

The 31st International Conference on Computational Linguistics
COLING 2025

**The First Workshop on
Multilingual Counterspeech Generation
(MCG)
with Shared Task on Multilingual Counterspeech Generation**

Workshop Proceedings

January 19, 2025

Sponsored by the project TSI100923-2023-1, funded by MTDFP,
Secretary of State of Digitization and Artificial intelligence,
ENIA, and by the European Union-Next Generation EU / PRTR

©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 979-8-89176-207-7

Preface

Welcome to the first workshop and shared task on Multilingual Counterspeech Generation, co-located at COLING 2025. Counterspeech is a promising strategy for fighting online hate, and it consists of refuting it via thoughtful and cogent reasons and fact-bound arguments. Despite the increasing interest from the NLP community in automating counterspeech production, existing studies and resources in this area still mainly focus on English. In this first edition of the workshop, we aim to fill this gap by encouraging researchers to delve into the challenging task of counterspeech generation in a multilingual setting. In fact, hate is an international phenomenon, and requires being addressed in all languages. The workshop also hosted a shared task, where participants were tasked with generating counterspeech in four different languages (English, Italian, Spanish and Basque). They also had the opportunity to employ specific background knowledge that was provided, or any additional external knowledge of their choice, to obtain generations of higher quality. Overall, the workshop received 13 paper submissions, including 10 related to the shared task. In total, 10 papers will be presented. Additionally, the program includes one non-archival paper presentation and two invited talks.

We extend our gratitude to the program committee for their constructive and insightful reviews, as well as to all authors for their submissions and engagement with the workshop.

Workshop organizers.

Website link: <https://sites.google.com/view/multilang-counterspeech-gen/>

Organizing Committee

Helena Bonaldi
María Estrella Vallecillo-Rodríguez
Irene Zubiaga
Arturo Montejo-Ráez
Aitor Soroa
María Teresa Martín-Valdivia
Marco Guerini
Rodrigo Agerri

Table of Contents

<i>PANDA - Paired Anti-hate Narratives Dataset from Asia: Using an LLM-as-a-Judge to Create the First Chinese Counterspeech Dataset</i>	
Michael Bennie, Demi Zhang, Bushi Xiao, Jing Cao, Chryseis Xinyi Liu, Jian Meng and Alayo Tripp	1
<i>RSSN at Multilingual Counterspeech Generation: Leveraging Lightweight Transformers for Efficient and Context-Aware Counter-Narrative Generation</i>	
Ravindran V	13
<i>Northeastern Uni at Multilingual Counterspeech Generation: Enhancing Counter Speech Generation with LLM Alignment through Direct Preference Optimization</i>	
Sahil Wadhwa, Chengtian Xu, Haoming Chen, Aakash Mahalingam, Akankshya Kar and Divya Chaudhary	19
<i>NLP@IIMAS-CLTL at Multilingual Counterspeech Generation: Combating Hate Speech Using Contextualized Knowledge Graph Representations and LLMs</i>	
David Salvador Márquez, Helena Montserrat Gómez Adorno, Iliia Markov and Selene Báez Santamaría	29
<i>CODEOFCONDUCT at Multilingual Counterspeech Generation: A Context-Aware Model for Robust Counterspeech Generation in Low-Resource Languages</i>	
Michael Bennie, Bushi Xiao, Chryseis Xinyi Liu, Demi Zhang and Jian Meng	37
<i>HW-TSC at Multilingual Counterspeech Generation</i>	
Xinglin Lyu, Haolin Wang, Min Zhang and Hao Yang	47
<i>MNLP@Multilingual Counterspeech Generation: Evaluating Translation and Background Knowledge Filtering</i>	
Emanuele Moscato, Arianna Muti and Debora Nozza	56
<i>Hyderabadi Pearls at Multilingual Counterspeech Generation : HALT : Hate Speech Alleviation using Large Language Models and Transformers</i>	
Md Shariq Farhan	65
<i>TrenTeam at Multilingual Counterspeech Generation: Multilingual Passage Re-Ranking Approaches for Knowledge-Driven Counterspeech Generation Against Hate</i>	
Daniel Russo	77
<i>The First Workshop on Multilingual Counterspeech Generation at COLING 2025: Overview of the Shared Task</i>	
Helena Bonaldi, María Estrella Vallecillo-Rodríguez, Irune Zubiaga, Arturo Montejo-Raez, Aitor Soroa, María-Teresa Martín-Valdivia, Marco Guerini and Rodrigo Agerri	92

Conference Program

PANDA - Paired Anti-hate Narratives Dataset from Asia: Using an LLM-as-a-Judge to Create the First Chinese Counterspeech Dataset

Michael Bennie, Demi Zhang, Bushi Xiao, Jing Cao, Chryseis Xinyi Liu, Jian Meng and Alayo Tripp

RSSN at Multilingual Counterspeech Generation: Leveraging Lightweight Transformers for Efficient and Context-Aware Counter-Narrative Generation

Ravindran V

Northeastern Uni at Multilingual Counterspeech Generation: Enhancing Counter Speech Generation with LLM Alignment through Direct Preference Optimization

Sahil Wadhwa, Chengtian Xu, Haoming Chen, Aakash Mahalingam, Akankshya Kar and Divya Chaudhary

NLP@IIMAS-CLTL at Multilingual Counterspeech Generation: Combating Hate Speech Using Contextualized Knowledge Graph Representations and LLMs

David Salvador Márquez, Helena Montserrat Gómez Adorno, Iliia Markov and Selene Báez Santamaría

CODEOFCONDUCT at Multilingual Counterspeech Generation: A Context-Aware Model for Robust Counterspeech Generation in Low-Resource Languages

Michael Bennie, Bushi Xiao, Chryseis Xinyi Liu, Demi Zhang and Jian Meng

HW-TSC at Multilingual Counterspeech Generation

Xinglin Lyu, Haolin Wang, Min Zhang and Hao Yang

MNLP@Multilingual Counterspeech Generation: Evaluating Translation and Background Knowledge Filtering

Emanuele Moscato, Arianna Muti and Debora Nozza

Hyderabadi Pearls at Multilingual Counterspeech Generation : HALT : Hate Speech Alleviation using Large Language Models and Transformers

Md Shariq Farhan

TrenTeam at Multilingual Counterspeech Generation: Multilingual Passage Re-Ranking Approaches for Knowledge-Driven Counterspeech Generation Against Hate

Daniel Russo

The First Workshop on Multilingual Counterspeech Generation at COLING 2025: Overview of the Shared Task

Helena Bonaldi, María Estrella Vallecillo-Rodríguez, Irune Zubiaga, Arturo Montejo-Raez, Aitor Soroa, María-Teresa Martín-Valdivia, Marco Guerini and Rodrigo Agerri

