WMT 2025

Tenth Conference on Machine Translation

Proceedings of the Conference

November 8-9, 2025

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
317 Sidney Baker St. S
Suite 400 - 134
Kerrville, TX 78028
USA
Tel: +1-855-225-1962
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 2028 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 eight 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-Chairs

Organizing Committee

Chairs

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

Program Committee

Chairs

Idris Abdulmumin, University of Pretoria
Laurie Burchell, Common Crawl Foundation
Isaac Caswell, Google Research
Daryna Dementieva, Technical University of Munich
Marion Di Marco, TUM
Lukas Edman, TU Munich
Mark Fishel, University of Tartu
Barry Haddow, University of Edinburgh and Aveni
Tom Kocmi, Cohere
Philipp Koehn, Johns Hopkins University
Jean Maillard, Meta AI
Christof Monz, University of Amsterdam
Shu Okabe, TUM Heilbronn
Lisa Yankovskaya, University of Tartu

Program Committee

Sadaf Abdul Rauf, Fatima Jinnah Women Unversity
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 and Aveni
Colin Cherry, Google
Pranjal Chitale, Indian Institute of Technology, Madras
Katsuki Chousa, NTT
Chenhui Chu, Kyoto University
Josep Crego, CHAPSVISION
David Dale, Meta AI
Brian Davis, Dublin City University
Ona De Gibert, University of Helsinki
Nisansa De Silva, University of Moratuwa
Hiroyuki Deguchi, NTT, Inc.
Steve Deneefe, RWS Language Weaver
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
Sören Dreano, Dublin City University
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, imaxin|software
Mercedes García-Martínez, Pangeanic
Harritxu Gete, Vicomtech
Jesús González-Rubio, WebInterpret
Isao Goto, Ehime University
Thamme Gowda, Microsoft
Varun Gumma, Microsoft
Kamil Guttmann, Laniqo, 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
Tosho 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
Jillaphat 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 and 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
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
Ben Li, Nomura Research Institute, Ltd.
Zhenhao Li, Imperial College London
Baohao Liao, University of Amsterdam
Xixian Liao, Barcelona Supercomputing Center
Lisa Liu, University of California, San Diego
Wuying Liu, Ludong University
Danni Liu, Karlsruhe Institute of Technology

Siyou Liu, University of Macau
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
Bhavitvya 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 and Technology Research Laboratories
Sthembiso 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, Laniqo / 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, Laniqo, 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
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 Steingrimsson, The Arni Magnusson Institute for Icelandic Studies
Steinthór Steingrímsson, The Árni Magnússon Institute for Icelandic Studies
Katsuhito Sudoh, Nara Women's University
Haoxiang Sun, ShanghaiJiaotong University
Eduardo Sánchez, Meta/UCL
George Tambouratzis, ILSP/Athena R.C.
Aleš Tamchyna, Memsource
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
Xiaotian Wang, University of Tsukuba
Wei Wang, Apple AI/ML
Taro Watanabe, Nara Institute of Science and Technology
Rachel Wicks, Johns Hopkins University
Guillaume Wisniewski, Universite Paris Cite and LLF
Derek F. Wong, University of Macau
Yulong Wu, The University of Manchester
Minghao Wu, Monash University
Di Wu, University of Amsterdam
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
Hyeongu Yun, LG AI Research
François Yvon, ISIR CNRS and Sorbonne Université
Armel Randy Zebaze Dongmo, Inria
Hui Zeng, Jiaguyi Language Technology Co., Ltd
Chrysoula Zerva, Instituto de Instituto de Telecomunicações, Instituto Superior Técnico, University of Lisbon
Runzhe Zhan, University of Macau
Xuan Zhang, Johns Hopkins University
Dakun Zhang, CHAPSVISION
Min Zhang, Huawei
Shaolin Zhu, Tianjin university
Lichao Zhu, Paris CitÃ© University
Maria Zimina, Paris Diderot – Sorbonne Paris Cité, CLILLAC-ARP (EA 3967)
Maria Zimina-Poirot, ALTAE, UniversitÃ© Paris CitÃ©
Vilém Zouhar, ETH Zurich, Charles University

Table of Contents

An Empirical Analysis of Machine Translation for Expanding Multilingual Benchmarks
Sara Rajaee, Rochelle Choenni, Ekaterina Shutova and Christof Monz 1

Cross-lingual Human-Preference Alignment for Neural Machine Translation with Direct Quality Optimization
Kaden Uhlig, Joern Wuebker, Raphael Reinauer and John Denero 31

Audio-Based Crowd-Sourced Evaluation of Machine Translation Quality
Sami Haq, Sheila Castilho and Yvette Graham 52

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 64

Context Is Ubiquitous, but Rarely Changes Judgments: Revisiting Document-Level MT Evaluation
Ahrii Kim 81

GIIFT: Graph-guided Inductive Image-free Multimodal Machine Translation
Jiafeng Xiong and Yuting Zhao 98

Specification-Aware Machine Translation and Evaluation for Purpose Alignment
Yoko Kayano and Saku Sugawara 113

OpenWHO: A Document-Level Parallel Corpus for Health Translation in Low-Resource Languages
Raphael Merx, Hanna Suominen, Trevor Cohn and Ekaterina Vylomova 142

Factors Affecting Translation Quality in In-context Learning for Multilingual Medical Domain
Jonathan Mutal, Raphael Rubino and Pierrette Bouillon 161

Character-Aware English-to-Japanese Translation of Fictional Dialogue Using Speaker Embeddings and Back-Translation
Ayuna Nagato and Takuya Matsuzaki 180

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 191

Targeted Source Text Editing for Machine Translation: Exploiting Quality Estimators and Large Language Models
Hyuga Koretaka, Atsushi Fujita and Tomoyuki Kajiwara 200

Self-Retrieval from Distant Contexts for Document-Level Machine Translation
Ziqian Peng, Rachel Bawden and François Yvon 220

Using Encipherment to Isolate Conditions for the Successful Fine-tuning of Massively Multilingual Translation Models
Carter Louchheim, Denis Sotnichenko, Yukina Yamaguchi and Mark Hopkins 241

Translate, Then Detect: Leveraging Machine Translation for Cross-Lingual Toxicity Classification
Samuel Bell, Eduardo Sánchez, David Dale, Pontus Stenetorp, Mikel Artetxe and Marta R. Costa-Jussà 253

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 269

DocHPLT: A Massively Multilingual Document-Level Translation Dataset
Dayyán O'brien, Bhavitvya Malik, Ona De Gibert, Pinzhen Chen, Barry Haddow and Jörg Tiedemann 286

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 301

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 314

Implementing and Evaluating Multi-source Retrieval-Augmented Translation
Tommi Nieminen, Jörg Tiedemann and Sami Virpioja 327

A Cross-Lingual Perspective on Neural Machine Translation Difficulty
Esther Ploeger, Johannes Bjerva, Jörg Tiedemann and Robert Oestling 340

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 Lakougna, Jessica Lundin, Christof Monz, Kenton Murray, Masaaki Nagata, Stefano Perrella, Lorenzo Proietti, Martin Popel, Maja Popović, Parker Riley, Mariya Shmatova, Steinthór Steingrímsson, Lisa Yankovskaya and Vilém Zouhar 355

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 414

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 Deoghare, Archchana Sindhujan, Jiayi Wang, David Ifeoluwa Adelani, Brian Thompson, Tom Kocmi, Markus Freitag and Daniel Deutsch436

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 484

Findings of the WMT 2025 Shared Task of the Open Language Data Initiative
David Dale, Laurie Burchell, Jean Maillard, Idris Abdulmumin, Antonios Anastasopoulos, Isaac Caswell and Philipp Koehn 495

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 503

Findings of the First Shared Task for Creole Language Machine Translation at WMT25
Nathaniel Robinson, Claire Bizon Monroc, Rasul Dent, Stefan Watson, Kenton Murray, Raj Dabre, Andre Coy and Heather Lent 520

Findings of WMT 2025 Shared Task on Low-resource Indic Languages Translation
Partha Pakray, Reddi Krishna, Santanu Pal, Advaitha Vetagiri, Sandeep Dash, Arnab Kumar Maji, Saralin A. Lyngdoh, Lenin Laitonjam, Anupam Jamatia, Koj Sambyo, Ajit Das and Riyanka Manna532

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 554

Midheind at WMT25 General Machine Translation Task
Svanhvít Lilja Ingólfsdóttir, Haukur Jónsson, Kári Steinn Adhalsteinsson, Róbert Fjölnir Birkisson, Sveinbjörn Thórdharson and Thorvaldur Páll Helgason 577

A Preliminary Study of AI Agent Model in Machine Translation
Ahrii Kim 583

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 587

Evaluation of QWEN-3 for English to Ukrainian Translation
Cristian Grozea and Oleg Verbitsky 594

SYSTRAN @ WMT 2025 General Translation Task
Dakun Zhang, Yara Khater, Ramzi Rahli, Anna Rebollo and Josep Crego 599

Shy-hunyuan-MT at WMT25 General Machine Translation Shared Task
Mao Zheng, Zheng Li, Yang Du, Bingxin Qu and Mingyang Song 607

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 614

Instruction-Tuned English to Bhojpuri Neural Machine Translation Using Contrastive Preference Optimization
Kshetrimayum Boynao Singh, Deepak Kumar and Asif Ekbal 638

SH at WMT25 General Machine Translation Task
Hayate Shiroma 644

Simple Test Time Scaling for Machine Translation: Kaze-MT at the WMT25 General Translation Task
Shaomu Tan 651

NTTSU at WMT2025 General Translation Task
Zhang Yin, Hiroyuki Deguchi, Haruto Azami, Guanyu Ouyang, Kosei Buma, Yingyi Fu, Katsuki Chousa and Takehito Utsuro 657

A Decoding for Machine Translation in LLMs - SRPOL Participation in WMT2025*
Adam Dobrowolski, Paweł Przewłocki, Paweł Przybysz, Marcin Szymański and Dawid Siwicki 666

In2X at WMT25 Translation Task
Lei Pang, Hanyi Mao, Quanjia Xiao, Chen Ruihan, Jingjun Zhang, Haixiao Liu and Xiangyi Li 671

CUNI at WMT25 General Translation Task
Josef Jon, Miroslav Hrabal, Martin Popel and Ondřej Bojar 680

UvA-MT's Participation in the WMT25 General Translation Shared Task
Di Wu, Yan Meng, Maya Konstantinovna Nachesa, Seth Aycock and Christof Monz 688

AMI at WMT25 General Translation Task: How Low Can We Go? Finetuning Lightweight Llama Models for Low Resource Machine Translation
Atli Jasonarson and Steinthor Steingrimsson 695

KIKIS at WMT 2025 General Translation Task
Koichi Iwakawa, Keito Kudo, Subaru Kimura, Takumi Ito and Jun Suzuki 705

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 723

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 732

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 740

IRB-MT at WMT25 Translation Task: A Simple Agentic System Using an Off-the-Shelf LLM
Ivan Grubišić and Damir Korencic 753

Improving Low-Resource Japanese Translation with Fine-Tuning and Backtranslation for the WMT 25 General Translation Task
Felipe Fujita and Hideyuki Takada 765

Multi-agentMT: Deploying AI Agent in the WMT25 Shared Task
Ahrii Kim 769

Laniqo at WMT25 General Translation Task: Self-Improved and Retrieval-Augmented Translation
Kamil Guttmann, Zofia Rostek, Adrian Charkiewicz, Antoni Solarski, Mikołaj Pokrywka and Artur Nowakowski 778

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 . . . 789

GENDER1PERSON: Test Suite for Estimating Gender Bias of First-person Singular Forms
Maja Popović and Ekaterina Lapshinova-Koltunski 800

Evaluation of LLM for English to Hindi Legal Domain Machine Translation Systems
Kshetrimayum Boynao Singh, Deepak Kumar and Asif Ekbal 823

RoCS-MT v2 at WMT 2025: Robust Challenge Set for Machine Translation
Rachel Bawden and Benoît Sagot 834

Automated Evaluation for Terminology Translation Related to the EEA Agreement
Selma Dis Hauksdottir and Steinthor Steingrimsson 850

Up to Par? MT Systems Take a Shot at Sports Terminology
Einar Sigurdsson, Magnús Magnússon, Atli Jasonarson and Steinthor Steingrimsson 856

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 866

Tagged Span Annotation for Detecting Translation Errors in Reasoning LLMs
Taemin Yeom, Yonghyun Ryu, Yoonjung Choi and Jinyeong Bak 878

COMET-poly: Machine Translation Metric Grounded in Other Candidates
Maike Züfle, Vilém Zouhar, Tu Anh Dinh, Felipe Maia Polo, Jan Niehues and Mrinmaya Sachan 887

Long-context Reference-based MT Quality Estimation
Sami Haq, Chinonso Osuji, Sheila Castilho, Brian Davis and Thiago Castro Ferreira 905

Evaluating WMT 2025 Metrics Shared Task Submissions on the SSA-MTE African Challenge Set
Senyu Li, Felermino Dario Mario Ali, Jiayi Wang, Rui Sousa-Silva, Henrique Lopes Cardoso, Pontus Stenetorp, Colin Cherry and David Ifeoluwa Adelani 913

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 920

GEMBA V2: Ten Judgments Are Better Than One
Marcin Junczys-Dowmunt 926

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 934

MSLC25: Metric Performance on Low-Quality Machine Translation, Empty Strings, and Language Variants
Rebecca Knowles, Samuel Larkin and Chi-Kiu Lo 945

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 957

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 969

UvA-MT at WMT25 Evaluation Task: LLM Uncertainty as a Proxy for Translation Quality
Di Wu and Christof Monz 974

Submission for WMT25 Task 3
Govardhan Padmanabhan 984

RankedCOMET: Elevating a 2022 Baseline to a Top-5 Finish in the WMT 2025 QE Task
Sujal Maharjan and Astha Shrestha 994

Quality-Informed Segment-Level Error Correction Using Natural Language Explanations from xTower and Large Language Models
Prashant Sharma 999

TASER: Translation Assessment via Systematic Evaluation and Reasoning
Monishwaran Maheswaran, Marco Carini, Christian Federmann and Tony Diaz 1004

Vicomtech@WMT 2025: Evolutionary Model Compression for Machine Translation
David Ponce, Harritxu Gete and Thierry Etchegoyhen 1011

Iterative Layer Pruning for Efficient Translation Inference
Yasmin Moslem, Muhammad Hazim Al Farouq and John Kelleher 1022

Expanding the WMT24++ Benchmark with Rumantsch Grischun, Sursilvan, Sutsilvan, Surmiran, Puter, and Vallader
Jannis Vamvas, Ignacio Pérez Prat, Not 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 1028

A French Version of the OLDI Seed Corpus
Malik Marmonier, Benoît Sagot and Rachel Bawden 1048

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 . . . 1061

Correcting the Tamazight Portions of FLORES+ and OLDI Seed Datasets
Alp Oktem, Mohamed Aymane Farhi, Brahim Essaidi, Naceur Jabouja and Farida Boudichat 1072

Filling the Gap for Uzbek: Creating Translation Resources for Southern Uzbek
Mukhammadsaid Mamasaidov, Azizullah Aral, Abror Shopulatov and Mironshoh Inomjonov 1081

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 Washington 1088

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 Doumbouya, Djibrila Diane, Baba Mamadi Diane, Solo Farabado, Edoardo Ferrante, Alessandro Guasoni, Mamadou Keita, Sudhamoy Debbarma, Ali Kuzhuget, David Anugraha, Muhammad Ravi Shulthan Habibi, Sina Ahmadi, Mingfei Liu and Jonathan Eng 1103

Improved Norwegian Bokmål Translations for FLORES
Petter Mæhlum, Anders Næss Evensen and Yves Scherrer 1124

NRC Systems for the WMT2025-LRSL Shared Task
Samuel Larkin, Chi-Kiu Lo and Rebecca Knowles 1133

TartuNLP at WMT25 LLMs with Limited Resources for Slavic Languages Shared Task
Taido Purason and Mark Fishel 1143

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 1151

Krey-All WMT 2025 CreoleMT System Description: Language Agnostic Strategies for Low-Resource Translation
Ananya Ayasi . . . 1158

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 . . . 1166

KozKreolMRU WMT 2025 CreoleMT System Description: Koz Kreol: Multi-Stage Training for English–Mauritian Creole MT
Yush Rajcoomar . . . 1183

JHU WMT 2025 CreoleMT System Description: Data for Belizean Kriol and French Guianese Creole MT
Nathaniel Robinson . . . 1191

WMT 2025 CreoleMT Systems Description : Martinican Creole and French
Ludovic Mompelat . . . 1198

JU-NLP: Improving Low-Resource Indic Translation System with Efficient LoRA-Based Adaptation
Priyobroto Acharya, Haranath Mondal, Dipanjan Saha, Dipankar Das and Sivaji Bandyopadhyay 1201

An Attention-Based Neural Translation System for English to Bodo
Subhash Wary, Birhang Borgoyary, Akher Ahmed, Mohanji Sah and Apurbalal Senapati . . . 1210

Tackling Low-Resource NMT with Instruction-Tuned LLaMA: A Study on Kokborok and Bodo
Deepak Kumar, Kshetrimayum Boynao Singh and Asif Ekbal . . . 1215

DELAB-IIITM WMT25: Enhancing Low-Resource Machine Translation for Manipuri and Assamese
Dingku Oinam and Navanath Saharia . . . 1222

Transformers: Leveraging OpenNMT and Transfer Learning for Low-Resource Indian Language Translation
Bhagyashree Wagh, Harish Bapat, Neha Gupta and Saurabh Salunkhe . . . 1227

RBG-AI: Benefits of Multilingual Language Models for Low-Resource Languages
Barathi Ganesh Hb and Michal Ptaszynski . . . 1233

ANVITA : A Multi-pronged Approach for Enhancing Machine Translation of Extremely Low-Resource Indian Languages
Sivabhavani J, Daneshwari Kankanwadi, Abhinav Mishra and Biswajit Paul . . . 1240

DoDS-IITPKD:Submissions to the WMT25 Low-Resource Indic Language Translation Task
Ontiwell Khongthaw, G.l. Salvin, Shrikant Budde, Abigairl Chigwededza, Dhruvadeep Malkar and Swapnil Hingmire . . . 1248

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 . . . 1253

AkibaNLP-TUT: Injecting Language-Specific Word-Level Noise for Low-Resource Language Translation
Shoki Hamada, Tomoyosi Akiba and Hajime Tsukada . . . 1259

BVSLP: Machine Translation Using Linguistic Embellishments for IndicMT Shared Task 2025
Nisheeth Joshi, Palak Arora, Anju Krishnia, Riya Lonchenpa and Mhasilenuo Vizo 1265

TranssionMT's Submission to the Indic MT Shared Task in WMT 2025
Zebiao Zhou, Hui Li, Xiangxun Zhu and Kangzhen Liu 1271

Laniqo at WMT25 Terminology Translation Task: A Multi-Objective Reranking Strategy for Terminology-Aware Translation via Pareto-Optimal Decoding
Kamil Guttmann, Adrian Charkiewicz, Zofia Rostek, Mikołaj Pokrywka and Artur Nowakowski 1276

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 1284

Contextual Selection of Pseudo-terminology Constraints for Terminology-aware Neural Machine Translation in the IT Domain
Benjamin Pong 1292

IRB-MT at WMT25 Terminology Translation Task: Metric-guided Multi-agent Approach
Ivan Grubišić and Damir Korencic 1302

Terminology-Constrained Translation from Monolingual Data Using GRPO
Javier Garcia Gilabert, Carlos Escolano, Xixian Liao and Maite Melero 1335

It Takes Two: A Dual Stage Approach for Terminology-Aware Translation
Akshat Jaswal 1344

Program

Saturday, November 8, 2025

09:00 - 09:10 *Opening Remarks - 20 Years of WMT*

09:10 - 10:30 *Session 1 - Shared Task Overview Papers I*

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 Lakougna, Jessica Lundin, Christof Monz, Kenton Murray, Masaaki Nagata, Stefano Perrella, Lorenzo Proietti, Martin Popel, Maja Popović, Parker Riley, Mariya Shmatova, Steinthór Steingrímsson, Lisa Yankovskaya and Vilém Zouhar

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

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 Deoghare, Archchana Sindhujan, Jiayi Wang, David Ifeoluwa Adelani, Brian Thompson, Tom Kocmi, Markus Freitag and Daniel Deutsch

10:30 - 11:00 *Coffee Break*

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

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

Midheind at WMT25 General Machine Translation Task
Svanhvít Lilja Ingólfsdóttir, Haukur Jónsson, Kári Steinn Adhalsteinsson, Róbert Fjölnir Birkisson, Sveinbjörn Thórdharson and Thorvaldur Páll Helgason

A Preliminary Study of AI Agent Model in Machine Translation
Ahrii Kim

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

Evaluation of QWEN-3 for English to Ukrainian Translation
Cristian Grozea and Oleg Verbitsky

SYSTRAN @ WMT 2025 General Translation Task
Dakun Zhang, Yara Khater, Ramzi Rahli, Anna Rebollo and Josep Crego

Shy-hunyuan-MT at WMT25 General Machine Translation Shared Task
Mao Zheng, Zheng Li, Yang Du, Bingxin Qu and Mingyang Song

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

Instruction-Tuned English to Bhojpuri Neural Machine Translation Using Contrastive Preference Optimization
Kshetrimayum Boynao Singh, Deepak Kumar and Asif Ekbal

SH at WMT25 General Machine Translation Task
Hayate Shiroma

Simple Test Time Scaling for Machine Translation: Kaze-MT at the WMT25 General Translation Task
Shaomu Tan

NTTSU at WMT2025 General Translation Task
Zhang Yin, Hiroyuki Deguchi, Haruto Azami, Guanyu Ouyang, Kosei Buma, Yingyi Fu, Katsuki Chousa and Takehito Utsuro

A Decoding for Machine Translation in LLMs - SRPOL Participation in WMT2025*
Adam Dobrowolski, Paweł Przewłocki, Paweł Przybysz, Marcin Szymański and Dawid Siwicki

In2X at WMT25 Translation Task
Lei Pang, Hanyi Mao, Quanjia Xiao, Chen Ruihan, Jingjun Zhang, Haixiao Liu and Xiangyi Li

CUNI at WMT25 General Translation Task
Josef Jon, Miroslav Hrabal, Martin Popel and Ondřej Bojar

UvA-MT's Participation in the WMT25 General Translation Shared Task
Di Wu, Yan Meng, Maya Konstantinovna Nachesa, Seth Aycock and Christof Monz

AMI at WMT25 General Translation Task: How Low Can We Go? Finetuning Lightweight Llama Models for Low Resource Machine Translation
Atli Jasonarson and Steinthor Steingrimsson

KIKIS at WMT 2025 General Translation Task
Koichi Iwakawa, Keito Kudo, Subaru Kimura, Takumi Ito and Jun Suzuki

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

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

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

IRB-MT at WMT25 Translation Task: A Simple Agentic System Using an Off-the-Shelf LLM
Ivan Grubišić and Damir Korencic

Improving Low-Resource Japanese Translation with Fine-Tuning and Backtranslation for the WMT 25 General Translation Task
Felipe Fujita and Hideyuki Takada

Multi-agentMT: Deploying AI Agent in the WMT25 Shared Task
Ahrii Kim

Laniqo at WMT25 General Translation Task: Self-Improved and Retrieval-Augmented Translation
Kamil Guttmann, Zofia Rostek, Adrian Charkiewicz, Antoni Solarski, Mikołaj Pokrywka and Artur Nowakowski

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

11:00 - 12:00 *Test Suites Shared Task*

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

Evaluation of LLM for English to Hindi Legal Domain Machine Translation Systems
Kshetrimayum Boynao Singh, Deepak Kumar and Asif Ekbal

RoCS-MT v2 at WMT 2025: Robust Challenge Set for Machine Translation
Rachel Bawden and Benoît Sagot

Automated Evaluation for Terminology Translation Related to the EEA Agreement
Selma Dis Hauksdottir and Steinthor Steingrimsson

Up to Par? MT Systems Take a Shot at Sports Terminology
Einar Sigurdsson, Magnús Magnússon, Atli Jasonarson and Steinthor Steingrimsson

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*

Tagged Span Annotation for Detecting Translation Errors in Reasoning LLMs
Taemin Yeom, Yonghyun Ryu, Yoonjung Choi and Jinyeong Bak

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

Long-context Reference-based MT Quality Estimation
Sami Haq, Chinonso Osuji, Sheila Castilho, Brian Davis and Thiago Castro Ferreira

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

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

GEMBA V2: Ten Judgments Are Better Than One
Marcin Junczys-Dowmunt

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

MSLC25: Metric Performance on Low-Quality Machine Translation, Empty Strings, and Language Variants
Rebecca Knowles, Samuel Larkin and Chi-Kiu Lo

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

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

UvA-MT at WMT25 Evaluation Task: LLM Uncertainty as a Proxy for Translation Quality
Di Wu and Christof Monz

Submission for WMT25 Task 3
Govardhan Padmanabhan

RankedCOMET: Elevating a 2022 Baseline to a Top-5 Finish in the WMT 2025 QE Task
Sujal Maharjan and Astha Shrestha

Quality-Informed Segment-Level Error Correction Using Natural Language Explanations from xTower and Large Language Models
Prashant Sharma

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 (Alibaba)*

15:00 - 16:00 *Coffee Break*

16:00 - 17:30 *Session 4 - Research Papers Talk Session*

An Empirical Analysis of Machine Translation for Expanding Multilingual Benchmarks
Sara Rajaee, Rochelle Choenni, Ekaterina Shutova and Christof Monz

Cross-lingual Human-Preference Alignment for Neural Machine Translation with Direct Quality Optimization
Kaden Uhlig, Joern Wuebker, Raphael Reinauer and John Denero

Audio-Based Crowd-Sourced Evaluation of Machine Translation Quality
Sami Haq, Sheila Castilho and Yvette Graham

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

Context Is Ubiquitous, but Rarely Changes Judgments: Revisiting Document-Level MT Evaluation
Ahrii Kim

GIIFT: Graph-guided Inductive Image-free Multimodal Machine Translation
Jiafeng Xiong and Yuting Zhao

Sunday, November 9, 2025

09:10 - 10:30 *Session 5 - Shared Task Overview Papers II*

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

Findings of the WMT 2025 Shared Task of the Open Language Data Initiative
David Dale, Laurie Burchell, Jean Maillard, Idris Abdulmumin, Antonios Anastasopoulos, Isaac Caswell and Philipp Koehn

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

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

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

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

10:30 - 11:00 *Coffee Break*

09:10 - 10:30 *Session 5 - Shared Task Posters II*

11:00 - 12:00 *Model Compression Shared Task*

Vicomtech@WMT 2025: Evolutionary Model Compression for Machine Translation
David Ponce, Harritxu Gete and Thierry Etchegoyhen

Iterative Layer Pruning for Efficient Translation Inference
Yasmin Moslem, Muhammad Hazim Al Farouq and John Kelleher

Sunday, November 9, 2025 (continued)

11:00 - 12:00 *Open Language Data Initiative*

Expanding the WMT24++ Benchmark with Rumantsch Grischun, Sursilvan, Sutsilvan, Surmiran, Puter, and Vallader
Jannis Vamvas, Ignacio Pérez Prat, Not 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

A French Version of the OLDI Seed Corpus
Malik Marmonier, Benoît Sagot and Rachel Bawden

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

Correcting the Tamazight Portions of FLORES+ and OLDI Seed Datasets
Alp Oktem, Mohamed Aymane Farhi, Brahim Essaidi, Naceur Jabouja and Farida Boudichat

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

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 Washington

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 Doumbouya, Djibrila Diane, Baba Mamadi Diane, Solo Farabado, Edoardo Ferrante, Alessandro Guasoni, Mamadou Keita, Sudhamoy Debbarma, Ali Kuzhuget, David Anugraha, Muhammad Ravi Shulthan Habibi, Sina Ahmadi, Mingfei Liu and Jonathan Eng

Improved Norwegian Bokmål Translations for FLORES
Petter Mæhlum, Anders Næss Evensen and Yves Scherrer

11:00 - 12:00 *Slavic Languages Shared Task*

NRC Systems for the WMT2025-LRSL Shared Task
Samuel Larkin, Chi-Kiu Lo and Rebecca Knowles

TartuNLP at WMT25 LLMs with Limited Resources for Slavic Languages Shared Task
Taido Purason and Mark Fishel

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*

Krey-All WMT 2025 CreoleMT System Description: Language Agnostic Strategies for Low-Resource Translation
Ananya Ayasi

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

KozKreolMRU WMT 2025 CreoleMT System Description: Koz Kreol: Multi-Stage Training for English–Mauritian Creole MT
Yush Rajcoomar

JHU WMT 2025 CreoleMT System Description: Data for Belizean Kriol and French Guianese Creole MT
Nathaniel Robinson

WMT 2025 CreoleMT Systems Description : Martinican Creole and French
Ludovic Mompelat

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

JU-NLP: Improving Low-Resource Indic Translation System with Efficient LoRA-Based Adaptation
Priyobroto Acharya, Haranath Mondal, Dipanjan Saha, Dipankar Das and Sivaji Bandyopadhyay

An Attention-Based Neural Translation System for English to Bodo
Subhash Wary, Birhang Borgoyary, Akher Ahmed, Mohanji Sah and Apurbalal Senapati

Tackling Low-Resource NMT with Instruction-Tuned LLaMA: A Study on Kokborok and Bodo
Deepak Kumar, Kshetrimayum Boynao Singh and Asif Ekbal

DELAB-IIITM WMT25: Enhancing Low-Resource Machine Translation for Manipuri and Assamese
Dingku Oinam and Navanath Saharia

Transformers: Leveraging OpenNMT and Transfer Learning for Low-Resource Indian Language Translation
Bhagyashree Wagh, Harish Bapat, Neha Gupta and Saurabh Salunkhe

RBG-AI: Benefits of Multilingual Language Models for Low-Resource Languages
Barathi Ganesh Hb and Michal Ptaszynski

ANVITA : A Multi-pronged Approach for Enhancing Machine Translation of Extremely Low-Resource Indian Languages
Sivabhavani J, Daneshwari Kankanwadi, Abhinav Mishra and Biswajit Paul

DoDS-IITPKD:Submissions to the WMT25 Low-Resource Indic Language Translation Task
Ontiwell Khongthaw, G.l. Salvin, Shrikant Budde, Abigairl Chigwededza, Dhruvadeep Malkar and Swapnil Hingmire

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

AkibaNLP-TUT: Injecting Language-Specific Word-Level Noise for Low-Resource Language Translation
Shoki Hamada, Tomoyosi Akiba and Hajime Tsukada

BVSLP: Machine Translation Using Linguistic Embellishments for IndicMT Shared Task 2025
Nisheeth Joshi, Palak Arora, Anju Krishnia, Riya Lonchenpa and Mhasilenuo Vizo

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*

Laniqo at WMT25 Terminology Translation Task: A Multi-Objective Reranking Strategy for Terminology-Aware Translation via Pareto-Optimal Decoding
Kamil Guttmann, Adrian Charkiewicz, Zofia Rostek, Mikołaj Pokrywka and Artur Nowakowski

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

Contextual Selection of Pseudo-terminology Constraints for Terminology-aware Neural Machine Translation in the IT Domain
Benjamin Pong

IRB-MT at WMT25 Terminology Translation Task: Metric-guided Multi-agent Approach
Ivan Grubišić and Damir Korencic

Terminology-Constrained Translation from Monolingual Data Using GRPO
Javier Garcia Gilabert, Carlos Escolano, Xixian Liao and Maite Melero

It Takes Two: A Dual Stage Approach for Terminology-Aware Translation
Akshat Jaswal

12:00 - 14:00 *Lunch Break*

14:00 - 15:30 *Session 7 - Research Paper Boaster Session*

15:30 - 16:00 *Coffee Break*

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

Context Is Ubiquitous, but Rarely Changes Judgments: Revisiting Document-Level MT Evaluation
Ahrii Kim

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 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

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