GeBNLP 2026

The 7th Workshop on Gender Bias in Natural Language Processing

Proceedings of the Workshop

November 9, 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 979-8-89176-447-7

Message from the Organisation Committee

This volume contains the proceedings of the 7th Workshop on Gender Bias in Natural Language Processing, held in conjunction with the 5th Asia-Pacific Chapter of the Association for Computational Linguistics and the 15th International Joint Conference on Natural Language Processing (AACL-IJCNLP 2026). The workshop will take place in a fully virtual format on November 9, 2026. This year marks a structural transition for our team; we extend our deepest gratitude to Marta R. Costa-jussà and Christian Hardmeier, who founded this workshop and have now transitioned to a senior steering committee role. We are also delighted to welcome Giuseppe Attanasio and Vera Neplenbroek to the organizing committee.

This year, we received 23 submissions, accepting 14 archival papers for the proceedings alongside one non-archival presentation. The accepted contributions explore gender bias across a diverse spectrum of languages, NLP tasks, and sociolinguistic contexts. We received many studies on the intersection of linguistic structure and model behavior. Rather than evaluating isolated tasks, authors target studying specific grammatical demands of varied languages.

The research presented in this volume broadens the scope of bias evaluation and mitigation techniques. Most importantly, the community continues to expand beyond purely binary frameworks; numerous studies explicitly model and evaluate the representation of non-binary, LGBTQ+, and intersectional identities, paving the way for more inclusive natural language systems. Moreover, this year's research introduces novel diagnostic datasets, multilingual evaluation methodologies, and advanced mitigation frameworks designed to better detect, explain, and reduce implicit biases and misogyny.

Collectively, these contributions underscore that understanding gender bias in NLP requires looking far beyond familiar stereotypes. By bringing together innovative research spanning diverse languages, technical methods, and social contexts, this workshop provides a vital forum for advancing the fair and inclusive NLP. We look forward to an enriching discussion during our virtual gathering on November 9, 2026.

October 2026
Giuseppe Attanasio, Christine Basta, Marta R. Costa-jussà, Agnieszka Faleńska, Christian Hardmeier, Vera Neplenbroek, Debora Nozza, and Karolina Stańczak (Alphabetically ordered)

Organizing Committee

Organizing Committee

Giuseppe Attanasio, Instituto de Telecomunicações
Christine Basta, FCDS, Alexandria University and HiTZ, the University of the Basque Country
Agnieszka Falenska, University of Stuttgart
Vera Neplenbroek, University of Amsterdam
Debora Nozza, Bocconi University
Karolina Stanczak, ETHZ - ETH Zurich and McGill University, McGill University

Steering Committee

Marta R. Costa-jussà, Meta
Christian Hardmeier, IT University of Copenhagen

Program Committee

Reviewers

Neelima Agarwal, Ashutosh Ahuja, Héctor Allende-Cid, Duygu Altinok, Andrew Aquilina, Noor Mairukh Khan Arnob

Bhiman Kumar Baghel, Tom Södahl Bladsjö, Claudia Borg

Filipa Calado, Ankani Chattoraj, Hongyu Chen, Lina Conti

Joke Daems, Hannah Devinney, Enzo Doyen

Iuliia Eremenko

Dennis Fucci

Eleni Gkovedarou, Lavanya Gupta

Supriya Khadka, Gaurav Kumar

Manuel Lardelli, Arun Balajiee Lekshmi Narayanan, Marlene Lutz

Orfeas Menis Mastromichalakis, Arianna Muti, Ricardo Muñoz Sánchez

Maya K. Nachesa, Naveen Jafer Nizar

Abigail Oppong

Dhruv Patel, Olatz Perez-de-Vinaspre, Andrea Piergentili, Alexandre Puttick

Yangge Qian, Xiaolin Qin

Mohan Raj, Varsha Kuppur Rajendra, Milankumar Rana

Hamidreza Saffari, Beatrice Savoldi, Mohammadamin Shafiei, Christina Skelton

Samia Touileb

Dennis Ulmer

Anirudh Vasudevan, Huy Hien Vu

Ej Zhou, Haotian Zhu

Keynote Talk

Measurement and impact in bias research

Vagrant Gautam
Heidelberg Institute for Theoretical Studies

Abstract: After nearly a decade of bias research, I argue we still do not have consensus on how to measure bias, and how our work impacts NLP systems deployed in the real world. In this talk, I will provide an overview of some of these issues, such as defining gender and bias, and operationalizing them via imperfect proxies. I will then revisit older debates such as intrinsic versus extrinsic bias, as well as newer debates like natural versus synthetic data, studying open or closed models, and measuring the impact of our work. I will close with some of my own suggestions for where bias research in NLP should go next, but hope to encourage discussion afterwards with the community on how we can address these problems going forward.

Bio: Vagrant Gautam is an Independent Postdoctoral Fellow at the Heidelberg Institute for Theoretical Studies, where they work on both social and technical aspects of trustworthy natural language processing. On the technical side, they focus on the interplay of biases and reasoning, proposing methods for and evaluations of these aspects that draw from the social sciences and humanities. On the social side, they investigate gaps and opportunities in how we define and operationalize abstract concepts like 'trustworthiness', 'gender', etc., as well their societal implications. They hold a PhD in computer science from Saarland University, before which they worked on speech recognition, discourse processing, and experimental phonology in academia and industry. In their free time, they think about birds, listen to metal, and experiment with weird new hobbies.

Keynote Talk

Gender, Users, and Inclusive Multilingual Language Technologies

Beatrice Savoldi
Fondazione Bruno Kessler, MT Unit

Abstract: As language technologies increasingly permeate society and mediate how people communicate, questions of gender bias reach a broader and more diverse set of users and scenarios. What shapes the usability of these technologies and their potential for harm in real-world settings? How do these issues emerge differently across languages, cultures, and marginalized communities? And how are they reconfigured in frontier, natively multimodal models?
Machine translation (MT), among the most extensively studied and user-facing tasks in NLP, has long been at the forefront of gender bias research, during a period marked by societal, linguistic, and technological shifts. Using MT and multilinguality as a litmus test for bias research, I will reflect on the evolution and directions of the field, bringing together work that centers users, pairs inclusive design with sociolinguistic innovations, and extends beyond text to speech technologies.

Bio: Beatrice Savoldi is a researcher in the Machine Translation Unit at Fondazione Bruno Kessler (FBK), Italy. Her research sits at the intersection of language, AI, and society, focusing on fairness and safety in multilingual and multimodal tasks, such as machine translation for both text and speech, and drawing on human-centered approaches to study how language technologies affect everyday users. She obtained her joint international Ph.D. from the Universities of Trento and Augsburg, with a thesis on gender bias in automatic translation that received the University of Augsburg Best PhD Thesis Prize. Her work is regularly published in top-tier NLP venues and has received several awards, including an Outstanding Paper Award at COLING 2020 and two Social Impact Awards at EMNLP 2024. Beatrice is also a co-founder and organizer of the international Gender-Inclusive Translation Technologies (GITT) workshop.

Table of Contents

Gender-Competence Associations in Computer Science Discourse: A Longitudinal Analysis
Lauren Chen 1

Valicə: A gender-neutralized dataset of learner Italian
Céline Leuzinger, Caroline Grand-Clement, Ricardo Muñoz Sánchez and Tom Södahl Bladsjö . 7

WinoQueer-NL: Assessing Bias in Dutch Language Models toward LGBTQ+ Identities
Jiska Beuk and Gerasimos Spanakis 21

GRUFF: LLM Pronoun Fidelity, Reasoning, and Biases in German
Fabian Mewes, Anne Lauscher and Vagrant Gautam 35

Beyond Binary Blame: Gender Identity and Relationship-Norm Bias in LLM Fault Attribution
Meredith McFadden 52

Gender Bias Beyond Stereotypes: Towards Linguistically Gender-Fair Conversational LLMs
Enzo Doyen and Amalia Todirascu 70

"Not Polite but Not Rude." LLM Formality Adaptation Across Gendered Linguistic Styles in Japanese
Momo Takamatsu, Filip Miletić and Agnieszka Falenska 92

RAGing Against Misogyny: Improving Implicit Misogyny Explanations
Ivana Crescenzi, Arianna Muti, Maarten Sap, Tommaso Caselli and Debora Nozza 111

BaP-Bench: A Bias-Aware Bangla Paraphrase Benchmark for Measuring Instruction-Guided Gender Debiasing and Evaluating Transformer Models
Most. Sharmin Sultana Samu and Md. Rakibul Islam 126

Love, Anger, Sadness, or Just a Rating? Gender Assignment in Human and LLM Translations of Amazon Reviews
Shushen Manakhimova, Ekaterina Lapshinova-Koltunski, Khushi Pitroda and Maja Popović . 145

Analyzing Occupational Gender Bias across the LLM Training Lifecycle
Manthan Gadhia, Alexander Hoyle and Karolina Stańczak 157

When the Language Must Choose: A Position on Gender Bias Evaluation for Arabic NLP
Wajdi Zaghouani 173

DEval-MT: An Intersectional Gender Bias Evaluation Framework for German Machine Translation
Michelle Kappl, Leon Schreiber and Stefan Hillmann 181

I Am Working So Hard, Can't You See?: Physical vs. Cognitive Labor Gender Bias in LLMs in Persian and English
Elham Vatankhahan Esfahani, Sahar Bahrami Khorshid and Ramin Golshaie 193