat Politècnica de València)
Kevin Duh (Johns Hopkins University)
Marc Dymetman (Naver Labs Europe)
Hiroshi Echizen'ya (Hokkai-Gakuen University)
Sergey Edunov (Faceook AI Research)
Marcello Federico (Amazon AI)
Yang Feng (Institute of Computing Technology, Chinese Academy of Sciences)
Andrew Finch (Apple Inc.)
Orhan Firat (Google AI)
George Foster (Google)
Alexander Fraser (Ludwig-Maximilians-Universität München)
Atsushi Fujita (National Institute of Information and Communications Technology)
Juri Ganitkevitch (Google)
Mercedes García-Martínez (Pangeanic)
Ekaterina Garmash (KLM Royal Dutch Airlines)
Jesús González-Rubio (WebInterpret)
Isao Goto (NHK)
Miguel Graça (RWTH Aachen University)
Roman Grundkiewicz (School of Informatics, University of Edinburgh)
Mandy Guo (Google)
Jeremy Gwinnup (Air Force Research Laboratory)
Thanh-Le Ha (Karlsruhe Institute of Technology)
Nizar Habash (New York University Abu Dhabi)
Gholamreza Haffari (Monash University)
Greg Hanneman (Amazon)
Christian Hardmeier (Uppsala universitet)
Eva Hasler (SDL Research)
Yifan He (Alibaba Group)
John Henderson (MITRE)
Christian Herold (RWTH Aachen University)
Felix Hieber (Amazon Research)
Hieu Hoang (University of Edinburgh)

Vu Cong Duy Hoang (The University of Melbourne)
Bojie Hu (Tencent Research, Beijing, China)
Junjie Hu (Carnegie Mellon University)
Mika Hämäläinen (University of Helsinki)
Gonzalo Iglesias (SDL)
Kenji Imamura (National Institute of Information and Communications Technology)
Aizhan Imankulova (Tokyo Metropolitan University)
Julia Ive (University of Sheffield)
Marcin Junczys-Dowmunt (Microsoft)
Shahram Khadivi (eBay)
Huda Khayrallah (Johns Hopkins University)
Douwe Kiela (Facebook)
Yunsu Kim (RWTH Aachen University)
Rebecca Knowles (Johns Hopkins University)
Julia Kreutzer (Department of Computational Linguistics, Heidelberg University)
Shankar Kumar (Google)
Anoop Kunchukuttan (Microsoft AI and Research)
Surafel Melaku Lakew (University of Trento and Fondazione Bruno Kessler)
Ekaterina Lapshinova-Koltunski (Universität des Saarlandes)
Alon Lavie (Amazon/Carnegie Mellon University)
Gregor Leusch (eBay)
William Lewis (Microsoft Research)
Jindřich Libovický (Charles University)
Patrick Littell (National Research Council of Canada)
Qun Liu (Huawei Noah's Ark Lab)
Samuel Läubli (University of Zurich)
Pranava Madhyastha (Imperial College London)
Andreas Maletti (Universität Leipzig)
Saab Mansour (Apple)
Sameen Maruf (Monash University)
Arne Mauser (Google, Inc)
Arya D. McCarthy (Johns Hopkins University)
Antonio Valerio Miceli Barone (The University of Edinburgh)
Paul Michel (Carnegie Mellon University)
Aaron Mueller (The Johns Hopkins University)
Kenton Murray (University of Notre Dame)
Tomáš Musil (Charles University)
Mathias Müller (University of Zurich)
Masaaki Nagata (NTT Corporation)
Toshiaki Nakazawa (The University of Tokyo)
Preslav Nakov (Qatar Computing Research Institute, HBKU)
Graham Neubig (Carnegie Mellon University)
Jan Niehues (Maastricht University)
Nikola Nikolov (University of Zurich and ETH Zurich)
Xing Niu (University of Maryland)
Tsuyoshi Okita (Kyushuu institute of technology)
Daniel Ortiz-Martínez (Technical University of Valencia)

Myle Ott (Facebook AI Research)
Santanu Pal (Saarland University)
Carla Parra Escartín (Unbabel)
Pavel Pecina (Charles University)
Stephan Peitz (Apple)
Sergio Penkale (Lingo24)
Mārcis Pinnis (Tilde)
Martin Popel (Charles University, Faculty of Mathematics and Physics, UFAL)
Maja Popović (ADAPT Centre @ DCU)
Matīss Rikters (Tilde)
Annette Rios (Institute of Computational Linguistics, University of Zurich)
Jan Rosendahl (RWTH Aachen University)
Raphael Rubino (DFKI)
Devendra Sachan (CMU / Petuum Inc.)
Elizabeth Salesky (Carnegie Mellon University)
Hassan Sawaf (Amazon Web Services)
Jean Senellart (SYSTRAN)
Rico Sennrich (University of Edinburgh)
Patrick Simianer (Lilt)
Linfeng Song (University of Rochester)
Felix Stahlberg (University of Cambridge, Department of Engineering)
Dario Stojanovski (LMU Munich)
Katsuhito Sudoh (Nara Institute of Science and Technology (NAIST))
Felipe Sánchez-Martínez (Universitat d'Alacant)
Aleš Tamchyna (Charles University in Prague, UFAL MFF)
Gongbo Tang (Uppsala University)
Jörg Tiedemann (University of Helsinki)
Antonio Toral (University of Groningen)
Ke Tran (Amazon)
Marco Turchi (Fondazione Bruno Kessler)
Ferhan Ture (Comcast Applied AI Research)
Nicola Ueffing (eBay)
Masao Utiyama (NICT)
Dušan Variš (Charles University, Institute of Formal and Applied Linguistics)
David Vilar (Amazon)
Ivan Vulić (University of Cambridge)
Ekaterina Vylomova (University of Melbourne)
Wei Wang (Google Research)
Weiyue Wang (RWTH Aachen University)
Taro Watanabe (Google)
Philip Williams (University of Edinburgh)
Hua Wu (Baidu)
Joern Wuebker (Lilt, Inc.)
Hainan Xu (Johns Hopkins University)
Yinfei Yang (Google)
François Yvon (LIMSI/CNRS)
Dakun Zhang (SYSTRAN)
Xuan Zhang (Johns Hopkins University)

Table of Contents

Saliency-driven Word Alignment Interpretation for Neural Machine Translation
Shuoyang Ding, Hainan Xu and Philipp Koehn 1

Improving Zero-shot Translation with Language-Independent Constraints
Ngoc-Quan Pham, Jan Niehues, Thanh-Le Ha and Alexander Waibel 13

Incorporating Source Syntax into Transformer-Based Neural Machine Translation
Anna Currey and Kenneth Heafield 24

APE at Scale and Its Implications on MT Evaluation Biases
Markus Freitag, Isaac Caswell and Scott Roy 34

Generalizing Back-Translation in Neural Machine Translation
Miguel Graça, Yunsu Kim, Julian Schamper, Shahram Khadivi and Hermann Ney 45

Tagged Back-Translation
Isaac Caswell, Ciprian Chelba and David Grangier 53

Hierarchical Document Encoder for Parallel Corpus Mining
Mandy Guo, Yinfei Yang, Keith Stevens, Daniel Cer, Heming Ge, Yun-hsuan Sung, Brian Strope and Ray Kurzweil 64

The Effect of Translationese in Machine Translation Test Sets
Mike Zhang and Antonio Toral 73

Customizing Neural Machine Translation for Subtitling
Evgeny Matusov, Patrick Wilken and Yota Georgakopoulou 82

Integration of Dubbing Constraints into Machine Translation
Ashutosh Saboo and Timo Baumann 94

Widening the Representation Bottleneck in Neural Machine Translation with Lexical Shortcuts
Denis Emelin, Ivan Titov and Rico Sennrich 102

A High-Quality Multilingual Dataset for Structured Documentation Translation
Kazuma Hashimoto, Raffaella Buschiazzo, James Bradbury, Teresa Marshall, Richard Socher and Caiming Xiong 116

Conference Program

Thursday, August 1, 2019

8:45–9:00 ***Opening Remarks***

9:00–10:30 **Session 1: Shared Tasks Overview Presentations I (chair: Barry Haddow)**

9:00–9:35 *Findings of the 2019 Conference on Machine Translation (WMT19)*
Loïc Barrault, Ondřej Bojar, Marta R. Costa-jussà, Christian Federmann, Mark Fishel, Yvette Graham, Barry Haddow, Matthias Huck, Philipp Koehn, Shervin Malmasi, Christof Monz, Mathias Müller, Santanu Pal, Matt Post and Marcos Zampieri

9:35–9:50 *Test Suites*

9:50–10:10 *Results of the WMT19 Metrics Shared Task: Segment-Level and Strong MT Systems Pose Big Challenges*
Qingsong Ma, Johnny Wei, Ondřej Bojar and Yvette Graham

10:10–10:30 *Findings of the First Shared Task on Machine Translation Robustness*
Xian Li, Paul Michel, Antonios Anastasopoulos, Yonatan Belinkov, Nadir Durrani, Orhan Firat, Philipp Koehn, Graham Neubig, Juan Pino and Hassan Sajjad

10:30-11:00 ***Coffee Break***

11:00–12:30 **Session 2: Shared Task Poster Session I**

11:00–12:30 *Shared Task: News Translation*

The University of Edinburgh's Submissions to the WMT19 News Translation Task
Rachel Bawden, Nikolay Bogoychev, Ulrich Germann, Roman Grundkiewicz, Faheem Kirefu, Antonio Valerio Miceli Barone and Alexandra Birch

GTCOM Neural Machine Translation Systems for WMT19
Chao Bei, Hao Zong, Conghu Yuan, Qingming Liu and Baoyong Fan

Machine Translation with parfda, Moses, kenlm, nplm, and PRO
Ergun Biçici

Thursday, August 1, 2019 (continued)

LIUM's Contributions to the WMT2019 News Translation Task: Data and Systems for German-French Language Pairs
Fethi Bougares, Jane Wottawa, Anne Baillot, Loïc Barrault and Adrien Bardet

The University of Maryland's Kazakh-English Neural Machine Translation System at WMT19
Eleftheria Briakou and Marine Carpuat

DBMS-KU Interpolation for WMT19 News Translation Task
Sari Dewi Budiwati, Al Hafiz Akbar Maulana Siagian, Tirana Noor Fatyanosa and Masayoshi Aritsugi

Lingua Custodia at WMT'19: Attempts to Control Terminology
Franck Burlot

The TALP-UPC Machine Translation Systems for WMT19 News Translation Task: Pivoting Techniques for Low Resource MT
Noe Casas, José A. R. Fonollosa, Carlos Escolano, Christine Basta and Marta R. Costa-jussà

Kyoto University Participation to the WMT 2019 News Shared Task
Fabien Cromieres and Sadao Kurohashi

NICT's Supervised Neural Machine Translation Systems for the WMT19 News Translation Task
Raj Dabre, Kehai Chen, Benjamin Marie, Rui Wang, Atsushi Fujita, Masao Utiyama and Eiichiro Sumita

The University of Sydney's Machine Translation System for WMT19
Liang Ding and Dacheng Tao

UdS-DFKI Participation at WMT 2019: Low-Resource (en-gu) and Coreference-Aware (en-de) Systems
Cristina España-Bonet and Dana Ruiter

The IIIT-H Gujarati-English Machine Translation System for WMT19
Vikrant Goyal and Dipti Misra Sharma

Kingsoft's Neural Machine Translation System for WMT19
Xinze Guo, Chang Liu, Xiaolong Li, Yiran Wang, Guoliang Li, Feng Wang, Zhitao Xu, Liuyi Yang, Li Ma and Changliang Li

The AFRL WMT19 Systems: Old Favorites and New Tricks
Jeremy Gwinnup, Grant Erdmann and Tim Anderson

Evaluating the Supervised and Zero-shot Performance of Multi-lingual Translation Models
Chris Hokamp, John Glover and Demian Gholipour Ghalandari

The MLLP-UPV Supervised Machine Translation Systems for WMT19 News Translation Task
Javier Iranzo-Sánchez, Gonçal Garcés Díaz-Munío, Jorge Civera and Alfons Juan

Microsoft Translator at WMT 2019: Towards Large-Scale Document-Level Neural Machine Translation
Marcin Junczys-Dowmunt

CUNI Submission for Low-Resource Languages in WMT News 2019
Tom Kocmi and Ondřej Bojar

CUNI Systems for the Unsupervised News Translation Task in WMT 2019
Ivana Kvapilíková, Dominik Macháček and Ondřej Bojar

A Comparison on Fine-grained Pre-trained Embeddings for the WMT19Chinese-English News Translation Task
Zhenhao Li and Lucia Specia

The NiuTrans Machine Translation Systems for WMT19
Bei Li, Yinqiao Li, Chen Xu, Ye Lin, Jiqiang Liu, Hui Liu, Ziyang Wang, Yuhao Zhang, Nuo Xu, Zeyang Wang, Kai Feng, Hexuan Chen, Tengbo Liu, Yanyang Li, Qiang Wang, Tong Xiao and Jingbo Zhu

Multi-Source Transformer for Kazakh-Russian-English Neural Machine Translation
Patrick Littell, Chi-kiu Lo, Samuel Larkin and Darlene Stewart

Incorporating Word and Subword Units in Unsupervised Machine Translation Using Language Model Rescoring
Zihan Liu, Yan Xu, Genta Indra Winata and Pascale Fung

JUMT at WMT2019 News Translation Task: A Hybrid Approach to Machine Translation for Lithuanian to English
Sainik Kumar Mahata, Avishek Garain, Adityar Rayala, Dipankar Das and Sivaji Bandyopadhyay

Johns Hopkins University Submission for WMT News Translation Task
Kelly Marchisio, Yash Kumar Lal and Philipp Koehn

NICT's Unsupervised Neural and Statistical Machine Translation Systems for the WMT19 News Translation Task
Benjamin Marie, Haipeng Sun, Rui Wang, Kehai Chen, Atsushi Fujita, Masao Utiyama and Eiichiro Sumita

Thursday, August 1, 2019 (continued)

PROMT Systems for WMT 2019 Shared Translation Task
Alexander Molchanov

JU-Saarland Submission to the WMT2019 English–Gujarati Translation Shared Task
Riktim Mondal, Shankha Raj Nayek, Aditya Chowdhury, Santanu Pal, Sudip Kumar Naskar and Josef van Genabith

Facebook FAIR's WMT19 News Translation Task Submission
Nathan Ng, Kyra Yee, Alexei Baevski, Myle Ott, Michael Auli and Sergey Edunov

eTranslation's Submissions to the WMT 2019 News Translation Task
Csaba Oravecz, Katina Bontcheva, Adrien Lardilleux, László Tihanyi and Andreas Eisele

Tilde's Machine Translation Systems for WMT 2019
Marcis Pinnis, Rihards Krišlauks and Matiss Rikters

Apertium-fin-eng–Rule-based Shallow Machine Translation for WMT 2019 Shared Task
Tommi Pirinen

English-Czech Systems in WMT19: Document-Level Transformer
Martin Popel, Dominik Macháček, Michal Auersperger, Ondřej Bojar and Pavel Pecina

The RWTH Aachen University Machine Translation Systems for WMT 2019
Jan Rosendahl, Christian Herold, Yunsu Kim, Miguel Graça, Weiyue Wang, Parnia Bahar, Yingbo Gao and Hermann Ney

The Universitat d'Alacant Submissions to the English-to-Kazakh News Translation Task at WMT 2019
Víctor M. Sánchez-Cartagena, Juan Antonio Pérez-Ortiz and Felipe Sánchez-Martínez

CUED@WMT19:EWC&LMs
Felix Stahlberg, Danielle Saunders, Adrià de Gispert and Bill Byrne

Baidu Neural Machine Translation Systems for WMT19
Meng Sun, Bojian Jiang, Hao Xiong, Zhongjun He, Hua Wu and Haifeng Wang

University of Tartu's Multilingual Multi-domain WMT19 News Translation Shared Task Submission
Andre Tättar, Elizaveta Korotkova and Mark Fishel

Neural Machine Translation for English–Kazakh with Morphological Segmentation and Synthetic Data
Antonio Toral, Lukas Edman, Galiya Yeshmagambetova and Jennifer Spenader

The LMU Munich Unsupervised Machine Translation System for WMT19
Dario Stojanovski, Viktor Hangya, Matthias Huck and Alexander Fraser

Combining Local and Document-Level Context: The LMU Munich Neural Machine Translation System at WMT19
Dario Stojanovski and Alexander Fraser

IITP-MT System for Gujarati-English News Translation Task at WMT 2019
Sukanta Sen, Kamal Kumar Gupta, Asif Ekbal and Pushpak Bhattacharyya

The University of Helsinki Submissions to the WMT19 News Translation Task
Aarne Talman, Umut Sulubacak, Raúl Vázquez, Yves Scherrer, Sami Virpioja, Alessandro Raganato, Arvi Hurskainen and Jörg Tiedemann

Microsoft Research Asia's Systems for WMT19
Yingce Xia, Xu Tan, Fei Tian, Fei Gao, Di He, Weicong Chen, Yang Fan, Linyuan Gong, Yichong Leng, Renqian Luo, Yiren Wang, Lijun Wu, Jinhua Zhu, Tao Qin and Tie-Yan Liu

The En-Ru Two-way Integrated Machine Translation System Based on Transformer
Doron Yu

DFKI-NMT Submission to the WMT19 News Translation Task
Jingyi Zhang and Josef van Genabith

11:00–12:30 *Shared Task: Test Suites*

Linguistic Evaluation of German-English Machine Translation Using a Test Suite
Eleftherios Avramidis, Vivien Macketanz, Ursula Strohriegel and Hans Uszkoreit

A Test Suite and Manual Evaluation of Document-Level NMT at WMT19
Kateřina Rysová, Magdaléna Rysová, Tomáš Musil, Lucie Poláková and Ondřej Bojar

Evaluating Conjunction Disambiguation on English-to-German and French-to-German WMT 2019 Translation Hypotheses
Maja Popović

Thursday, August 1, 2019 (continued)

The MuCoW Test Suite at WMT 2019: Automatically Harvested Multilingual Contrastive Word Sense Disambiguation Test Sets for Machine Translation
Alessandro Raganato, Yves Scherrer and Jörg Tiedemann

SAO WMT19 Test Suite: Machine Translation of Audit Reports
Tereza Vojtěchová, Michal Novák, Miloš Klouček and Ondřej Bojar

11:00–12:30 *Shared Task: Metrics*

WMDO: Fluency-based Word Mover's Distance for Machine Translation Evaluation
Julian Chow, Lucia Specia and Pranava Madhyastha

Meteor++ 2.0: Adopt Syntactic Level Paraphrase Knowledge into Machine Translation Evaluation
Yinuo Guo and Junfeng Hu

YiSi - a Unified Semantic MT Quality Evaluation and Estimation Metric for Languages with Different Levels of Available Resources
Chi-kiu Lo

EED: Extended Edit Distance Measure for Machine Translation
Peter Stanchev, Weiyue Wang and Hermann Ney

Filtering Pseudo-References by Paraphrasing for Automatic Evaluation of Machine Translation
Ryoma Yoshimura, Hiroki Shimanaka, Yukio Matsumura, Hayahide Yamagishi and Mamoru Komachi

11:00–12:30 *Shared Task: Robustness*

Naver Labs Europe's Systems for the WMT19 Machine Translation Robustness Task
Alexandre Berard, Ioan Calapodescu and Claude Roux

NICT's Supervised Neural Machine Translation Systems for the WMT19 Translation Robustness Task
Raj Dabre and Eiichiro Sumita

System Description: The Submission of FOKUS to the WMT 19 Robustness Task
Cristian Grozea

CUNI System for the WMT19 Robustness Task
Jindřich Helcl, Jindřich Libovický and Martin Popel

NTT's Machine Translation Systems for WMT19 Robustness Task
Soichiro Murakami, Makoto Morishita, Tsutomu Hirao and Masaaki Nagata

JHU 2019 Robustness Task System Description
Matt Post and Kevin Duh

Robust Machine Translation with Domain Sensitive Pseudo-Sources: Baidu-OSU WMT19 MT Robustness Shared Task System Report
Renjie Zheng, Hairong Liu, Mingbo Ma, Baigong Zheng and Liang Huang

Improving Robustness of Neural Machine Translation with Multi-task Learning
Shuyan Zhou, Xiangkai Zeng, Yingqi Zhou, Antonios Anastasopoulos and Graham Neubig

12:30–14:00 ***Lunch***

14:00–15:30 **Panel on "Open Problems in Machine Translation" (chair Ondrej Bojar): Alex Fraser (Ludwig-Maximilians-Universität München), Alon Lavie (Unbabel), Marcin Junczys-Dowmunt (Microsoft), Yvette Graham (Dublin City University)**

15:30–16:00 ***Coffee Break***

16:00–17:30 **Session 4: Research Papers on Modeling and Analysis (chair: Matthias Huck)**

16:00–16:15 *Saliency-driven Word Alignment Interpretation for Neural Machine Translation*
Shuoyang Ding, Hainan Xu and Philipp Koehn

16:15–16:30 *Improving Zero-shot Translation with Language-Independent Constraints*
Ngoc-Quan Pham, Jan Niehues, Thanh-Le Ha and Alexander Waibel

16:30–16:45 *Incorporating Source Syntax into Transformer-Based Neural Machine Translation*
Anna Currey and Kenneth Heafield

Thursday, August 1, 2019 (continued)

16:45–17:00 *APE at Scale and Its Implications on MT Evaluation Biases*
Markus Freitag, Isaac Caswell and Scott Roy

17:00–17:15 *Generalizing Back-Translation in Neural Machine Translation*
Miguel Graça, Yunsu Kim, Julian Schamper, Shahram Khadivi and Hermann Ney

17:15–17:30 *Tagged Back-Translation*
Isaac Caswell, Ciprian Chelba and David Grangier

Friday, August 2, 2019

9:00–10:30 Session 5: Shared Tasks Overview Presentations II (chair: Mark Fishel)

9:00–9:20 *Findings of the WMT 2019 Shared Tasks on Quality Estimation*
Erick Fonseca, Lisa Yankovskaya, André F. T. Martins, Mark Fishel and Christian Federmann

9:20–9:40 *Findings of the WMT 2019 Shared Task on Automatic Post-Editing*
Rajen Chatterjee, Christian Federmann, Matteo Negri and Marco Turchi

9:40–10:00 *Findings of the WMT 2019 Biomedical Translation Shared Task: Evaluation for MEDLINE Abstracts and Biomedical Terminologies*
Rachel Bawden, Kevin Bretonnel Cohen, Cristian Grozea, Antonio Jimeno Yepes, Madeleine Kittner, Martin Krallinger, Nancy Mah, Aurelie Neveol, Mariana Neves, Felipe Soares, Amy Siu, Karin Verspoor and Maika Vicente Navarro

10:00–10:15 *Shared Task on Similar Languages*

10:15–10:30 *Findings of the WMT 2019 Shared Task on Parallel Corpus Filtering for Low-Resource Conditions*
Philipp Koehn, Francisco Guzmán, Vishrav Chaudhary and Juan Pino

10:30-11:00 ***Coffee Break***

Friday, August 2, 2019 (continued)

11:00–12:30 Session 2: Shared Task Poster Session I

11:00–12:30 *Shared Task: Quality Estimation*

RTM Stacking Results for Machine Translation Performance Prediction
Ergun Biçici

Unbabel's Participation in the WMT19 Translation Quality Estimation Shared Task
Fabio Kepler, Jonay Trénous, Marcos Treviso, Miguel Vera, António Góis, M. Amin Farajian, António V. Lopes and André F. T. Martins

QE BERT: Bilingual BERT Using Multi-task Learning for Neural Quality Estimation
Hyun Kim, Joon-Ho Lim, Hyun-Ki Kim and Seung-Hoon Na

MIPT System for World-Level Quality Estimation
Mikhail Mosyagin and Varvara Logacheva

NJU Submissions for the WMT19 Quality Estimation Shared Task
Hou Qi

Quality Estimation and Translation Metrics via Pre-trained Word and Sentence Embeddings
Elizaveta Yankovskaya, Andre Tättar and Mark Fishel

SOURCE: SOURce-Conditional Elmo-style Model for Machine Translation Quality Estimation
Junpei Zhou, Zhisong Zhang and Zecong Hu

11:00–12:30 *Shared Task: Automatic Post-Editing*

Transformer-based Automatic Post-Editing Model with Joint Encoder and Multi-source Attention of Decoder
WonKee Lee, Jaehun Shin and Jong-Hyeok Lee

Unbabel's Submission to the WMT2019 APE Shared Task: BERT-Based Encoder-Decoder for Automatic Post-Editing
António V. Lopes, M. Amin Farajian, Gonçalo M. Correia, Jonay Trénous and André F. T. Martins

Friday, August 2, 2019 (continued)

USAAR-DFKI – The Transference Architecture for English–German Automatic Post-Editing
Santanu Pal, Hongfei Xu, Nico Herbig, Antonio Krüger and Josef van Genabith

APE through Neural and Statistical MT with Augmented Data. ADAPT/DCU Submission to the WMT 2019 APE Shared Task
Dimitar Shterionov, Joachim Wagner and Félix do Carmo

Effort-Aware Neural Automatic Post-Editing
Amirhossein Tebbifakhr, Matteo Negri and Marco Turchi

UdS Submission for the WMT 19 Automatic Post-Editing Task
Hongfei Xu, Qiuhui Liu and Josef van Genabith

11:00–12:30 *Shared Task: Biomedical Translation*

Terminology-Aware Segmentation and Domain Feature for the WMT19 Biomedical Translation Task
Casimiro Pio Carrino, Bardia Rafieian, Marta R. Costa-jussà and José A. R. Fonollosa

Exploring Transfer Learning and Domain Data Selection for the Biomedical Translation
Noor-e- Hira, Sadaf Abdul Rauf, Kiran Kiani, Ammara Zafar and Raheel Nawaz

Huawei's NMT Systems for the WMT 2019 Biomedical Translation Task
Wei Peng, Jianfeng Liu, Liangyou Li and Qun Liu

UCAM Biomedical Translation at WMT19: Transfer Learning Multi-domain Ensembles
Danielle Saunders, Felix Stahlberg and Bill Byrne

BSC Participation in the WMT Translation of Biomedical Abstracts
Felipe Soares and Martin Krallinger

11:00–12:30 *Shared Task: Similar Languages*

The MLLP-UPV Spanish-Portuguese and Portuguese-Spanish Machine Translation Systems for WMT19 Similar Language Translation Task
Pau Baquero-Arnal, Javier Iranzo-Sánchez, Jorge Civera and Alfons Juan

Friday, August 2, 2019 (continued)

The TALP-UPC System for the WMT Similar Language Task: Statistical vs Neural Machine Translation
Magdalena Biesialska, Lluis Guardia and Marta R. Costa-jussà

Machine Translation from an Intercomprehension Perspective
Yu Chen and Tania Avgustinova

Utilizing Monolingual Data in NMT for Similar Languages: Submission to Similar Language Translation Task
Jyotsana Khatri and Pushpak Bhattacharyya

Neural Machine Translation: Hindi-Nepali
Sahinur Rahman Laskar, Partha Pakray and Sivaji Bandyopadhyay

NICT's Machine Translation Systems for the WMT19 Similar Language Translation Task
Benjamin Marie, Raj Dabre and Atsushi Fujita

Panlingua-KMI MT System for Similar Language Translation Task at WMT 2019
Atul Kr. Ojha, Ritesh Kumar, Akanksha Bansal and Priya Rani

UDS–DFKI Submission to the WMT2019 Czech–Polish Similar Language Translation Shared Task
Santanu Pal, Marcos Zampieri and Josef van Genabith

Neural Machine Translation of Low-Resource and Similar Languages with Back-translation
Michael Przystupa and Muhammad Abdul-Mageed

The University of Helsinki Submissions to the WMT19 Similar Language Translation Task
Yves Scherrer, Raúl Vázquez and Sami Virpioja

11:00–12:30 *Shared Task: Parallel Corpus Filtering*

Dual Monolingual Cross-Entropy Delta Filtering of Noisy Parallel Data
Amittai Axelrod, Anish Kumar and Steve Sloto

NRC Parallel Corpus Filtering System for WMT 2019
Gabriel Bernier-Colborne and Chi-kiu Lo

Friday, August 2, 2019 (continued)

Low-Resource Corpus Filtering Using Multilingual Sentence Embeddings
Vishrav Chaudhary, Yuqing Tang, Francisco Guzmán, Holger Schwenk and Philipp Koehn

Quality and Coverage: The AFRL Submission to the WMT19 Parallel Corpus Filtering for Low-Resource Conditions Task
Grant Erdmann and Jeremy Gwinnup

Webinterpret Submission to the WMT2019 Shared Task on Parallel Corpus Filtering
Jesús González-Rubio

Noisy Parallel Corpus Filtering through Projected Word Embeddings
Murathan Kurfalı and Robert Östling

Filtering of Noisy Parallel Corpora Based on Hypothesis Generation
Zuzanna Parcheta, Germán Sanchis-Trilles and Francisco Casacuberta

Parallel Corpus Filtering Based on Fuzzy String Matching
Sukanta Sen, Asif Ekbal and Pushpak Bhattacharyya

The University of Helsinki Submission to the WMT19 Parallel Corpus Filtering Task
Raúl Vázquez, Umut Sulubacak and Jörg Tiedemann

12:30–14:00 ***Lunch***

Friday, August 2, 2019 (continued)

14:00–15:30 Session 7: Invited Talk (chair: Matt Post)

14:00–15:30 *Marine Carpuat (University of Maryland): Semantic, Style and Other Data Divergences in Neural Machine Translation*

15:30–16:00 ***Coffee Break***

16:00–17:30 Session 8: Research Papers on Applications (chair: Marco Turchi)

16:00–16:15 *Hierarchical Document Encoder for Parallel Corpus Mining*
Mandy Guo, Yinfei Yang, Keith Stevens, Daniel Cer, Heming Ge, Yun-hsuan Sung, Brian Strope and Ray Kurzweil

16:15–16:30 *The Effect of Translationese in Machine Translation Test Sets*
Mike Zhang and Antonio Toral

16:30–16:45 *Customizing Neural Machine Translation for Subtitling*
Evgeny Matusov, Patrick Wilken and Yota Georgakopoulou

16:45–17:00 *Integration of Dubbing Constraints into Machine Translation*
Ashutosh Saboo and Timo Baumann

17:00–17:15 *Widening the Representation Bottleneck in Neural Machine Translation with Lexical Shortcuts*
Denis Emelin, Ivan Titov and Rico Sennrich

17:15–17:30 *A High-Quality Multilingual Dataset for Structured Documentation Translation*
Kazuma Hashimoto, Raffaella Buschiazzo, James Bradbury, Teresa Marshall, Richard Socher and Caiming Xiong