

WMT 2025

**The Tenth Conference
on Machine Translation**

Proceedings of the Conference

November 8-9, 2025

©2025 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 979-8-89176-341-8

Introduction

The Tenth Conference on Machine Translation (WMT 2025) took place on November 8–9, 2025, immediately following the 2025 Conference on Empirical Methods in Natural Language Processing (EMNLP 2025) in Suzhou, China.

This is the tenth 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, the third time at EMNLP 2018 in Brussels, Belgium, the fourth time at ACL 2019 in Florence, Italy, the fifth time at EMNLP-2020, which was held as an online event due to the COVID-19 pandemic, the sixth time at EMNLP 2021 at Punta Cana, Dominican Republic, the seventh time at EMNLP 2022 in Abu Dhabi, United Arab Emirates, the eighth time at EMNLP 2023 in Singapore, and the ninth time at EMNLP 2024 in Miami, USA. 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, 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 10 shared tasks. These consisted of 5 translation tasks: General Translation, Low-Resource Indic Language Translation, Terminology Translation, Creole Language Translation, and Model Compression, two evaluation tasks: MT Test Suites (“Help us break LLMs, Vol. 2”) and Automated Translation Quality Evaluation Systems, two multilingual tasks: Multilingual Instruction and LLMs with Limited Resources for Slavic Languages and finally the Open Language Data Initiative.

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 2025 has received 60 full research paper submissions (not counting withdrawn submissions). In total, WMT 2025 featured 21 full research paper presentations and 79 shared task presentations.

The invited talk was given by Longyue Wang from Alibaba, China.

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.

Barry Haddow, Tom Kocmi, Philipp Koehn, and Christof Monz

Co-Organizers

Organizers:

Barry Haddow (University of Edinburgh)
Tom Kocmi (Cohere)
Philipp Koehn (Johns Hopkins University)
Christof Monz (University of Amsterdam)

Program Committee:

Sadaf Abdul Rauf (Fatima Jinnah Women University)
Kolawole Adebayo (Dublin City University)
David Ifeoluwa Adelani (McGill University / MILA)
Aniefon Akpan (Univeristy of Uyo)
Asim Alaskar (King Saud University)
Priscilla Amuok (Masakhane)
Antonios Anastasopoulos (George Mason University)
Anietie Andy (Howard University)
Ramakrishna Appicharla (Indian Institute of Technology Patna)
Ali Araabi (University of Amsterdam)
Anuoluwapo Aremu (Lelapa AI, University of Trento, Masakhane)
Joseph Attieh (University of Helsinki)
Mikko Aulamo (University of Helsinki)
Eleftherios Avramidis (German Research Center for Artificial Intelligence (DFKI))
Olumide Awokoya (Masakhane)
Seth Aycock (University of Amsterdam)
Fatemeh Azadi (National Research Council Canada)
Selene Baez Santamaria (University of Zurich)
Parnia Bahar (Apple)
Rachel Bawden (Inria)
Meriem Beloucif (Uppsala University)
Nathaniel Berger (Heidelberg University)
Toms Bergmanis (Tilde)
Siddharth Betala (Indian Institute of Technology Madras)
Alexandra Birch (University of Edinburgh)
Ondřej Bojar (Charles University, MFF UFAL)
Maharaj Brahma (Indian Institute of Technology Hyderabad)
Eleftheria Briakou (Google)
José G. C. de Souza (Unbabel)
Marine Carpuat (University of Maryland)
Antonio Castaldo (University of Naples “L’Orientale”)
Sheila Castilho (Dublin City University)
Rajen Chatterjee (Apple Inc.)
Pinzhen Chen (University of Edinburgh & Aveni)
Colin Cherry (Google)
Pranjal Chitale (Indian Institute of Technology, Madras)
Katsuki Chousa (NTT)
Chenhui Chu (Kyoto University)

Josep Crego (CHAPSVISION)
Sören DREANO (Dublin City University)
David Dale (Meta AI)
Brian Davis (Dublin City University)
Steve DeNeefe (RWS Language Weaver)
Hiroyuki Deguchi (NTT, Inc.)
Michael Denkowski (Amazon)
Sourabh Deoghare (IIT Bombay)
Daniel Deutsch (Google)
Giorgio Maria Di Nunzio (University of Padua)
Shuoyang Ding (NVIDIA)
Miguel Domingo (Universitat Politècnica de València)
Kevin Duh (Johns Hopkins University)
Koel Dutta Chowdhury (Saarland Informatics Campus, Saarland University)
Hiroshi Echizen'ya (Hokkai-Gakuen University)
Hafsteinn Einarsson (University of Iceland)
Muhammad ElNokrashy (Microsoft)
Grant Erdmann (Air Force Research Laboratory)
Carlos Escolano (Universitat Politècnica de Catalunya, Barcelona Supercomputing Center)
Miquel Esplà-Gomis (Universitat d'Alacant)
Marcello Federico (Amazon)
Patrick Fernandes (Carnegie Mellon University, Instituto de Telecomunicações)
Aloka Fernando (University of Moratuwa)
Edoardo Ferrante (Conseggio pe-o patrimonio linguistico ligure)
Ryo Fujii (Future Corporation)
Atsushi Fujita (National Institute of Information and Communications Technology)
Marco Gaido (Fondazione Bruno Kessler, University of Trento)
Baban Gain (Indian Institute of Technology, Patna)
Javier Garcia Gilabert (Barcelona Super Computing Center)
Sofía García (imaxinlsoftware)
Mercedes García-Martínez (Pangeanic)
Harritxu Gete (Vicomtech)
Ona de Gibert (University of Helsinki)
Jesús González-Rubio (WebInterpret)
Isao Goto (Ehime University)
Thamme Gowda (Microsoft)
Varun Gumma (Microsoft)
Kamil Guttman (Lanigo, Adam Mickiewicz University in Poznań)
Jeremy Gwinnup (Air Force Research Laboratory)
Anne Göhring (University of Zurich)
Nizar Habash (New York University Abu Dhabi)
Sami Haq (Dublin City University)
Ali Hatami (University of Galway)
Xiaoyu He (University of Edinburgh)
Jindřich Helcl (Charles University in Prague)
John Henderson (Mechanical Learning)
Amr Hendy (Microsoft)

Shohei Higashiyama (National Institute of Information and Communications Technology)
 Toshio Hirasawa (OMRON SINIC X Corporation)
 Miroslav Hrabal (Charles University)
 Shujian Huang (National Key Laboratory for Novel Software Technology, Nanjing University)
 Rudali Huidrom (ADAPT Research Centre, Dublin City University)
 Ifeoluwatayo Ige (Rochester Institute of Technology)
 Kenji Imamura (National Institute of Information and Communications Technology)
 Mert Inan (Northeastern University)
 Jilaphat Jaroenkantasima (Kasetsart University)
 Advait Joglekar (Indian Institute of Technology Madras)
 Josef Jon (Charles University)
 Nisheeth Joshi (Banasthali Vidyapith)
 Takatomo Kano (NTT)
 Marzena Karpinska (University of Massachusetts Amherst)
 Kishore Kashyap (Department of Information Technology , Gauhati University)
 Kazutaka Kinugawa (NHK Science & Technology Research Laboratories)
 Mateusz Klimaszewski (Warsaw University of Technology)
 Sai Koneru (Karlsruhe Institute of Technology)
 Maarit Koponen (University of Eastern Finland)
 Elizaveta Korotkova (University of Tartu)
 Mikołaj Koszowski (ML Research at Allegro)
 Artur Kot (Allegro.com)
 Lea Krause (Vrije Universiteit Amsterdam)
 Parameswari Krishnamurthy (Assistant Professor, IIIT Hyderabad)
 Mateusz Krubinski (Snowflake)
 Roland Kuhn (National Research Council of Canada)
 Ashish Kulkarni (Krutrim)
 Anoop Kunchukuttan (Microsoft AI and Research)
 Ali Kuzhuget (tyvan.ru)
 Ben LI (Nomura Research Institute, Ltd.)
 Wen Lai (Technical University of Munich)
 Lenin Laitonjam (NIT Mizoram)
 Ekaterina Lapshinova-Koltunski (University of Hildesheim)
 Samuel Larkin (National Research Council Canada)
 Sahinur Rahman Laskar (UPES Dehradun)
 Alon Lavie (Unbabel/Carnegie Mellon University)
 Annie Lee (University of Toronto)
 Gregor Leusch (eBay)
 William Lewis (University of Washington)
 Zhenhao Li (Imperial College London)
 Baohao Liao (University of Amsterdam)
 Xixian Liao (Barcelona Supercomputing Center)
 Danni Liu (Karlsruhe Institute of Technology)
 Lisa Liu (University of California, San Diego)
 Siyou Liu (University of Macau)
 Wuying Liu (Ludong University)
 Chi-kiu Lo (National Research Council of Canada)

Henrique Lopes Cardoso (University of Porto)
Jiaming Luo (Google)
Nam Luu (Charles University)
Chenyang Lyu (MBZUAI)
Andreas Maletti (Universität Leipzig)
Bhavivy Malik (University of Edinburgh)
Shushen Manakhimova (German Research Center for Artificial Intelligence (DFKI))
Euan McGill (Universitat Pompeu Fabra)
Maite Melero (BSC)
Yan Meng (University of Amsterdam)
Antonio Valerio Miceli Barone (The University of Edinburgh)
Nikolay Mikhaylovskiy (NTR Labs / Higher IT School of Tomsk State University)
Hideya Mino (NHK Science & Technology Research Laboratories)
sthemviso Mkhwanazi (CSIR)
Wafaa Mohammed (University of Amsterdam)
Amit Moryossef (Bar-Ilan university, University of Zurich)
Yasmin Moslem (ADAPT Centre, Trinity College Dublin)
Ananya Mukherjee (International Institute of Information Technology Hyderabad)
Aniruddha Mukherjee (Kalinga Institute of Industrial Technology)
Kenton Murray (Johns Hopkins University)
Jonathan Mutal (Unige)
Mathias Müller (University of Zurich)
Masaaki Nagata (NTT Inc.)
Toshiaki Nakazawa (The University of Tokyo)
Subhajit Naskar (Google)
Angel Navarro (PRHLT)
Prashanth Nayak (KantanAI)
Mariana Neves (German Federal Institute for Risk Assessment)
Jan Niehues (Karlsruhe Institut of Technology)
Takashi Ninomiya (Ehime University)
Lydia Nishimwe (Inria)
Artur Nowakowski (Lanigo / Adam Mickiewicz University)
Perez Ogayo (Carnegie Mellon University)
Tsuyoshi Okita (Kyushu institute of technology)
Antoni Oliver (Universitat Oberta de Catalunya)
Constantin Orasan (University of Surrey)
Lucía Ormaechea (Université de Genève)
Daniel Ortiz-Martinez (University of Barcelona)
Partha Pakray (National Institute of Technology Silchar)
Santanu Pal (Wipro)
Jianhui Pang (University of Macau)
Chanjun Park (Korea University)
Pavel Pecina (Charles University)
Stephan Peitz (Apple)
Sergio Penkale (Lingo24)
Frithjof Petrick (AppTek)
Pavel Petrushkov (eBay)

Sanjita Phijam (National Institute of Technology Silchar)
Mārcis Pinnis (Tilde)
Mikołaj Pokrywka (Lanigo, Allegro, Adam Mickiewicz University)
Jose Pombal (Unbabel)
Martin Popel (Charles University, Faculty of Mathematics and Physics, UFAL)
Maja Popović (IU University)
Matt Post (Microsoft)
Pawan Rajpoot (Self)
Surangika Ranathunga (Massey University)
Vikas Raunak (Microsoft)
Pretam Ray (Indian Institute of Technology Kharagpur)
Stephen Richardson (Brigham Young University)
Matiss Rikters (AIST)
Nathaniel Robinson (Johns Hopkins University)
Raphael Rubino (UNIGE)
Pavel Rychly (NLP Centre, Faculty of Informatics, Masaryk University)
Benoît Sagot (Inria)
Pramit Sahoo (Indian Institute of Technology Hyderabad)
Yusuke Sakai (Nara Institute of Science and Technology)
Elizabeth Salesky (Johns Hopkins University)
Loitongbam Sanayai Meetei (National Institute of Technology Silchar)
Marcelo Sandoval-Castaneda (Toyota Technological Institute at Chicago)
Aleix Sant (Barcelona Supercomputing Center)
Ashish Sardana (NVIDIA)
Hamees Sayed (Indian Institute of Technology Madras)
Yves Scherrer (University of Oslo)
Florian Schottmann (TextShuttle, ETH Zurich)
Rico Sennrich (University of Zurich)
Hendra Setiawan (Apple Inc.)
Apurva Shah (Google Inc)
Lia Shahnazaryan (Paderborn University)
Mohammed Shaikheldin (Independent Researcher)
Sherrie Shen (University of Edinburgh)
Ketaki Shetye (International Institute of Information Technology)
Manish Shrivastava (International Institute of Information Technology Hyderabad)
Ammon Shurtz (Brigham Young University)
Edoardo Signoroni (NLP Centre, Faculty of Informatics, Masaryk University)
Nisansa de Silva (University of Moratuwa)
Thoudam Doren Singh (National Institute of Technology Meghalaya)
Huacheng Song (The Hong Kong Polytechnic University)
Rui Sousa-Silva (University of Porto - Faculty of Arts and Humanities)
Felix Stahlberg (Google Research)
Steinthor Steingrímsson (The Árni Magnússon Institute for Icelandic Studies)
Katsuhito Sudoh (Nara Women's University)
Haoxiang Sun (Shanghai Jiaotong University)
Eduardo Sánchez (Meta/UCL)
George Tambouratzis (ILSP/Athena R.C.)

Aleš Tamchyna (Memsources)
 Shaomu Tan (University of Amsterdam)
 Gongbo Tang (Beijing Language and Culture University)
 Jörg Tiedemann (University of Helsinki)
 Evgeniia Tokarchuk (University of Amsterdam)
 Isidora Tourni (Boston University)
 Marco Turchi (Zoom Video Communications)
 Masao Utiyama (NICT)
 Takehito Utsuro (University of Tsukuba)
 Jannis Vamvas (Department of Computational Linguistics, University of Zurich)
 Dusan Varis (Charles University, Institute of Formal and Applied Linguistics)
 Raul Vazquez (University of Helsinki)
 Menan Velayuthan (University of Jaffna)
 Nitin Venkateswaran (University of Florida)
 Federica Vezzani (University of Padua)
 David Vilar (Google)
 Ekaterina Vylomova (University of Melbourne)
 Wei Wang (Apple AI/ML)
 Xiaotian Wang (University of Tsukuba)
 Taro Watanabe (Nara Institute of Science and Technology)
 Rachel Wicks (Johns Hopkins University)
 Guillaume Wisniewski (Université Paris Cité and LLF)
 Derek F. Wong (University of Macau)
 Di Wu (University of Amsterdam)
 Minghao Wu (Monash University)
 Yulong Wu (The University of Manchester)
 Tong Xiao (Northeastern University)
 Hongzhi Xu (Shanghai International Studies University)
 Saumitra Yadav (International Institute of Information Technology, Hyderabad)
 Hao Yang (Huawei Co. Ltd)
 Kazuki Yano (Tohoku University)
 Zekai Ye (Harbin Institute of Technology)
 Hyeon-gu Yun (LG AI Research)
 François Yvon (ISIR CNRS & Sorbonne Université)
 Armel Randy Zebaze Dongmo (Inria)
 Hui Zeng (Jiaguyi Language Technology Co., Ltd)
 Chrysoula Zerva (Instituto de Telecomunicações, Instituto Superior Técnico, University of Lisbon)
 Runzhe Zhan (University of Macau)
 Dakun Zhang (CHAPSVISION)
 Min Zhang (Huawei)
 Xuan Zhang (Johns Hopkins University)
 Lichao Zhu (Paris Cité University)
 Shaolin Zhu (Tianjin university)
 Maria Zimina (Paris Diderot Sorbonne Paris Cité, CLILLAC-ARP (EA 3967))
 Vilém Zouhar (ETH Zurich, Charles University)

Table of Contents

Findings of the WMT25 General Machine Translation Shared Task: Time to Stop Evaluating on Easy Test Sets

Tom Kocmi, Ekaterina Artemova, Eleftherios Avramidis, Rachel Bawden, Ondřej Bojar, Konstantin Dranch, Anton Dvorkovich, Sergey Dukanov, Mark Fishel, Markus Freitag, Thamme Gowda, Roman Grundkiewicz, Barry Haddow, Marzena Karpinska, Philipp Koehn, Howard Lakounga, Jessica Lundin, Christof Monz, Kenton Murray, Masaaki Nagata, Stefano Perrella, Lorenzo Proietti, Martin Popel, Maja Popović, Parker Riley, Mariya Shmatova, Steinþór Steingrímsson, Lisa Yankovskaya and Vilém Zouhar 1

Findings of the WMT25 Shared Task on Automated Translation Evaluation Systems: Linguistic Diversity is Challenging and References Still Help

Alon Lavie, Greg Hanneman, Sweta Agrawal, Diptesh Kanojia, Chi-kiu Lo, Vilém Zouhar, Frederic Blain, Chrysoula Zerva, Eleftherios Avramidis, Sourabh Dattatray Deoghare, Archchana Sindhuja, Jiayi Wang, David Ifeoluwa Adelani, Brian Thompson, Tom Kocmi, Markus Freitag and Daniel Deutsch . 60

Findings of the WMT25 Multilingual Instruction Shared Task: Persistent Hurdles in Reasoning, Generation, and Evaluation

Tom Kocmi, Sweta Agrawal, Ekaterina Artemova, Eleftherios Avramidis, Eleftheria Briakou, Pinzhen Chen, Marzieh Fadaee, Markus Freitag, Roman Grundkiewicz, Yupeng Hou, Philipp Koehn, Julia Kreutzer, Saab Mansour, Stefano Perrella, Lorenzo Proietti, Parker Riley, Eduardo Sánchez, Patricia Schmidova, Mariya Shmatova and Vilém Zouhar 108

Miðeind at WMT25 General Machine Translation Task

Svanhvít Lilja Ingólfssdóttir, Haukur Páll Jónsson, Kári Steinn Aðalsteinsson, Róbert Fjölfnir Biriksson, Sveinbjörn Þórðarson and Þorvaldur Páll Helgason 130

A Preliminary Study of AI Agent Model in Machine Translation

Ahrii Kim 136

Marco Large Translation Model at WMT2025: Transforming Translation Capability in LLMs via Quality-Aware Training and Decoding

Hao Wang, Linlong Xu, Heng Liu, Yangyang Liu, Xiaohu Zhao, Bo Zeng, Longyue Wang, Weihua Luo and Kaifu Zhang 140

Evaluation of QWEN-3 for English to Ukrainian Translation

Cristian Grozea and Oleg Verbitsky 147

SYSTRAN @ WMT 2025 General Translation Task

Dakun Zhang, Yara Khater, Ramzi Rahli, Anna Rebollo and Josep Crego 152

Shy-hunyuan-MT at WMT25 General Machine Translation Shared Task

Mao Zheng, Zheng Li, Yang Du, Bingxin Qu and Mingyang Song 160

From SALAMANDRA to SALAMANDRATA: BSC Submission for WMT25 General Machine Translation Shared Task

Javier Garcia Gilabert, Xixian Liao, Severino Da Dalt, Ella Bohman, Audrey Mash, Francesca De Luca Fornaciari, Irene Baucells, Joan Llop, Miguel Claramunt, Carlos Escolano and Maite Melero . 167

Instruction-Tuned English to Bhojpuri Neural Machine Translation Using Contrastive Preference Optimization

Kshetrimayum Boynao Singh, Deepak Kumar and Asif Ekbal 191

<i>SH at WMT25 General Machine Translation Task</i>	
Hayate Shiroma	197
<i>Simple Test Time Scaling for Machine Translation: Kaze-MT at the WMT25 General Translation Task</i>	
Shaomu Tan	204
<i>NTTSU at WMT2025 General Translation Task</i>	
Zhang Yin, Hiroyuki Deguchi, Haruto Azami, Guanyu Ouyang, Kosei Buma, YINGYI FU, Katsuki Chousa and Takehito Utsuro	210
<i>A* Decoding for Machine Translation in LLMs - SRPOL Participation in WMT2025</i>	
Adam Dobrowolski, Paweł Przewłocki, Paweł Przybysz, Marcin Szymański and Dawid Siwicki	219
<i>In2X at WMT25 Translation Task</i>	
Lei Pang, Hanyi Mao, Quanjia Xiao, Chen Ruihan, Jingjun Zhang, Haixiao Liu and Xiangyi Li	224
<i>CUNI at WMT25 General Translation Task</i>	
Josef Jon, Miroslav Hrabal, Martin Popel and Ondřej Bojar	233
<i>UvA-MT's Participation in the WMT25 General Translation Shared Task</i>	
Di Wu, Yan Meng, Maya Konstantinovna Nachesa, Seth Aycok and Christof Monz	241
<i>AMI at WMT25 General Translation Task: How Low Can We Go? Finetuning Lightweight Llama Models for Low Resource Machine Translation</i>	
Atli Jasonarson and Steinthor Steingrímsson	248
<i>KIKIS at WMT 2025 General Translation Task</i>	
Koichi Iwakawa, Keito Kudo, Subaru Kimura, Takumi Ito and Jun Suzuki	258
<i>Google Translate's Research Submission to WMT2025</i>	
Mara Finkelstein, Geza Kovacs, Isaac Caswell, Tobias Domhan, Jan-Thorsten Peter, Juraj Juraska, Markus Freitag and David Vilar	276
<i>DLUT and GTCOM's Large Language Model Based Translation System for WMT25</i>	
Hao Zong, Chao Bei, Wentao Chen, Conghu Yuan, Huan Liu and Degen Huang	285
<i>Yandex Submission to the WMT25 General Machine Translation Task</i>	
Nikolay Karpachev, Ekaterina Enikeeva, Dmitry Popov, Arsenii Bulgakov, Daniil Panteleev, Dmitrii Ulianov, Artem Kryukov and Artem Mekhraliev	293
<i>IRB-MT at WMT25 Translation Task: A Simple Agentic System Using an Off-the-Shelf LLM</i>	
Ivan Grubišić and Damir Korencić	306
<i>Improving Low-Resource Japanese Translation with Fine-Tuning and Backtranslation for the WMT 25 General Translation Task</i>	
Felipe Fujita and Hideyuki Takada	318
<i>Multi-agentMT: Deploying AI Agent in the WMT25 Shared Task</i>	
Ahrii Kim	322
<i>Lanigo at WMT25 General Translation Task: Self-Improved and Retrieval-Augmented Translation</i>	
Kamil Guttman, Zofia Rostek, Adrian Charkiewicz, Antoni Solarski, Mikołaj Pokrywka and Artur Nowakowski	331

<i>Command-A-Translate: Raising the Bar of Machine Translation with Difficulty Filtering</i>	
Tom Kocmi, Arkady Arkhangorodsky, Alexandre Berard, Phil Blunsom, Samuel Cahyawijaya, Théo Dehaze, Marzieh Fadaee, Nicholas Frosst, Matthias Galle, Aidan Gomez, Nithya Govindarajan, Wei-Yin Ko, Julia Kreutzer, Kelly Marchisio, Ahmet Üstün, Sebastian Vincent and Ivan Zhang	342
<i>GENDERPERSON: Test Suite for Estimating Gender Bias of First-person Singular Forms</i>	
Maja Popović and Ekaterina Lapshinova-Koltunski	353
<i>Evaluation of LLM for English to Hindi Legal Domain Machine Translation Systems</i>	
Kshetrimayum Boynao Singh, Deepak Kumar and Asif Ekbal	376
<i>RoCS-MT v2 at WMT 2025: Robust Challenge Set for Machine Translation</i>	
Rachel Bawden and Benoît Sagot	387
<i>Automated Evaluation for Terminology Translation Related to the EEA Agreement</i>	
Selma Dis Hauksdottir and Steinthor Steingrímsson	403
<i>Up to Par? MT Systems Take a Shot at Sports Terminology</i>	
Einar Sigurdsson, Magnús Már Magnússon, Atli Jasonarson and Steinthor Steingrímsson	409
<i>Fine-Grained Evaluation of English-Russian MT in 2025: Linguistic Challenges Mirroring Human Translator Training</i>	
Shushen Manakhimova, Maria Kunilovskaya, Ekaterina Lapshinova-Koltunski and Eleftherios Avramidis	419
<i>Tagged Span Annotation for Detecting Translation Errors in Reasoning LLMs</i>	
Taemin Yeom, Yonghyun Ryu, Yoonjung Choi and JinYeong Bak	431
<i>COMET-poly: Machine Translation Metric Grounded in Other Candidates</i>	
Maike Züfle, Vilém Zouhar, Tu Anh Dinh, Felipe Maia Polo, Jan Niehues and Mrinmaya Sachan	440
<i>Long-context Reference-based MT Quality Estimation</i>	
Sami Ul Haq, Chinonso Cynthia Osuji, Sheila Castilho, Brian Davis and Thiago Castro Ferreira	458
<i>Evaluating WMT 2025 Metrics Shared Task Submissions on the SSA-MTE African Challenge Set</i>	
Senyu Li, Felermينو Dario Mario Ali, Jiayi Wang, Rui Sousa-Silva, Henrique Lopes Cardoso, Pontus Stenetorp, Colin Cherry and David Ifeoluwa Adelani	466
<i>Nvidia-Nemo's WMT 2025 Metrics Shared Task Submission</i>	
Brian Yan, Shuoyang Ding, Kuang-Da Wang, Siqi Ouyang, Oleksii Hrinchuk, Vitaly Lavrukhin and Boris Ginsburg	473
<i>GEMBA V2: Ten Judgments Are Better Than One</i>	
Marcin Junczys-Dowmunt	479
<i>CUNI and Phrase at WMT25 MT Evaluation Task</i>	
Miroslav Hrabal, Ondrej Glembek, Aleš Tamchyna, Almut Silja Hildebrand, Alan Eckhard, Miroslav Štola, Sergio Penkale, Zuzana Šimečková, Ondřej Bojar, Alon Lavie and Craig Stewart	487
<i>MSLC25: Metric Performance on Low-Quality Machine Translation, Empty Strings, and Language Variants</i>	
Rebecca Knowles, Samuel Larkin and Chi-kiu Lo	498

<i>MetricX-25 and GemSpanEval: Google Translate Submissions to the WMT25 Evaluation Shared Task</i>	
Juraj Juraska, Tobias Domhan, Mara Finkelstein, Tetsuji Nakagawa, Geza Kovacs, Daniel Deutsch, Pidong Wang and Markus Freitag	510
<i>HW-TSC's Submissions to the WMT 2025 Segment-level Quality Score Prediction Task</i>	
Yuanchang Luo, Jiaxin GUO, Daimeng Wei, Hengchao Shang, Zongyao Li, Zhiqiang Rao, Jinlong Yang, Zhanglin Wu, xiaoyu chen and Hao Yang	522
<i>UvA-MT at WMT25 Evaluation Task: LLM Uncertainty as a Proxy for Translation Quality</i>	
Di Wu and Christof Monz	527
<i>Submission for WMT25 Task 3</i>	
Govardhan Padmanabhan	537
<i>RankedCOMET: Elevating a 2022 Baseline to a Top-5 Finish in the WMT 2025 QE Task</i>	
Sujal Maharjan and Astha Shrestha	547
<i>Quality-Informed Segment-Level Error Correction Using Natural Language Explanations from xTower and Large Language Models</i>	
Prashant K. Sharma	552
<i>TASER: Translation Assessment via Systematic Evaluation and Reasoning</i>	
Monishwaran Maheswaran, Marco Carini, Christian Federmann and Tony Diaz	557
<i>An Empirical Analysis of Machine Translation for Expanding Multilingual Benchmarks</i>	
Sara Rajae, Rochelle Choenni, Ekaterina Shutova and Christof Monz	564
<i>Cross-lingual Human-Preference Alignment for Neural Machine Translation with Direct Quality Optimization</i>	
Kaden Uhlig, Joern Wuebker, Raphael Reinauer and John DeNero	594
<i>Audio-Based Crowd-Sourced Evaluation of Machine Translation Quality</i>	
Sami Ul Haq, Sheila Castilho and Yvette Graham	615
<i>Meaningful Pose-Based Sign Language Evaluation</i>	
Zifan Jiang, Colin Leong, Amit Moryossef, Oliver Cory, Maksym Ivashechkin, Neha Tarigopula, Biao Zhang, Anne Göhring, Annette Rios, Rico Sennrich and Sarah Ebling	627
<i>Context Is Ubiquitous, but Rarely Changes Judgments: Revisiting Document-Level MT Evaluation</i>	
Ahrii Kim	644
<i>GIIFT: Graph-guided Inductive Image-free Multimodal Machine Translation</i>	
Jiafeng Xiong and Yuting Zhao	661
<i>Findings of the WMT 2025 Shared Task on Model Compression: Early Insights on Compressing LLMs for Machine Translation</i>	
Marco Gaido, Roman Grundkiewicz, Thamme Gowda and Matteo Negri	676
<i>Findings of the WMT 2025 Shared Task of the Open Language Data Initiative</i>	
Laurie V. Burchell, David Dale, Jean Maillard, Idris Abdulmumin, Antonios Anastasopoulos, Isaac Caswell and Philipp Koehn	687

<i>Findings of the WMT 2025 Shared Task LLMs with Limited Resources for Slavic Languages: MT and QA</i>	
Shu Okabe, Daryna Dementieva, Marion Di Marco, Lukas Edman, Katharina Haemmerl, Marko Měškank, Anita Hendrichowa and Alexander Fraser	695
<i>Findings of the First Shared Task for Creole Language Machine Translation at WMT25</i>	
Nathaniel Romney Robinson, Claire Bizon Monroc, Rasul Dent, Stefan Watson, Kenton Murray, Raj Dabre, Andre Coy and Heather Lent	712
<i>Findings of WMT 2025 Shared Task on Low-resource Indic Languages Translation</i>	
Partha Pakray, Reddi Mohana Krishna, Santanu Pal, Advaita Vetagiri, Sandeep Kumar Dash, Arnab Kumar Maji, Saralin A. Lyngdoh, Lenin Laitonjam, Anupam Jamatia, Koj Sambyo, Ajit Das and Riyanka Manna	724
<i>Findings of the WMT25 Terminology Translation Task: Terminology is Useful Especially for Good MTs</i>	
Kirill Semenov, Xu Huang, Vilém Zouhar, Nathaniel Berger, Dawei Zhu, Arturo Oncevay and Pinzhen Chen	746
<i>Vicomtech@WMT 2025: Evolutionary Model Compression for Machine Translation</i>	
David Ponce, Harritxu Gete and Thierry Etchegoyhen	769
<i>Iterative Layer Pruning for Efficient Translation Inference</i>	
Yasmin Moslem, Muhammad Hazim Al Farouq and John D. Kelleher	780
<i>Expanding the WMT24++ Benchmark with Rumantsch Grischun, Sursilvan, Sutsilvan, Surmiran, Puter, and Vallader</i>	
Jannis Vamvas, Ignacio Pérez Prat, Not Battesta Soliva, Sandra Baltermia-Guetg, Andrina Beeli, Simona Beeli, Madlaina Capeder, Laura Decurtins, Gian Peder Gregori, Flavia Hobi, Gabriela Holderegger, Arina Lazzarini, Viviana Lazzarini, Walter Rosselli, Bettina Vital, Anna Rutkiewicz and Rico Senrich	786
<i>A French Version of the OLDI Seed Corpus</i>	
Malik Marmonier, Benoît Sagot and Rachel Bawden	806
<i>Bringing Ladin to FLORES+</i>	
Samuel Frontull, Thomas Ströhle, Carlo Zoli, Werner Pescosta, Ulrike Frenademez, Matteo Ruggeri, Daria Valentin, Karin Comploj, Gabriel Perathoner, Silvia Liotto and Paolo Anvidalfarei	819
<i>Correcting the Tamazight Portions of FLORES+ and OLDI Seed Datasets</i>	
Alp Oktem, Mohamed Aymane Farhi, Brahim Essaidi, Naceur Jabouja and Farida Boudichat	830
<i>Filling the Gap for Uzbek: Creating Translation Resources for Southern Uzbek</i>	
Mukhammadsaid Mamasaidov, Azizullah Aral, Abror Shopulatov and Mironshoh Inomjonov	839
<i>The Kyrgyz Seed Dataset Submission to the WMT25 Open Language Data Initiative Shared Task</i>	
Murat Jumashev, Alina Tillabaeva, Aida Kasieva, Turgunbek Omurkanov, Akylai Musaeva, Meerim Emil kyzy, Gulaiym Chagataeva and Jonathan North Washington	846
<i>SMOL: Professionally Translated Parallel Data for 115 Under-represented Languages</i>	
Isaac Caswell, Elizabeth Nielsen, Jiaming Luo, Colin Cherry, Geza Kovacs, Hadar Shemtov, Partha Talukdar, Dinesh Tewari, Moussa Koulako Bala Doumbouya, Djibrila Diane, Baba Mamadi Diane, Solo Farabado, Edoardo Ferrante, Alessandro Guasoni, Mamadou K. Keita, Sudhamoy DebBarma, Ali Kuzhuget, David Anugraha, Muhammad Ravi Shulthan Habibi, Sina Ahmadi, Mingfei Liu and Jonathan Eng	861

<i>Improved Norwegian Bokmål Translations for FLORES</i>	
Petter Mæhlum, Anders Næss Evensen and Yves Scherrer	882
<i>NRC Systems for the WMT2025-LRSL Shared Task</i>	
Samuel Larkin, Chi-kiu Lo and Rebecca Knowles	891
<i>TartuNLP at WMT25 LLMs with Limited Resources for Slavic Languages Shared Task</i>	
Taido Purason and Mark Fishel	901
<i>JGU Mainz's Submission to the WMT25 Shared Task on LLMs with Limited Resources for Slavic Languages: MT and QA</i>	
Hossain Shaikh Saadi, Minh Duc Bui, Mario Sanz-Guerrero and Katharina von der Wense	909
<i>Krey-All WMT 2025 CreoleMT System Description: Language Agnostic Strategies for Low-Resource Translation</i>	
Ananya Ayasi	916
<i>EdinHelsOW WMT 2025 CreoleMT System Description: Improving Lusophone Creole Translation through Data Augmentation, Model Merging and LLM Post-editing</i>	
Jacqueline Rowe, Ona de Gibert, Mateusz Klimaszewski, Coleman Haley, Alexandra Birch and Yves Scherrer	924
<i>KozKreolMRU WMT 2025 CreoleMT System Description: Koz Kreol: Multi-Stage Training for English–Mauritian Creole MT</i>	
Yush Rajcoomar	941
<i>JHU WMT 2025 CreoleMT System Description: Data for Belizean Kriol and French Guianese Creole MT</i>	
Nathaniel Romney Robinson	949
<i>WMT 2025 CreoleMT Systems Description : Martinican Creole and French</i>	
Ludovic V. Mompelat	956
<i>JU-NLP: Improving Low-Resource Indic Translation System with Efficient LoRA-Based Adaptation</i>	
Priyobroto Acharya, Haranath Mondal, Dipanjan Saha, Dipankar Das and Sivaji Bandyopadhyay	959
<i>An Attention-Based Neural Translation System for English to Bodo</i>	
Subhash Kumar Wary, Birhang Borgoyary, Akher Uddin Ahmed, Mohanji Prasad Sah and Apur- balal Senapati	968
<i>Tackling Low-Resource NMT with Instruction-Tuned LLaMA: A Study on Kokborok and Bodo</i>	
Deepak Kumar, Kshetrimayum Boynao Singh and Asif Ekbal	973
<i>DELAB-IIITM WMT25: Enhancing Low-Resource Machine Translation for Manipuri and Assamese</i>	
Dingku Singh Oinam and Navanath Saharia	980
<i>Transformers: Leveraging OpenNMT and Transfer Learning for Low-Resource Indian Language Trans- lation</i>	
Bhagyashree Wagh, Harish Bapat, Neha Gupta and Saurabh Salunkhe	985
<i>RBG-AI: Benefits of Multilingual Language Models for Low-Resource Languages</i>	
Barathi Ganesh HB and Michal Ptaszynski	991

<i>ANVITA : A Multi-pronged Approach for Enhancing Machine Translation of Extremely Low-Resource Indian Languages</i>	
SIVABHAVANI J, Daneshwari Kankanwadi, Abhinav Mishra and Biswajit Paul	998
<i>DoDS-IITPKD:Submissions to the WMT25 Low-Resource Indic Language Translation Task</i>	
Ontiwell Khongthaw, G.L. John Salvin, Shrikant Tryambak Budde, Abigail Nyasha Chigweddedza, Dhruvadeep Malkar and Swapnil Hingmire	1006
<i>A Preliminary Exploration of Phrase-Based SMT and Multi-BPE Segmentations through Concatenated Tokenised Corpora for Low-Resource Indian Languages</i>	
Saumitra Yadav and Manish Shrivastava	1011
<i>AkibaNLP-TUT: Injecting Language-Specific Word-Level Noise for Low-Resource Language Translation</i>	
Shoki Hamada, Tomoyosi Akiba and Hajime Tsukada	1017
<i>BVSLP: Machine Translation Using Linguistic Embellishments for IndicMT Shared Task 2025</i>	
Nisheeth Joshi, Palak Arora, Anju Krishnia, Riya Lonchenpa and Mhasilenuo Vizo	1023
<i>TranssionMT's Submission to the Indic MT Shared Task in WMT 2025</i>	
zebiao zhou, hui li, xiangxun zhu and kangzhen liu	1029
<i>Laniqo at WMT25 Terminology Translation Task: A Multi-Objective Reranking Strategy for Terminology-Aware Translation via Pareto-Optimal Decoding</i>	
Kamil Gutmman, Adrian Charkiewicz, Zofia Rostek, Mikołaj Pokrywka and Artur Nowakowski	1034
<i>Fine-tuning NMT Models and LLMs for Specialised EN-ES Translation Using Aligned Corpora, Glossaries, and Synthetic Data: MULTITAN at WMT25 Terminology Shared Task</i>	
Lichao Zhu, Maria Zimina-Poirot, Stephane Patin and Cristian Valdez	1042
<i>Contextual Selection of Pseudo-terminology Constraints for Terminology-aware Neural Machine Translation in the IT Domain</i>	
Benjamin Pong	1050
<i>IRB-MT at WMT25 Terminology Translation Task: Metric-guided Multi-agent Approach</i>	
Ivan Grubišić and Damir Korencic	1060
<i>Terminology-Constrained Translation from Monolingual Data Using GRPO</i>	
Javier Garcia Gilabert, Carlos Escolano, Xixian Liao and Maite Melero	1093
<i>It Takes Two: A Dual Stage Approach for Terminology-Aware Translation</i>	
Akshat Singh Jaswal	1102
<i>Specification-Aware Machine Translation and Evaluation for Purpose Alignment</i>	
Yoko Kayano and Saku Sugawara	1109
<i>OpenWHO: A Document-Level Parallel Corpus for Health Translation in Low-Resource Languages</i>	
Raphael Merx, Hanna Suominen, Trevor Cohn and Ekaterina Vylomova	1138
<i>Factors Affecting Translation Quality in In-context Learning for Multilingual Medical Domain</i>	
Jonathan Mutal, Raphael Rubino and Pierrette Bouillon	1157
<i>Character-Aware English-to-Japanese Translation of Fictional Dialogue Using Speaker Embeddings and Back-Translation</i>	
Ayuna Nagato and Takuya Matsuzaki	1176

<i>DTW-Align: Bridging the Modality Gap in End-to-End Speech Translation with Dynamic Time Warping Alignment</i>	
Abderrahmane Issam, Yusuf Can Semerci, Jan Scholtes and Gerasimos Spanakis	1187
<i>Targeted Source Text Editing for Machine Translation: Exploiting Quality Estimators and Large Language Models</i>	
Hyuga Koretaka, Atsushi Fujita and Tomoyuki Kajiwara	1196
<i>Self-Retrieval from Distant Contexts for Document-Level Machine Translation</i>	
Ziqian Peng, Rachel Bawden and François Yvon	1216
<i>Using Encipherment to Isolate Conditions for the Successful Fine-tuning of Massively Multilingual Translation Models</i>	
Carter Louchheim, Denis Sotnichenko, Yukina Yamaguchi and Mark Hopkins	1237
<i>Translate, Then Detect: Leveraging Machine Translation for Cross-Lingual Toxicity Classification</i>	
Samuel J. Bell, Eduardo Sánchez, David Dale, Pontus Stenetorp, Mikel Artetxe and Marta R. Costajussà	1249
<i>Feeding Two Birds or Favoring One? Adequacy–Fluency Tradeoffs in Evaluation and Meta-Evaluation of Machine Translation</i>	
Behzad Shayegh, Jan-Thorsten Peter, David Vilar, Tobias Domhan, Juraj Juraska, Markus Freitag and Lili Mou	1265
<i>DocHPLT: A Massively Multilingual Document-Level Translation Dataset</i>	
Dayyán O’Brien, Bhavitvya Malik, Ona de Gibert, Pinzhen Chen, Barry Haddow and Jörg Tiedemann	1282
<i>SONAR-SLT: Multilingual Sign Language Translation via Language-Agnostic Sentence Embedding Supervision</i>	
Yasser HAMIDULLAH, Shakib Yazdani, Cennet Oguz, Josef van Genabith and Cristina España-Bonet	1297
<i>GAMBIT+: A Challenge Set for Evaluating Gender Bias in Machine Translation Quality Estimation Metrics</i>	
George Filandrianos, Orfeas Menis Mastromichalakis, Wafaa Mohammed, Giuseppe Attanasio and Chrysoula Zerva	1310
<i>Implementing and Evaluating Multi-source Retrieval-Augmented Translation</i>	
Tommi Nieminen, Jörg Tiedemann and Sami Virpioja	1323
<i>A Cross-Lingual Perspective on Neural Machine Translation Difficulty</i>	
Esther Ploeger, Johannes Bjerva, Jörg Tiedemann and Robert Oestling	1336

Conference Program

Saturday, November 8, 2025

9:00–9:10 *Opening Remarks*

9:10–10:30 **Session 1: Shared Task Overview Papers I**

9:10–9:40 *Findings of the WMT25 General Machine Translation Shared Task: Time to Stop Evaluating on Easy Test Sets*

Tom Kocmi, Ekaterina Artemova, Eleftherios Avramidis, Rachel Bawden, Ondřej Bojar, Konstantin Dranch, Anton Dvorkovich, Sergey Dukanov, Mark Fishel, Markus Freitag, Thamme Gowda, Roman Grundkiewicz, Barry Haddow, Marzena Karpinska, Philipp Koehn, Howard Lakouгна, Jessica Lundin, Christof Monz, Kenton Murray, Masaaki Nagata, Stefano Perrella, Lorenzo Proietti, Martin Popel, Maja Popović, Parker Riley, Mariya Shmatova, Steinþór Steingrímsson, Lisa Yankovskaya and Vilém Zouhar

9:40–10:00 *Findings of the WMT25 Shared Task on Automated Translation Evaluation Systems: Linguistic Diversity is Challenging and References Still Help*

Alon Lavie, Greg Hanneman, Sweta Agrawal, Diptesh Kanojia, Chi-kiu Lo, Vilém Zouhar, Frederic Blain, Chrysoula Zerva, Eleftherios Avramidis, Sourabh Dattatray Deoghare, Archchana Sindhuja, Jiayi Wang, David Ifeoluwa Adelani, Brian Thompson, Tom Kocmi, Markus Freitag and Daniel Deutsch

10:10–10:30 *Findings of the WMT25 Multilingual Instruction Shared Task: Persistent Hurdles in Reasoning, Generation, and Evaluation*

Tom Kocmi, Sweta Agrawal, Ekaterina Artemova, Eleftherios Avramidis, Eleftheria Briakou, Pinzhen Chen, Marzieh Fadaee, Markus Freitag, Roman Grundkiewicz, Yupeng Hou, Philipp Koehn, Julia Kreutzer, Saab Mansour, Stefano Perrella, Lorenzo Proietti, Parker Riley, Eduardo Sánchez, Patricia Schmidtova, Mariya Shmatova and Vilém Zouhar

10:30–11:00 *Coffee Break*

11:00–12:00 **Session 2: Shared Task Posters**

11:00–12:00 **General Translation Task**

11:00–12:00 *Miðeind at WMT25 General Machine Translation Task*

Svanhvít Lilja Ingólfssdóttir, Haukur Páll Jónsson, Kári Steinn Aðalsteinsson, Róbert Fjölfnir Birkisson, Sveinbjörn Þórðarson and Þorvaldur Páll Helgason

11:00–12:00 *A Preliminary Study of AI Agent Model in Machine Translation*

Ahrii Kim

11:00–12:00 *Marco Large Translation Model at WMT2025: Transforming Translation Capability in LLMs via Quality-Aware Training and Decoding*

Hao Wang, Linlong Xu, Heng Liu, Yangyang Liu, Xiaohu Zhao, Bo Zeng, Longyue Wang, Weihua Luo and Kaifu Zhang

11:00–12:00 *Evaluation of OWEN-3 for English to Ukrainian Translation*

Saturday, November 8, 2025 (continued)

- 11:00–12:00 *SYSTRAN @ WMT 2025 General Translation Task*
Dakun Zhang, Yara Khater, Ramzi Rahli, Anna Rebollo and Josep Crego
- 11:00–12:00 *Shy-hunyuan-MT at WMT25 General Machine Translation Shared Task*
Mao Zheng, Zheng Li, Yang Du, Bingxin Qu and Mingyang Song
- 11:00–12:00 *From SALAMANDRA to SALAMANDRATA: BSC Submission for WMT25 General Machine Translation Shared Task*
Javier Garcia Gilabert, Xixian Liao, Severino Da Dalt, Ella Bohman, Audrey Mash, Francesca De Luca Fornaciari, Irene Baucells, Joan Llop, Miguel Claramunt, Carlos Escolano and Maite Melero
- 11:00–12:00 *Instruction-Tuned English to Bhojpuri Neural Machine Translation Using Contrastive Preference Optimization*
Kshetrimayum Boynao Singh, Deepak Kumar and Asif Ekbal
- 11:00–12:00 *SH at WMT25 General Machine Translation Task*
Hayate Shiroma
- 11:00–12:00 *Simple Test Time Scaling for Machine Translation: Kaze-MT at the WMT25 General Translation Task*
Shaomu Tan
- 11:00–12:00 *NTTSU at WMT2025 General Translation Task*
Zhang Yin, Hiroyuki Deguchi, Haruto Azami, Guanyu Ouyang, Kosei Buma, YINGYI FU, Katsuki Chousa and Takehito Utsuro
- 11:00–12:00 *A* Decoding for Machine Translation in LLMs - SRPOL Participation in WMT2025*
Adam Dobrowolski, Paweł Przewłocki, Paweł Przybysz, Marcin Szymański and Dawid Siwicki
- 11:00–12:00 *In2X at WMT25 Translation Task*
Lei Pang, Hanyi Mao, Quanjia Xiao, Chen Ruihan, Jingjun Zhang, Haixiao Liu and Xiangyi Li
- 11:00–12:00 *CUNI at WMT25 General Translation Task*
Josef Jon, Miroslav Hrabal, Martin Popel and Ondřej Bojar
- 11:00–12:00 *UvA-MT's Participation in the WMT25 General Translation Shared Task*
Di Wu, Yan Meng, Maya Konstantinovna Nachesa, Seth Aycock and Christof Monz
- 11:00–12:00 *AMI at WMT25 General Translation Task: How Low Can We Go? Finetuning Lightweight Llama Models for Low Resource Machine Translation*
Atli Jasonarson and Steinthor Steingrímsson

Saturday, November 8, 2025 (continued)

- 11:00–12:00 *KIKIS at WMT 2025 General Translation Task*
Koichi Iwakawa, Keito Kudo, Subaru Kimura, Takumi Ito and Jun Suzuki
- 11:00–12:00 *Google Translate’s Research Submission to WMT2025*
Mara Finkelstein, Geza Kovacs, Isaac Caswell, Tobias Domhan, Jan-Thorsten Peter, Juraj Juraska, Markus Freitag and David Vilar
- 11:00–12:00 *DLUT and GTCOM’s Large Language Model Based Translation System for WMT25*
Hao Zong, Chao Bei, Wentao Chen, Conghu Yuan, Huan Liu and Degen Huang
- 11:00–12:00 *Yandex Submission to the WMT25 General Machine Translation Task*
Nikolay Karpachev, Ekaterina Enikeeva, Dmitry Popov, Arsenii Bulgakov, Daniil Panteleev, Dmitrii Ulianov, Artem Kryukov and Artem Mekhraliev
- 11:00–12:00 *IRB-MT at WMT25 Translation Task: A Simple Agentic System Using an Off-the-Shelf LLM*
Ivan Grubišić and Damir Korencic
- 11:00–12:00 *Improving Low-Resource Japanese Translation with Fine-Tuning and Backtranslation for the WMT 25 General Translation Task*
Felipe Fujita and Hideyuki Takada
- 11:00–12:00 *Multi-agentMT: Deploying AI Agent in the WMT25 Shared Task*
Ahrii Kim
- 11:00–12:00 *Lanigo at WMT25 General Translation Task: Self-Improved and Retrieval-Augmented Translation*
Kamil Guttman, Zofia Rostek, Adrian Charkiewicz, Antoni Solarski, Mikołaj Pokrywka and Artur Nowakowski
- 11:00–12:00 *Command-A-Translate: Raising the Bar of Machine Translation with Difficulty Filtering*
Tom Kocmi, Arkady Arkhangorodsky, Alexandre Berard, Phil Blunsom, Samuel Cahyawijaya, Théo Dehaze, Marzieh Fadaee, Nicholas Frosst, Matthias Galle, Aidan Gomez, Nithya Govindarajan, Wei-Yin Ko, Julia Kreutzer, Kelly Marchisio, Ahmet Üstün, Sebastian Vincent and Ivan Zhang

Saturday, November 8, 2025 (continued)

11:00–12:00 MT Test Suites Shared Task

11:00–12:00 *GENDER1PERSON: Test Suite for Estimating Gender Bias of First-person Singular Forms*
Maja Popović and Ekaterina Lapshinova-Koltunski

11:00–12:00 *Evaluation of LLM for English to Hindi Legal Domain Machine Translation Systems*
Kshetrimayum Boynao Singh, Deepak Kumar and Asif Ekbal

11:00–12:00 *RoCS-MT v2 at WMT 2025: Robust Challenge Set for Machine Translation*
Rachel Bawden and Benoît Sagot

11:00–12:00 *Automated Evaluation for Terminology Translation Related to the EEA Agreement*
Selma Dis Hauksdottir and Steinthor Steingrímsson

11:00–12:00 *Up to Par? MT Systems Take a Shot at Sports Terminology*
Einar Sigurdsson, Magnús Már Magnússon, Atli Jasonarson and Steinthor Steingrímsson

11:00–12:00 *Fine-Grained Evaluation of English-Russian MT in 2025: Linguistic Challenges Mirroring Human Translator Training*
Shushen Manakhimova, Maria Kunilovskaya, Ekaterina Lapshinova-Koltunski and Eleftherios Avramidis

11:00–12:00 Translation Quality Evaluation Shared Task

11:00–12:00 *Tagged Span Annotation for Detecting Translation Errors in Reasoning LLMs*
Taemin Yeom, Yonghyun Ryu, Yoonjung Choi and JinYeong Bak

11:00–12:00 *COMET-poly: Machine Translation Metric Grounded in Other Candidates*
Maike Züfle, Vilém Zouhar, Tu Anh Dinh, Felipe Maia Polo, Jan Niehues and Mrinmaya Sachan

11:00–12:00 *Long-context Reference-based MT Quality Estimation*
Sami Ul Haq, Chinonso Cynthia Osuji, Sheila Castilho, Brian Davis and Thiago Castro Ferreira

11:00–12:00 *Evaluating WMT 2025 Metrics Shared Task Submissions on the SSA-MTE African Challenge Set*
Senyu Li, Felermينو Dario Mario Ali, Jiayi Wang, Rui Sousa-Silva, Henrique Lopes Cardoso, Pontus Stenetorp, Colin Cherry and David Ifeoluwa Adelani

Saturday, November 8, 2025 (continued)

- 11:00–12:00 *Nvidia-Nemo’s WMT 2025 Metrics Shared Task Submission*
Brian Yan, Shuoyang Ding, Kuang-Da Wang, Siqi Ouyang, Oleksii Hrinchuk, Vitaly Lavrukhin and Boris Ginsburg
- 11:00–12:00 *GEMBA V2: Ten Judgments Are Better Than One*
Marcin Junczys-Dowmunt
- 11:00–12:00 *CUNI and Phrase at WMT25 MT Evaluation Task*
Miroslav Hrabal, Ondrej Glembek, Aleš Tamchyna, Almut Silja Hildebrand, Alan Eckhard, Miroslav Štola, Sergio Penkale, Zuzana Šimečková, Ondřej Bojar, Alon Lavie and Craig Stewart
- 11:00–12:00 *MSLC25: Metric Performance on Low-Quality Machine Translation, Empty Strings, and Language Variants*
Rebecca Knowles, Samuel Larkin and Chi-kiu Lo
- 11:00–12:00 *MetricX-25 and GemSpanEval: Google Translate Submissions to the WMT25 Evaluation Shared Task*
Juraj Juraska, Tobias Domhan, Mara Finkelstein, Tetsuji Nakagawa, Geza Kovacs, Daniel Deutsch, Pidong Wang and Markus Freitag
- 11:00–12:00 *HW-TSC’s Submissions to the WMT 2025 Segment-level Quality Score Prediction Task*
Yuanchang Luo, Jiaxin GUO, Daimeng Wei, Hengchao Shang, Zongyao Li, Zhiqiang Rao, Jinlong Yang, Zhanglin Wu, xiaoyu chen and Hao Yang
- 11:00–12:00 *UvA-MT at WMT25 Evaluation Task: LLM Uncertainty as a Proxy for Translation Quality*
Di Wu and Christof Monz
- 11:00–12:00 *Submission for WMT25 Task 3*
Govardhan Padmanabhan
- 11:00–12:00 *RankedCOMET: Elevating a 2022 Baseline to a Top-5 Finish in the WMT 2025 QE Task*
Sujal Maharjan and Astha Shrestha
- 11:00–12:00 *Quality-Informed Segment-Level Error Correction Using Natural Language Explanations from xTower and Large Language Models*
Prashant K. Sharma
- 11:00–12:00 *TASER: Translation Assessment via Systematic Evaluation and Reasoning*
Monishwaran Maheswaran, Marco Carini, Christian Federmann and Tony Diaz
- 12:00–14:00 *Lunch Break*

Saturday, November 8, 2025 (continued)

14:00–15:00 Session 3: Invited Talk by Longyue Wang

15:00–16:00 *Coffee Break*

16:00–17:30 Session 4: Featured Research Papers Oral Presentations

16:00–16:15 *An Empirical Analysis of Machine Translation for Expanding Multilingual Benchmarks*

Sara Rajaei, Rochelle Choenni, Ekaterina Shutova and Christof Monz

16:15–16:30 *Cross-lingual Human-Preference Alignment for Neural Machine Translation with Direct Quality Optimization*

Kaden Uhlig, Joern Wuebker, Raphael Reinauer and John DeNero

16:30–16:45 *Audio-Based Crowd-Sourced Evaluation of Machine Translation Quality*

Sami Ul Haq, Sheila Castilho and Yvette Graham

16:45–17:00 *Meaningful Pose-Based Sign Language Evaluation*

Zifan Jiang, Colin Leong, Amit Moryossef, Oliver Cory, Maksym Ivashechkin, Neha Tarigopula, Biao Zhang, Anne Göhring, Annette Rios, Rico Sennrich and Sarah Ebling

17:00–17:15 *Context Is Ubiquitous, but Rarely Changes Judgments: Revisiting Document-Level MT Evaluation*

Ahrii Kim

17:15–17:30 *GIIFT: Graph-guided Inductive Image-free Multimodal Machine Translation*

Jiafeng Xiong and Yuting Zhao

Sunday, November 9, 2025

9:00–10:30 Session 5: Shared Task Overview Papers II

9:00–9:15 *Findings of the WMT 2025 Shared Task on Model Compression: Early Insights on Compressing LLMs for Machine Translation*
Marco Gaido, Roman Grundkiewicz, Thamme Gowda and Matteo Negri

9:15–9:30 *Findings of the WMT 2025 Shared Task of the Open Language Data Initiative*
Laurie V. Burchell, David Dale, Jean Maillard, Idris Abdulmumin, Antonios Anastopoulos, Isaac Caswell and Philipp Koehn

9:30–9:45 *Findings of the WMT 2025 Shared Task LLMs with Limited Resources for Slavic Languages: MT and QA*
Shu Okabe, Daryna Dementieva, Marion Di Marco, Lukas Edman, Katharina Haemmerl, Marko Měškank, Anita Hendrichowa and Alexander Fraser

9:45–10:00 *Findings of the First Shared Task for Creole Language Machine Translation at WMT25*
Nathaniel Romney Robinson, Claire Bizon Monroc, Rasul Dent, Stefan Watson, Kenton Murray, Raj Dabre, Andre Coy and Heather Lent

10:00–10:15 *Findings of WMT 2025 Shared Task on Low-resource Indic Languages Translation*
Partha Pakray, Reddi Mohana Krishna, Santanu Pal, Advaita Vetagiri, Sandeep Kumar Dash, Arnab Kumar Maji, Saralin A. Lyngdoh, Lenin Laitonjam, Anupam Jamatia, Koj Sambyo, Ajit Das and Riyanka Manna

10:15–10:30 *Findings of the WMT25 Terminology Translation Task: Terminology is Useful Especially for Good MTs*
Kirill Semenov, Xu Huang, Vilém Zouhar, Nathaniel Berger, Dawei Zhu, Arturo Oncevay and Pinzhen Chen

11:00–12:00 Session 6: Shared Task Posters

Sunday, November 9, 2025 (continued)

11:00–12:00 Model Compression Shared Task

11:00–12:00 *Vicomtech@WMT 2025: Evolutionary Model Compression for Machine Translation*
David Ponce, Harritxu Gete and Thierry Etchegoyhen

11:00–12:00 *Iterative Layer Pruning for Efficient Translation Inference*
Yasmin Moslem, Muhammad Hazim Al Farouq and John D. Kelleher

11:00–12:00 Open Language Data Initiative

11:00–12:00 *Expanding the WMT24++ Benchmark with Rumantsch Grischun, Sursilvan, Sutsilvan, Surmiran, Puter, and Vallader*
Jannis Vamvas, Ignacio Pérez Prat, Not Battesta Soliva, Sandra Baltermia-Guetg, Andrina Beeli, Simona Beeli, Madlaina Capeder, Laura Decurtins, Gian Peder Gregori, Flavia Hobi, Gabriela Holderegger, Arina Lazzarini, Viviana Lazzarini, Walter Rosselli, Bettina Vital, Anna Rutkiewicz and Rico Sennrich

11:00–12:00 *A French Version of the OLDI Seed Corpus*
Malik Marmonier, Benoît Sagot and Rachel Bawden

11:00–12:00 *Bringing Ladin to FLORES+*
Samuel Frontull, Thomas Ströhle, Carlo Zoli, Werner Pescosta, Ulrike Frenademez, Matteo Ruggeri, Daria Valentin, Karin Comploj, Gabriel Perathoner, Silvia Liotto and Paolo Anvidalfarei

11:00–12:00 *Correcting the Tamazight Portions of FLORES+ and OLDI Seed Datasets*
Alp Oktem, Mohamed Aymane Farhi, Brahim Essaidi, Naceur Jabouja and Farida Boudichat

11:00–12:00 *Filling the Gap for Uzbek: Creating Translation Resources for Southern Uzbek*
Mukhammadsaid Mamasaidov, Azizullah Aral, Abror Shopulatov and Mironshoh Inomjonov

11:00–12:00 *The Kyrgyz Seed Dataset Submission to the WMT25 Open Language Data Initiative Shared Task*
Murat Jumashev, Alina Tillabaeva, Aida Kasieva, Turgunbek Omurkanov, Akylai Musaeva, Meerim Emil kyzy, Gulaiym Chagataeva and Jonathan North Washington

11:00–12:00 *SMOL: Professionally Translated Parallel Data for 115 Under-represented Languages*
Isaac Caswell, Elizabeth Nielsen, Jiaming Luo, Colin Cherry, Geza Kovacs, Hadar Shemtov, Partha Talukdar, Dinesh Tewari, Moussa Koulako Bala Doumbouya, Djibrila Diane, Baba Mamadi Diane, Solo Farabado, Edoardo Ferrante, Alessandro Guasconi, Mamadou K. Keita, Sudhamoy DebBarma, Ali Kuzhuget, David Anugraha, Muhammad Ravi Shulthan Habibi, Sina Ahmadi, Mingfei Liu and Jonathan Eng

11:00–12:00 *Improved Norwegian Bokmål Translations for FLORES*
Petter Mæhlum, Anders Næss Evensen and Yves Scherrer

Sunday, November 9, 2025 (continued)

11:00–12:00 Slavic Languages Shared Task

11:00–12:00 *NRC Systems for the WMT2025-LRSL Shared Task*

Samuel Larkin, Chi-kiu Lo and Rebecca Knowles

11:00–12:00 *TartuNLP at WMT25 LLMs with Limited Resources for Slavic Languages Shared Task*

Taido Purason and Mark Fishel

11:00–12:00 *JGU Mainz’s Submission to the WMT25 Shared Task on LLMs with Limited Resources for Slavic Languages: MT and QA*

Hossain Shaikh Saadi, Minh Duc Bui, Mario Sanz-Guerrero and Katharina von der Wense

11:00–12:00 Creole Language Translation Shared Task

11:00–12:00 *Krey-All WMT 2025 CreoleMT System Description: Language Agnostic Strategies for Low-Resource Translation*

Ananya Ayasi

11:00–12:00 *EdinHelsOW WMT 2025 CreoleMT System Description: Improving Lusophone Creole Translation through Data Augmentation, Model Merging and LLM Post-editing*

Jacqueline Rowe, Ona de Gibert, Mateusz Klimaszewski, Coleman Haley, Alexandra Birch and Yves Scherrer

11:00–12:00 *KozKreolMRU WMT 2025 CreoleMT System Description: Koz Kreol: Multi-Stage Training for English–Mauritian Creole MT*

Yush Rajcoomar

11:00–12:00 *JHU WMT 2025 CreoleMT System Description: Data for Belizean Kriol and French Guianese Creole MT*

Nathaniel Romney Robinson

11:00–12:00 *WMT 2025 CreoleMT Systems Description : Martinican Creole and French*

Ludovic V. Mompelat

Sunday, November 9, 2025 (continued)

11:00–12:00 Low-Resource Indic Language Translation Shared Task

11:00–12:00 *JU-NLP: Improving Low-Resource Indic Translation System with Efficient LoRA-Based Adaptation*

Priyobroto Acharya, Haranath Mondal, Dipanjan Saha, Dipankar Das and Sivaji Bandyopadhyay

11:00–12:00 *An Attention-Based Neural Translation System for English to Bodo*

Subhash Kumar Wary, Birhang Borgoyary, Akher Uddin Ahmed, Mohanji Prasad Sah and Apurbalal Senapati

11:00–12:00 *Tackling Low-Resource NMT with Instruction-Tuned LLaMA: A Study on Kokborok and Bodo*

Deepak Kumar, Kshetrimayum Boynao Singh and Asif Ekbal

11:00–12:00 *DELAB-IITM WMT25: Enhancing Low-Resource Machine Translation for Manipuri and Assamese*

Dingku Singh Oinam and Navanath Saharia

11:00–12:00 *Transformers: Leveraging OpenNMT and Transfer Learning for Low-Resource Indian Language Translation*

Bhagyashree Wagh, Harish Bapat, Neha Gupta and Saurabh Salunkhe

11:00–12:00 *RBG-AI: Benefits of Multilingual Language Models for Low-Resource Languages*

Barathi Ganesh HB and Michal Ptaszynski

11:00–12:00 *ANVITA : A Multi-pronged Approach for Enhancing Machine Translation of Extremely Low-Resource Indian Languages*

SIVABHAVANI J, Daneshwari Kankanwadi, Abhinav Mishra and Biswajit Paul

11:00–12:00 *DoDS-IITPKD: Submissions to the WMT25 Low-Resource Indic Language Translation Task*

Ontiwell Khongthaw, G.L. John Salvin, Shrikant Tryambak Budde, Abigail Nyasha Chigweddedza, Dhruvadeep Malkar and Swapnil Hingmire

11:00–12:00 *A Preliminary Exploration of Phrase-Based SMT and Multi-BPE Segmentations through Concatenated Tokenised Corpora for Low-Resource Indian Languages*

Saumitra Yadav and Manish Shrivastava

11:00–12:00 *AkibaNLP-TUT: Injecting Language-Specific Word-Level Noise for Low-Resource Language Translation*

Shoki Hamada, Tomoyosi Akiba and Hajime Tsukada

11:00–12:00 *BVSLP: Machine Translation Using Linguistic Embellishments for IndicMT Shared Task 2025*

Nisheeth Joshi, Palak Arora, Anju Krishnia, Riya Lonchenpa and Mhasilenuo Vizo

Sunday, November 9, 2025 (continued)

11:00–12:00 *TranssionMT's Submission to the Indic MT Shared Task in WMT 2025*
zebiao zhou, hui li, xiangxun zhu and kangzhen liu

11:00–12:00 Terminology Translation Shared Task

11:00–12:00 *Lanigo at WMT25 Terminology Translation Task: A Multi-Objective Reranking Strategy for Terminology-Aware Translation via Pareto-Optimal Decoding*
Kamil Guttman, Adrian Charkiewicz, Zofia Rostek, Mikołaj Pokrywka and Artur Nowakowski

11:00–12:00 *Fine-tuning NMT Models and LLMs for Specialised EN-ES Translation Using Aligned Corpora, Glossaries, and Synthetic Data: MULTITAN at WMT25 Terminology Shared Task*
Lichao Zhu, Maria Zimina-Poirot, Stephane Patin and Cristian Valdez

11:00–12:00 *Contextual Selection of Pseudo-terminology Constraints for Terminology-aware Neural Machine Translation in the IT Domain*
Benjamin Pong

11:00–12:00 *IRB-MT at WMT25 Terminology Translation Task: Metric-guided Multi-agent Approach*
Ivan Grubišić and Damir Korencic

11:00–12:00 *Terminology-Constrained Translation from Monolingual Data Using GRPO*
Javier Garcia Gilabert, Carlos Escolano, Xixian Liao and Maite Melero

11:00–12:00 *It Takes Two: A Dual Stage Approach for Terminology-Aware Translation*
Akshat Singh Jaswal

12:30–14:00 Lunch Break

Sunday, November 9, 2025 (continued)

14:00–15:30 Session 7: Research Paper Boaster Session

Specification-Aware Machine Translation and Evaluation for Purpose Alignment

Yoko Kayano and Saku Sugawara

OpenWHO: A Document-Level Parallel Corpus for Health Translation in Low-Resource Languages

Raphael Merx, Hanna Suominen, Trevor Cohn and Ekaterina Vylomova

Factors Affecting Translation Quality in In-context Learning for Multilingual Medical Domain

Jonathan Mutal, Raphael Rubino and Pierrette Bouillon

Character-Aware English-to-Japanese Translation of Fictional Dialogue Using Speaker Embeddings and Back-Translation

Ayuna Nagato and Takuya Matsuzaki

DTW-Align: Bridging the Modality Gap in End-to-End Speech Translation with Dynamic Time Warping Alignment

Abderrahmane Issam, Yusuf Can Semerci, Jan Scholtes and Gerasimos Spanakis

Targeted Source Text Editing for Machine Translation: Exploiting Quality Estimators and Large Language Models

Hyuga Koretaka, Atsushi Fujita and Tomoyuki Kajiwara

Self-Retrieval from Distant Contexts for Document-Level Machine Translation

Ziqian Peng, Rachel Bawden and François Yvon

Using Encipherment to Isolate Conditions for the Successful Fine-tuning of Massively Multilingual Translation Models

Carter Louchheim, Denis Sotnichenko, Yukina Yamaguchi and Mark Hopkins

Translate, Then Detect: Leveraging Machine Translation for Cross-Lingual Toxicity Classification

Samuel J. Bell, Eduardo Sánchez, David Dale, Pontus Stenetorp, Mikel Artetxe and Marta R. Costa-jussà

Feeding Two Birds or Favoring One? Adequacy–Fluency Tradeoffs in Evaluation and Meta-Evaluation of Machine Translation

Behzad Shayegh, Jan-Thorsten Peter, David Vilar, Tobias Domhan, Juraj Juraska, Markus Freitag and Lili Mou

DocHPLT: A Massively Multilingual Document-Level Translation Dataset

Dayyán O’Brien, Bhavitvya Malik, Ona de Gibert, Pinzhen Chen, Barry Haddow and Jörg Tiedemann

Sunday, November 9, 2025 (continued)

SONAR-SLT: Multilingual Sign Language Translation via Language-Agnostic Sentence Embedding Supervision

Yasser HAMIDULLAH, Shakib Yazdani, Cennet Oguz, Josef van Genabith and Cristina España-Bonet

GAMBIT+: A Challenge Set for Evaluating Gender Bias in Machine Translation Quality Estimation Metrics

George Filandrianos, Orfeas Menis Mastromichalakis, Wafaa Mohammed, Giuseppe Attanasio and Chrysoula Zerva

Implementing and Evaluating Multi-source Retrieval-Augmented Translation

Tommi Nieminen, Jörg Tiedemann and Sami Virpioja

A Cross-Lingual Perspective on Neural Machine Translation Difficulty

Esther Ploeger, Johannes Bjerva, Jörg Tiedemann and Robert Oestling

15:30–16:00 *Coffee Break*

16:00–17:00 **Session 8: Research Paper Poster Session**