

Third International Workshop on Multimodal Content Analysis for Social Good

Usman Naseem
Macquarie University
Sydney, New South Wales, Australia
usman.naseem@mq.edu

Roy Ka-Wei Lee
Singapore University of
Technology and Design
Singapore
roy_lee@sutd.edu.sg

Surendrabikram Thapa
Virginia Tech
Blacksburg, Virginia, USA
surendrabikram@vt.edu

Mehwish Nasim
The University of Western Australia
Flinders University
Australia
mehwish.nasim@uwa.edu.au

Abstract

The third edition of the International Workshop on Multimedia Content Analysis for Social Good (MM4SG 2025) was held alongside the prestigious Web Conference 2025. This workshop aimed to tackle the critical challenge of analyzing and moderating multimodal content across digital platforms. In today's era, where diverse forms of multimodal data—including memes, text-embedded images, and fabricated content—can rapidly shape public opinion and influence societal narratives, the demand for sophisticated and ethical content moderation strategies has become increasingly urgent. MM4SG 2025 provided a unique forum for interdisciplinary collaboration, bringing together researchers and practitioners from natural language processing, machine learning, computational social science, and ethics to address these pressing concerns. This paper highlights the key themes, discussions, and contributions of the third edition of the MM4SG workshop, with a particular focus on the intersection of computational linguistics and multimodal content analysis. It also explores future directions for the workshop, including expanding its scope and impact in subsequent editions.

CCS Concepts

• **Information systems** → *Information retrieval*; **World Wide Web**; **Multimedia information systems**; *Information retrieval*; • **Computing methodologies** → **Natural language processing**.

Keywords

Multimodal Content Analysis, Multimodal Foundational Models, Multimodality for Social Good, Multimodal Content in the Web

ACM Reference Format:

Usman Naseem, Surendrabikram Thapa, Roy Ka-Wei Lee, and Mehwish Nasim. 2025. Third International Workshop on Multimodal Content Analysis for Social Good. In *Companion Proceedings of the ACM Web Conference 2025*

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
WWW Companion '25, Sydney, NSW, Australia
© 2025 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-1331-6/2025/04
<https://doi.org/10.1145/3701716.3718378>

(*WWW Companion '25*), April 28-May 2, 2025, Sydney, NSW, Australia. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3701716.3718378>

1 Introduction

In today's rapidly evolving digital landscape, the proliferation of multimodal content—including memes, text-embedded images, and AI-generated visuals—has transformed the way information is disseminated and consumed [1, 2]. While these formats enhance online engagement, they also pose significant challenges related to misinformation, harmful content, and ethical content moderation [3]. Addressing these concerns requires sophisticated analytical frameworks that balance freedom of expression with the need for responsible digital governance [5].

The Third International Workshop on Multimedia Content Analysis for Social Good (MM4SG 2025), held in conjunction with the Web Conference 2025, provided a dedicated platform to advance research in this critical area. Bringing together experts from natural language processing, machine learning, computational social science, and digital ethics, the workshop fostered interdisciplinary collaboration to explore innovative approaches to multimodal content analysis. Participants examined methods for detecting misinformation, improving AI-driven content moderation, and leveraging multimodal data for social impact across domains such as public health, education, and civic engagement.

Building on the success of previous editions [4], MM4SG 2025 also emphasized the role of computational linguistics in understanding the complex interplay between text and visuals in online content. Discussions covered advancements in multimodal sentiment analysis, multilingual content processing, and bias mitigation in AI systems. As multimodal content continues to evolve, the insights generated from MM4SG 2025 will play a vital role in shaping future research directions, including the integration of emerging technologies such as augmented reality and ethical AI-driven moderation strategies.

2 Call for Papers

We invited original contributions on a wide range of topics related to multimodal content analysis for social good with a focus on computational linguistics, including, but not limited to:

- MM4SG: Hate, Troll, Cyberbullying, Scams and Abuse Detection

- MM4SG: Fake News, Misinformation, Rumor and Event Detection
- MM4SG: Multimodal Sentiment Analysis
- MM4SG: Disaster Response and Crisis Management in the Web
- MM4SG: Multimodal Healthcare applications using Web data
- MM4SG: Multimodal content analysis for sustainable development goals (SDGs)
- MM4SG: New Datasets for Multimodal Content Analysis on the internet
- MM4SG: Multimodal content generation and analysis
- MM4SG: Large Language Models for Multimodal Model Content Analysis on the internet
- MM4SG: Foundation Models for Multimodal Content Analysis on the internet
- MM4SG: Socially Responsible Multimodal Content Analysis: Fairness, Bias, Accountability, and Transparency

3 Workshop Overview

The workshop received 22 submissions, of which 11 were accepted, resulting in an acceptance rate of 50%. Each submission underwent a rigorous review process, with at least two reviews provided by members of the program committee. The third edition of MM4SG continued to attract global interest, with submissions from a diverse range of countries, including Australia, India, the USA, Singapore, Italy, Bangladesh, the UK, South Korea, and Norway. The workshop program featured paper presentations alongside keynote speeches.

4 Program Committee

Our workshop was supported by various program committee members across the globe. The program committee was responsible for reviewing the allocated papers. The program committee is listed below (in no specific order):

- Faseela Abdullakutty (University of Qatar)
- Girish A. Koushik (University of Surrey)
- Hariram Veeramani (UCLA, USA)
- Junaid Akram (University of Sydney)
- Junaid Rashid (Sejong University)
- Kairui Jin (University of Sydney)
- Kritesh Rauniar (Delhi Technological University, India)
- Pietro Bernardelle (University of Queensland)
- Raj Sonani (Cornell University, USA)
- Rajeev Kumar (Netflix, USA)
- Ravi Teja Potla (Nvidia)
- Rishik Sood (University of Technology Sydney)
- Rohan Kulkarni (Meta)
- Rui Cao (Singapore Management University, Singapore)
- Shuvam Shiwakoti (Virginia Tech)
- Siddhant Bikram Shah (Northeastern University)
- Stefano Civelli (University of Queensland)
- Surabhi Adhikari (Columbia University/ Google)
- Suryodaya Bikram Shahi (University of Maryland, College Park)
- Tabia Tanzin Prama (Jahangirnagar University)
- Venkatesh Velugubantla (Meridian Cooperative)
- Vijayrajsinh Gohil (New York University)

- Xing Zi (University of Technology Sydney)

5 Future Workshops

Future editions of MM4SG will focus on advancing multimodal content analysis by incorporating video and audio data, addressing the ethical implications of generative AI, and enhancing multilingual and cross-cultural analysis. The workshop will continue fostering collaborations between academia, industry, and policymakers to develop practical solutions for misinformation detection, bias mitigation, and responsible content moderation. Through interdisciplinary research, MM4SG aims to drive innovations that promote safer and more inclusive digital spaces.

References

- [1] Surendrabikram Thapa, Surabhi Adhikari, Imran Razzak, Roy Ka-Wei Lee, and Usman Naseem. 2024. Hate Speech Classification in Text-Embedded Images: Integrating Ontology, Contextual Semantics, and Vision-Language Representations. In *International Conference on Advances in Social Networks Analysis and Mining*. Springer, 331–342.
- [2] Surendrabikram Thapa, Farhan Ahmad Jafri, Ali Hürriyetoğlu, Francielle Vargas, Roy Ka Wei Lee, and Usman Naseem. 2023. Multimodal Hate Speech Event Detection-Shared Task 4. In *CASE 2023-Proceedings of the 6th Workshop on Challenges and Applications of Automated Extraction of Socio-Political Events from Text, associated with 14th International Conference on Recent Advances in Natural Language Processing, RANLP 2023*. Association for Computational Linguistics, 151–159.
- [3] Surendrabikram Thapa, Farhan Ahmad Jafri, Kritesh Rauniar, Mehwish Nasim, and Usman Naseem. 2024. RUHate-MM: Identification of Hate Speech and Targets using Multimodal Data from Russia-Ukraine Crisis. In *Companion Proceedings of the ACM on Web Conference 2024*. 1854–1863.
- [4] Surendrabikram Thapa, Usman Naseem, Roy Ka-Wei Lee, and Mehwish Nasim. 2024. Foreword for the First International Workshop on Multimodal Content Analysis for Social Good (MM4SG 2024). In *Companion Proceedings of the ACM on Web Conference 2024*. 1803–1804.
- [5] Jialin Yuan, Ye Yu, Gaurav Mittal, Matthew Hall, Sandra Sajeev, and Mei Chen. 2024. Rethinking Multimodal Content Moderation from an Asymmetric Angle with Mixed-modality. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 8532–8542.