

COLING 2025

**The 31st International Conference on
Computational Linguistics**

Proceedings of the System Demonstrations

January 19 - 24, 2025

The COLING 2025 organizers gratefully acknowledge the support from the following sponsors.



Host Institution



Diamond



Winter School Sponsor



Platinum



Gold



Silver



Supporter



©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 979-8-89176-198-8

Message from the Program Chairs

This volume contains papers from the system demonstration session of the 31st International Conference on Computational Linguistics (COLING 2025) held in Abu Dhabi, UAE.

The demonstration session complements the conference's presentation and poster sessions and is focused on working software systems that are the tangible outcomes of research on computational linguistics. While the two previous COLINGs did not have a separate demo track, COLING 2025 has returned to having one, as was the case with COLING 2020 and earlier, and several of the other main NLP conferences.

There were 69 submissions, with 62 valid submissions reviewed. The program committee consisted of 149 members, with two chairs overseeing the reviewing. The acceptance criteria we followed during the selection process included the quality of work as well as the utility and demonstrability potential of the presented systems. We accepted 22 papers, giving an acceptance rate of 35

We thank the reviewers, the organizers of COLING, and everyone who submitted papers.

Mark Dras and Brodie Mather

COLING2025 Demonstration Program Co-Chairs

9 January 2025

Program Committee

Program Chairs

Brodie Mather, The Institute of Human & Machine Cognition, USA

Mark Dras, Macquarie University, Australia

Reviewers

Abdulaziz Alhamadani, Alessandra Zarcone, Alessandro Bondielli, Alex Fabbri, Alexander O'Connor, Alexandre Rademaker, Anais Ollagnier, Animesh Prasad, Anna Corazza, Antonio Moreno-Sandoval, Bang Liu, Bharat Ram Ambati, Boliang Zhang, Bonnie Webber, Brian Davis, Chao Shang, Chen Jia, Chen Wu, Chenchen Ding, Chenwei Yan, Chenyang Lyu, Chris Biemann, Christa Womser-Hacker, Dang Tuan Nguyen, Daniel Baumartz, Daniel Campos, Daryna Dementieva, Di Wang, Dimitar Dimitrov, Dimitar Shterionov, Dingzirui Wang, Eleri Aedmaa, Emily Allaway, Erica Kido Shimomoto, Evelin Amorim, Fangzhao Wu, Federico Ruggeri, Fiona Anting Tan, Flammie Pirinen, Francesca Frontini, Gael de Chalendar, Genta Winata, Georg Rehm, German Rigau, Giuseppe Abrami, Gosuddin Kamaruddin Siddiqi, Greta Damo, Guoxiu He, Hans Ole Hatzel, Hansi Hettiarachchi, Hui Lu, Huiqiang Jiang, Isao Goto, James Thorne, Ji-Ung Lee, Jiale Han, Joan Byamugisha, John Judge, Junhui Li, Juntao Yu, Kaili Huang, Kaja Dobrovoljc, Karina Halevy, Katsuhito Sudoh, Kaushik Roy, Kazunori Komatani, Kiet Nguyen, Koji Inoue, Lauriane Aufrant, Lei Shu, Liling Tan, Longyue Wang, Lucia Pitarch, Luis Morgado da Costa, Marco Turchi, Matthew Shardlow, Mikio Nakano, Mingxu Tao, Miquel Esplà-Gomis, Monjoy Saha, Muhammad Abdo, Muhua Zhu, Murali Raghu Babu Balusu, Muyun Yang, Naoki Otani, Nazia Nafis, Neelamadhav Gantayat, Nishat Raihan, Normaziah Aziz, Oier Lopez de Lacalle, Panupong Pasupat, Parnia Bahar, Paul Buitelaar, Paul Rayson, Peifeng Li, Peilin Zhou, Peng Li, Petra Bago, Philipp Koehn, Piji Li, Qihui Xu, Quzhe Huang, Rachele Sprugnoli, Robert Geislinger, Rodolfo Delmonte, Rodrigo Agerri, Roy Ka-Wei Lee, Rui Wang, Ruifan Li, Rémi Cardon, Sandeep Mathias, Sangha Nam, Seid Muhie Yimam, Seung-Hoon Na, Shiyi Han, Shuqing Bian, Simbiat Ajao, Siyou Liu, Smita Bailur, Sourabh Deoghare, Stefano Faralli, Sudeep Gandhe, Tafseer Ahmed, Tanel Alumäe, Taro Watanabe, Tharindu Ranasinghe, Thien Nguyen, Thomas Brovelli (Meyer), Timo Baumann, Tong Xiao, Torsten Zesch, Vibhu Mittal, Voula Giouli, Wang Xu, Wen Zhang, Wolfgang Menzel, Wuwei Lan, Xianchao Wu, Xin Zhang, Xinyuan Lu, Xixin Wu, Yinghui Li, Yo Ehara, Yu-Jung Heo, Yuanliang Meng, Yukun Ma, Yves Lepage, Zhongqing Wang, Ziwei Xu

Table of Contents

<i>PolyMinder: A Support System for Entity Annotation and Relation Extraction in Polymer Science Documents</i>	
Truong Dinh Do, An Hoang Trieu, Van-Thuy Phi, Minh Le Nguyen and Yuji Matsumoto	1
<i>Streamlining Biomedical Research with Specialized LLMs</i>	
Linqing Chen	9
<i>LENS: Learning Entities from Narratives of Skin Cancer</i>	
Daisy Monika Lal, Paul Rayson, Christopher Peter, Ignatius Ezeani, Mo El-Haj, Yafei Zhu and Yufeng Liu	20
<i>Loki: An Open-Source Tool for Fact Verification</i>	
Haonan Li, Xudong Han, Hao Wang, Yuxia Wang, Minghan Wang, Rui Xing, Yilin Geng, Zenan Zhai, Preslav Nakov and Timothy Baldwin	28
<i>UnifiedGEC: Integrating Grammatical Error Correction Approaches for Multi-languages with a Unified Framework</i>	
Yike Zhao, Xiaoman Wang, Yunshi Lan and Weining Qian	37
<i>Reliable, Reproducible, and Really Fast Leaderboards with Evalica</i>	
Dmitry Ustalov	46
<i>BeefBot: Harnessing Advanced LLM and RAG Techniques for Providing Scientific and Technology Solutions to Beef Producers</i>	
Zhihao Zhang, Carrie-Ann Wilson, Rachel Hay, Yvette Everingham and Usman Naseem	54
<i>AI-Press: A Multi-Agent News Generating and Feedback Simulation System Powered by Large Language Models</i>	
Xiawei Liu, Shiyue Yang, Xinnong Zhang, Haoyu Kuang, Libo Sun, Yihang Yang, Siming Chen, Xuanjing Huang and Zhongyu Wei	63
<i>A Probabilistic Toolkit for Multi-grained Word Segmentation in Chinese</i>	
Xi Ma, Yang Hou, Xuebin Wang and Zhenghua Li	83
<i>EasyJudge: an Easy-to-use Tool for Comprehensive Response Evaluation of LLMs</i>	
Yijie LI and Yuan Sun	91
<i>LUCE: A Dynamic Framework and Interactive Dashboard for Opinionated Text Analysis</i>	
Omnia Zayed, Gaurav Negi, Sampritha Hassan Manjunath, Devishree Pillai and Paul Buitelaar	104
<i>RAGthoven: A Configurable Toolkit for RAG-enabled LLM Experimentation</i>	
Gregor Karetka, Demetris Skottis, Lucia Dutková, Peter Hraška and Marek Suppa	117
<i>MuRAR: A Simple and Effective Multimodal Retrieval and Answer Refinement Framework for Multimodal Question Answering</i>	
Zhengyuan Zhu, Daniel Lee, Hong Zhang, Sai Sree Harsha, Loic Feujio, Akash Maharaj and Yunyao Li	126
<i>Human-Like Embodied AI Interviewer: Employing Android ERICA in Real International Conference</i>	
Zi Haur Pang, Yahui Fu, Divesh Lala, Mikey Elmers, Koji Inoue and Tatsuya Kawahara	136

<i>CASE: Large Scale Topic Exploitation for Decision Support Systems</i> Lorena Calvo Bartolomé, Jerónimo Arenas-García and David Pérez Fernández	151
<i>GECTurk WEB: An Explainable Online Platform for Turkish Grammatical Error Detection and Correction</i> Ali Gebeşçe and Gözde Gül Şahin	163
<i>GR-NLP-TOOLKIT: An Open-Source NLP Toolkit for Modern Greek</i> Lefteris Loukas, Nikolaos Smyrnioudis, Chrysa Dikonomaki, Spiros Barbakos, Anastasios Toumazatos, John Koutsikakis, Manolis Kyriakakis, Mary Georgiou, Stavros Vassos, John Pavlopoulos and Ion Androutsopoulos	174
<i>ViSoLex: An Open-Source Repository for Vietnamese Social Media Lexical Normalization</i> Anh Thi-Hoang Nguyen, Dung Ha Nguyen and Kiet Van Nguyen	183
<i>CompUGE-Bench: Comparative Understanding and Generation Evaluation Benchmark for Comparative Question Answering</i> Ahmad Shallouf, Irina Nikishina and Chris Biemann	189
<i>Autonomous Machine Learning-Based Peer Reviewer Selection System</i> Nurmukhammed Aitymbetov and Dimitrios Zorbas	199
<i>CULTURALLY YOURS: A Reading Assistant for Cross-Cultural Content</i> Saurabh Kumar Pandey, Harshit Budhiraja, Sougata Saha and Monojit Choudhury	208
<i>FEAT-writing: An Interactive Training System for Argumentative Writing</i> Yuning Ding, Franziska Wehrhahn and Andrea Horbach	217

Conference Program

PolyMinder: A Support System for Entity Annotation and Relation Extraction in Polymer Science Documents

Truong Dinh Do, An Hoang Trieu, Van-Thuy Phi, Minh Le Nguyen and Yuji Matsumoto

Streamlining Biomedical Research with Specialized LLMs

Linqing Chen

LENS: Learning Entities from Narratives of Skin Cancer

Daisy Monika Lal, Paul Rayson, Christopher Peter, Ignatius Ezeani, Mo El-Haj, Yafei Zhu and Yufeng Liu

Loki: An Open-Source Tool for Fact Verification

Haonan Li, Xudong Han, Hao Wang, Yuxia Wang, Minghan Wang, Rui Xing, Yilin Geng, Zenan Zhai, Preslav Nakov and Timothy Baldwin

UnifiedGEC: Integrating Grammatical Error Correction Approaches for Multi-languages with a Unified Framework

Yike Zhao, Xiaoman Wang, Yunshi Lan and Weining Qian

Reliable, Reproducible, and Really Fast Leaderboards with Evalica

Dmitry Ustalov

BeefBot: Harnessing Advanced LLM and RAG Techniques for Providing Scientific and Technology Solutions to Beef Producers

Zhihao Zhang, Carrie-Ann Wilson, Rachel Hay, Yvette Everingham and Usman Naseem

AI-Press: A Multi-Agent News Generating and Feedback Simulation System Powered by Large Language Models

Xiawei Liu, Shiyue Yang, Xinnong Zhang, Haoyu Kuang, Libo Sun, Yihang Yang, Siming Chen, Xuanjing Huang and Zhongyu Wei

A Probabilistic Toolkit for Multi-grained Word Segmentation in Chinese

Xi Ma, Yang Hou, Xuebin Wang and Zhenghua Li

EasyJudge: an Easy-to-use Tool for Comprehensive Response Evaluation of LLMs

Yijie LI and Yuan Sun

LUCE: A Dynamic Framework and Interactive Dashboard for Opinionated Text Analysis

Omnia Zayed, Gaurav Negi, Sampritha Hassan Manjunath, Devishree Pillai and Paul Buitelaar

RAGthoven: A Configurable Toolkit for RAG-enabled LLM Experimentation

Gregor Karetka, Demetris Skottis, Lucia Dutková, Peter Hraška and Marek Suppa

MuRAR: A Simple and Effective Multimodal Retrieval and Answer Refinement Framework for Multimodal Question Answering

Zhengyuan Zhu, Daniel Lee, Hong Zhang, Sai Sree Harsha, Loic Feujio, Akash Maharaj and Yunyao Li

Human-Like Embodied AI Interviewer: Employing Android ERICA in Real International Conference

Zi Haur Pang, Yahui Fu, Divesh Lala, Mikey Elmers, Koji Inoue and Tatsuya Kawahara

CASE: Large Scale Topic Exploitation for Decision Support Systems

Lorena Calvo Bartolomé, Jerónimo Arenas-García and David Pérez Fernández

GECTurk WEB: An Explainable Online Platform for Turkish Grammatical Error Detection and Correction

Ali Gebeşçe and Gözde Gül Şahin

GR-NLP-TOOLKIT: An Open-Source NLP Toolkit for Modern Greek

Lefteris Loukas, Nikolaos Smyrnioudis, Chrysa Dikonomaki, Spiros Barbakos, Anastasios Toumazatos, John Koutsikakis, Manolis Kyriakakis, Mary Georgiou, Stavros Vassos, John Pavlopoulos and Ion Androutsopoulos

ViSoLex: An Open-Source Repository for Vietnamese Social Media Lexical Normalization

Anh Thi-Hoang Nguyen, Dung Ha Nguyen and Kiet Van Nguyen

CompUGE-Bench: Comparative Understanding and Generation Evaluation Benchmark for Comparative Question Answering

Ahmad Shallouf, Irina Nikishina and Chris Biemann

Autonomous Machine Learning-Based Peer Reviewer Selection System

Nurmukhammed Aitymbetov and Dimitrios Zorbas

CULTURALLY YOURS: A Reading Assistant for Cross-Cultural Content

Saurabh Kumar Pandey, Harshit Budhiraja, Sougata Saha and Monojit Choudhury

FEAT-writing: An Interactive Training System for Argumentative Writing

Yuning Ding, Franziska Wehrhahn and Andrea Horbach