

MeLLM 2026

**The 1st Workshop on Multilinguality in the Era of Large
Language Models**

Proceedings of the Workshop

July 4, 2026

©2026 Association for Computational Linguistics

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 979-8-89176-430-9

Introduction

The 1st Workshop on Multilinguality in the Era of Large Language Models (MeLLM) will be held on July 4, 2026, in San Diego, California, United States, co-located with the 64th Annual Meeting of the Association for Computational Linguistics (ACL). The workshop aims to provide a dedicated venue for advancing language technologies that serve users across the globe. In the era of large language models (LLMs), we focus on (1) improving representations and model architectures to address persistent challenges in multilinguality, and (2) developing application-level solutions that leverage these models to support real-world use cases.

We received 61 archival submissions. Following a double-blind peer review process, 31 papers were accepted for publication, corresponding to an acceptance rate of 49.1%. In addition, 25 non-archival papers were included in the program. These papers span a broad spectrum of multilingual LLM research, ranging from foundational work on cross-lingual transfer and model merging to applied research in agentic systems, model safety, and domain-specific applications.

The workshop program includes five oral presentations and a poster session. We are also pleased to host keynote talks by two distinguished researchers, Julia Hirschberg from Columbia University and Julia Kreutzer from Cohere Labs. Best paper awards will be announced during the workshop.

Finally, we gratefully acknowledge our authors, reviewers, participants, and the ACL conference organizers for making this workshop possible.

MeLLM 2026 General Chairs

Organizing Committee

General Chair

Kaiyu Huang, Beijing Jiaotong University, China
Fengran Mo, Université de Montréal, Canada
Pinzhen Chen, Queen's University Belfast, UK

Program Chair

Xue Zhang, Beijing Jiaotong University, China
Zheyuan Liu, University of Notre Dame, USA
Zhiqi Huang, Capital One, USA
Koel Dutta Chowdhury, Universität des Saarlandes, Germany
Barry Haddow, Aveni, UK
Meng Jiang, University of Notre Dame, USA
Alexandra Birch, University of Edinburgh, UK
Jian-Yun Nie, Université de Montréal, Canada

Program Committee

Program Committee

Zhan Su, Halmstad University, Sweden
Changhao Guan, Beijing Jiaotong University, China
Junpeng Liu, Hong Kong University of Science and Technology (Guangzhou), China
Chao Huang, Beijing Jiaotong University, China
Qingchen Hu, McGill University, Canada
Zhaoxuan Tan, University of Notre Dame, USA
Jiayang Tu, University of Iowa, USA
You Li, Beijing Jiaotong University, China
Guangyao Dou, Johns Hopkins University, USA
Hongliang Li, Beijing Jiaotong University, China
Xianda Du, University of Waterloo, Canada
Yihan Zhu, University of Notre Dame, USA
Linjian Yang, Clemson University, USA
Hao Yu, Dalian University of Technology, China
Songming Zhang, Beijing Jiaotong University, China
Zhibo Man, Beijing Jiaotong University, China
Yuxin Tian, Université de Montréal, Canada
Jinghan Zhang, Clemson University, USA
Yuchen Hui, Université de Montréal, Canada
Zengkui Sun, Beijing Jiaotong University, China
Xiangchi Yuan, Georgia Institute of Technology, USA
Yongan Yu, McGill University, Canada
Xinyan Wang, Portland State University, USA
Guang Yang, Beijing Jiaotong University, China
Weijiang Li, University of Notre Dame, USA
Chenchen Lin, University of Connecticut, USA
Rui Qi, Beijing Jiaotong University, China
Zihe Liu, Beijing Jiaotong University, China
Dongwhi Kim, Northwestern University, USA
Jiahao Liang, McGill University, Canada

Invited Speakers

Julia Hirschberg, Columbia University, USA
Julia Kreutzer, Cohere, Canada

Table of Contents

<i>Lost in Dialect: The Annotation Gap in Multilingual LLM Safety</i> Wajdi Zaghouani	1
<i>Evidence-Augmented Generation Reasoning for Extremely Low-Resource Language Decipherment</i> Xiaoyu Zhu, Long Yuan, Rui Qi and Jinan Xu	14
<i>Improving Korean-English Cross-Lingual Retrieval: A Data-Centric Study of Language Composition and Model Merging</i> Youngjoon Jang, Junyoung Son, Taemin Lee, Seongtae Hong, Hyeonseok Moon, Seungyoon Lee, Andrew Matteson and Heuseok Lim	30
<i>Query-Synergy: Leveraging High-Resource Languages for Improving Retrieval Performance Across Multiple Languages</i> Seongtae Hong, Jungseob Lee, Hyeonseok Moon, Seungyoon Lee, Youngjoon Jang and Heuseok Lim	44
<i>Kyrgyz Text Normalization: A Comparative Study of Neural and Rule-Based Approaches</i> Zarina Uvalieva, Bektemir Kumarbai Uulu, Adilet Metinov, Tynchtykbek Tashbaltaev and Nurtilek Alibekov	52
<i>MultiHaluDet: Multilingual Hallucination Detection via LLM Hidden State Probing</i> Riasad Alvi, Nurul Labib Sayeedi and Md. Faiyaz Abdullah Sayeedi	63
<i>MFMDQwen: Multilingual Financial Misinformation Detection Based on Large Language Model</i> Zhiwei Liu, Yuyan Wang, Yuechen Jiang, Yupeng Cao, Tianlei Zhu, Xiaorui Guo, Zhiyang Deng, Zhiyuan Yao, Xiao-Yang Liu, Jimin Huang and Sophia Ananiadou	75
<i>Multilingual Chain-of-Thought Compression via Cross-Lingual Distillation</i> Jiarui Wan, Songming Zhang and Yufeng Chen	83
<i>When Retrieval Hurts: Evidence Utilization, Script Fidelity, and Knowledge Conflicts in Multilingual RAG</i> Varalekshmy M Mohan, Swathi Jayakumar, Gadha Saji Menon, Sachin Kurup, Veena G and Vani Kanjirangat	92
<i>DIMAS-OMOP: A Deliberative Intelligence-Based Multi-Agent System for Chinese Medical Text Standardization toward OMOP</i> Hanlin Lv, Xiao Wang, Kesong Wu, Lei Li and Lei Wang	108
<i>Beyond Accuracy: A Structured Error Analysis of Multilingual LLMs on Marathi Script Variation and Syntax</i> Tejas Patil and Barnali Chetia	119
<i>Cross-Lingual Sentiment Misalignment: Auditing Multilingual Language Models for Inversion Risk, Dialectal Representation, and Affective Stability</i> Nusrat Jahan Lia and Shubhashis Roy Dipta	127
<i>GAIA-v2-LILT: Multilingual Adaptation of Agent Benchmark beyond Translation</i> Yunsu Kim, Kaden Uhlig and Joern Wuebker	140
<i>Do Thoughts Depth Affect Multilingual Reasoning?</i> Linjian Yang, Xinyan Wang and Kunpeng Liu	149

<i>On the Limits of Model Merging for Multilinguality in Pre-Training</i> Seth Aycock, Fedor Vitiugin, Aleksandr Umnov, Christof Monz and Khalil Sima'an	159
<i>mmPISA-bench: Do LLMs Reason Equally Well Across 43 Languages?</i> Yerzhan Sapenov and Jaromir Savelka	170
<i>Cross-Lingual Bias in Large Language Models: A Comparative Analysis of English and Swahili</i> Ruolei Zhang, Teddy Njuguna and Yue Feng	181
<i>Polyglot-Lion: Efficient Multilingual ASR for Singapore via Balanced Fine-Tuning of Qwen3-ASR</i> Quy-Anh Dang and Chris Ngo	191
<i>The Multilingual Curse at the Retrieval Layer: Evidence from Amharic</i> Yosef Worku Alemneh, Kidist Amde Mekonnen and Maarten de Rijke	201
<i>ShahiEmotion: A Benchmark Dataset for Punjabi Shahmukhi Emotion Detection</i> Usman Nawaz, Muhammad Junaid Iqbal, Tahir Alyas, Muhammad Asaf, Shumayla Yaqoob, Usman Ahmed Raza, Muhammad Amin Nadim, Aftab Rafique and Faisal Rehman	211
<i>Evaluating Multilingual Tokenization under Worst-N Parity-Aware BPE</i> Vani Kanjirangat, David Kletz, Tanja Samardzic, Ljiljana Dolamic and Fabio Rinaldi	221
<i>MLingualFC: Evaluating Jailbreak Vulnerabilities in Multilingual Vision-Language Models</i> Rishabh Makwana, Mamta Mamta, Deeksha Varshney and Oana Cocarascu	229
<i>P3B3: A Multi-Turn Conversational Benchmark for Measuring European and Brazilian Portuguese Variety Bias in LLMs</i> Rafael Ferreira, Inês Vieira, Inês Calvo, James Furtado, Iago Paulo, Diogo Glória-Silva, Diogo Tavares, David Semedo and Joao Magalhaes	240
<i>SN-WER: Script-Normalized WER for Multi-Script Indic ASR Evaluation</i> Priyaranjan Pattnayak	249
<i>Causal Localization of the English Pivot in LLaVA: Mechanistic VLM Analysis and Training-Free Multilingual Steering</i> Abrar Zahin Raihan and Aurchi Chowdhury	257
<i>Multilingual Disparities in LLM-Based Safety Judgments: Evidence from Brand Safety Applications</i> Songjiang Liu, Riley Grossman, Mike Smith, Cristian Borcea and Yi Chen	266
<i>Benchmarking Byte-Pair Encoding Tokenizers on Different Languages with Bits per Byte</i> Soham Chowdhury and Warren Woolf	275
<i>Where Privacy Risk Lives in English-Source Multilingual RAG: A Stage-Decomposed Audit Across Five Query Languages</i> Yanhang Li, Zhichao Fan and Zexin Zhuang	284
<i>The Broken Telephone Changes Tone: Examining Nuanced Linguistic Cues in LLM Chains-of-Translation</i> Quang Minh Nguyen, Maida Aizaz and Braahmi Padmakumar	294
<i>Group-Merger: A LoRA-based Framework for Multilingual Continual Learning</i> Weijian yi, Hongliang Li and Jinan Xu	308
<i>When English Isn't the Best Teacher: Source Language Effects in Cross-Lingual In-Context Learning</i> Fred Philippy, Siwen Guo, Jacques Klein and Tegawendé F. Bissyandé	317