

C3NLP 2026

**The 4th Workshop on Cross-Cultural Considerations in NLP  
(C3NLP 2026)**

**Proceedings of the Workshop**

July 4, 2026

The C3NLP organizers gratefully acknowledge the support from the following sponsors.

**Gold**



©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](mailto:acl@aclweb.org)

ISBN 979-8-89176-420-0

## Introduction

Natural Language Processing has made impressive advances in recent years, with models increasingly becoming useful technologies across a wide range of applications. However, NLP models are trained on data produced by humans, whose language use is shaped by cultural contexts, values, beliefs, and social practices. As a result, these cultural factors are reflected in both the data used to train NLP systems and the behaviors these systems exhibit. If left unexamined, this can lead to misalignments between the cultural assumptions embedded in NLP technologies and the multicultural contexts in which they are deployed. Such misalignments may create barriers for under-represented cultures, violate cultural norms and values, or contribute to the erasure of cultural knowledge.

This workshop aims to build a long-term research agenda on culture and NLP, focusing on: (1) understanding how global cultures and NLP technologies intersect beyond multilingual and cross-lingual research; (2) detecting, measuring, and mitigating biases and harms in ways that reflect local cultures and values; and (3) developing more cross-culturally competent NLP systems. Addressing these questions requires interdisciplinary collaboration across NLP, linguistics, the social sciences, and cultural anthropology.

The workshop brings together NLP researchers and scholars from related disciplines to ground this emerging area of inquiry in cultural theories and social-scientific perspectives. The program focuses on three main themes: inclusivity and representation of cultures in NLP, cultural harms of NLP technologies, and culturally sensitive perspectives on social biases and harms in NLP. To support a broad conversation across disciplinary norms, we invited both archival and non-archival submissions. Archival papers, including long papers of up to 9 pages and short papers of up to 5 pages, are published in the C3NLP proceedings and presented at the workshop. Non-archival submissions, including extended abstracts and previously reviewed or accepted papers, are not published in the proceedings but are presented during the workshop. All new submissions underwent mutually anonymous peer review by the program committee, unless they had already been reviewed through ACL Rolling Review or accepted at another peer-reviewed venue.

We finally received 34 direct submissions and 5 submissions through ARR, accepting 26 direct submissions (4 short and 22 long, including 15 archival and 11 non-archival) and 3 ARR submissions (3 long, including 1 archival and 2 non-archival). Additionally, we received 5 already accepted submissions.

Welcome to the 4th Workshop on Cross-Cultural Considerations in NLP! We extend our heartfelt gratitude to our program committee for their thorough and insightful reviews, as well as to our authors for contributing outstanding and innovative research to this workshop. We eagerly anticipate a day of engaging discussions and the fostering of future collaborations.

– Vinodkumar Prabhakaran, Sunipa Dev, Luciana Benotti, Daniel Hershcovich, Yong Cao, Li Zhou, Bolei Ma, Ife Adebara

## Organizing Committee

Vinodkumar Prabhakaran, Google Research

Sunipa Dev, Google Research

Luciana Benotti, Universidad Nacional de Córdoba

Daniel Hershcovich, University of Copenhagen

Yong Cao, University of Tübingen

Li Zhou, The Chinese University of Hong Kong, Shenzhen

Bolei Ma, LMU Munich

Ife Adebara, The University of British Columbia

## Program Committee

Jaehyeok Lee, Sung Kyun Kwan University  
Youngsik Yun, Dongguk University  
Faiz Ghifari Haznitrama, Korea Advanced Institute of Science & Technology  
Dan Simonson, BlackBoiler, Inc.  
Cao Yuxuan  
Shaoyang Xu, Singapore University of Technology and Design  
Dan Goldwasser, Purdue University and Purdue University  
Joan Nwatu, University of Michigan - Ann Arbor  
Junho Myung  
Minh Duc Bui, Johannes-Gutenberg Universität Mainz  
Alice Oh, Google and Korea Advanced Institute of Science and Technology  
TianZe Zhang  
Yi Zhou, Cardiff University  
Yuning Yang, FDA  
Afshin Rahimi, Amazon  
Haoran Jin  
Nayeon Lee, NAVER  
Srishti Yadav  
Sina Bagheri Nezhad  
Thierry Poibeau  
Parth Mehta  
Xiaoning Dong, Tsinghua University  
Kaijie Chen  
Geng Zhao  
Hongliang Li

## Table of Contents

<i>Annotation as Cultural Interpretation: Rethinking Data Labeling in NLP</i> Wajdi Zaghouani .....	1
<i>Somatic in the East, Psychological in the West?: A Clinically-Grounded Evaluation of Cross-Cultural Depression Symptoms in LLMs</i> Shintaro Sakai, Jisun An, Migyeong Kang and Haewoon Kwak .....	11
<i>Modeling Cultural and Subcultural Variation in Code-Switched Discourse with Topic Annotation</i> Nemika Tyagi, Nelvin Licon-Guevara and Olga Kellert.....	40
<i>GCCLA: Graph-Conditioned Cross-Lingual Adaptation of Large Language Models Under Extreme Data Scarcity (A Case Study in Tigrigna)</i> Hagos Gebremedhin Gebremeskel, Chong Feng and Asefa Mebrahtu Abera .....	50
<i>LLM-Adapted Colombian Spanish Lexicography: Proficiency Control, Hallucination, and Cultural Distortion</i> Johnatan E. Bonilla .....	67
<i>Soft Prompts for Adapting LLMs to Cultural Commonsense Knowledge</i> Gabrielle Le Bellier, Marine Carpuat, Benoît Sagot and Chloé Clavel.....	76
<i>The Mirage of Diversity: Unmasking the Cultural Vocabulary Ceiling in LLMs</i> Soumedhik Bharati, Subhrajit Mukherjee and Shibam Mandal .....	101
<i>The American Palimpsest: Quantifying South Asian English Dialect Erasure in LLMs</i> Soumedhik Bharati, Shibam Mandal, Swarup Kr Ghosh and Sayani Mondal .....	108
<i>Toward Culturally Grounded Natural Language Processing</i> Sina Bagheri Nezhad .....	119
<i>TabletCraft: Bridging a 4,000-Year Cultural Gap with Bidirectional Akkadian NMT and Cuneiform Rendering</i> Zhaohui Geoffrey Wang .....	132
<i>Lost in Translation? How Language Shapes Responsibility Attribution in Large Language Models</i> Pavithra P M Nair, Gilad Gressel and Krishnashree Achuthan.....	137
<i>Ontology-oriented lexico-semantic modeling and neural classification of Chinese chéngyǔ: A culture-aware NLP approach</i> Lian Chen.....	150
<i>Sorry, Can't Help You": How Large Language Models Judge Failures to Help Across Languages</i> Pavithra P M Nair, Gilad Gressel and Krishnashree Achuthan.....	161
<i>Does Reasoning Kill the Joke? Long-Context Humor Understanding in Hindi</i> Kaveri Anuranjana, Navya Shrivastava, Atharv Johar, Rishabh Sabharwal, Gautam Ranka, Aryan Lunawat, Punit Rathore and Radhika Mamidi .....	177
<i>One Style Fits All? Cultural Values Embedded in Conversational AI via a People-Pleasing Lens</i> Yi-Jun Chen, I-Tsen Hsieh and Li-Wun Chang .....	187
<i>Beyond Monolithic Culture: Evaluating Understandability of Online Text Across Cultural Dimensions</i> Saurabh Kumar Pandey, Harshit Gupta, Sougata Saha and Monojit Choudhury .....	204