

The 5th Workshop on Modeling Socio-Emotional and Cognitive Processes from Multimodal Data in the Wild (MSECP-Wild)

Dudzik, Bernd; Hrkalic, Tiffany Matej; Küster, Dennis; St-Onge, David; Putze, Felix; Devillers, Laurence

DOI

[10.1145/3577190.3616883](https://doi.org/10.1145/3577190.3616883)

Publication date

2023

Document Version

Final published version

Published in

ICMI 2023 - Proceedings of the 25th International Conference on Multimodal Interaction

Citation (APA)

Dudzik, B., Hrkalic, T. M., Küster, D., St-Onge, D., Putze, F., & Devillers, L. (2023). The 5th Workshop on Modeling Socio-Emotional and Cognitive Processes from Multimodal Data in the Wild (MSECP-Wild). In *ICMI 2023 - Proceedings of the 25th International Conference on Multimodal Interaction* (pp. 828-829). (ACM International Conference Proceeding Series). Association for Computing Machinery (ACM). <https://doi.org/10.1145/3577190.3616883>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.



The 5th Workshop on Modeling Socio-Emotional and Cognitive Processes from Multimodal Data in the Wild (MSECP-Wild)

Bernd Dudzik
Delft University of Technology
Delft, The Netherlands
B.J.W.Dudzik@tudelft.nl

Tiffany Matej Hrkalic
Delft University of Technology
Delft, The Netherlands
Vrije Universiteit Amsterdam
Amsterdam, The Netherlands
t.matejhrkalovic@vu.nl

Dennis Küster
University of Bremen
Bremen, Germany
dennis.kuester@uni-bremen.de

David St-Onge
École de technologie supérieure
Montreal, Canada
david.st-onge@etsmtl.ca

Felix Putze
University of Bremen
Bremen, Germany
felix.putze@uni-bremen.de

Laurence Devillers
LISN-CNRS
Orsay, France
Sorbonne University
Paris, France
devil@limsi.fr

ABSTRACT

The ability to automatically infer relevant aspects of human users' thoughts and feelings is crucial for technologies to intelligently adapt their behaviors in complex interactions. Research on multimodal analysis has demonstrated the potential of technology to provide such estimates for a broad range of internal states and processes. However, constructing robust approaches for deployment in real-world applications remains an open problem. The MSECP-Wild workshop series is a multidisciplinary forum to present and discuss research addressing this challenge. Submissions to this 5th iteration span efforts relevant to multimodal data collection, modeling, and applications. In addition, our workshop program builds on discussions emerging in previous iterations, highlighting ethical considerations when building and deploying technology modeling internal states in the wild. For this purpose, we host a range of relevant keynote speakers and interactive activities.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in ubiquitous and mobile computing.**

KEYWORDS

Multimodal Modeling, Affective Computing, Human-centered AI

ACM Reference Format:

Bernd Dudzik, Tiffany Matej Hrkalic, Dennis Küster, David St-Onge, Felix Putze, and Laurence Devillers. 2023. The 5th Workshop on Modeling Socio-Emotional and Cognitive Processes from Multimodal Data in the Wild (MSECP-Wild). In *INTERNATIONAL CONFERENCE ON MULTIMODAL INTERACTION (ICMI '23)*, October 09–13, 2023, Paris, France. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3577190.3616883>



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike International 4.0 License.

ICMI '23, October 09–13, 2023, Paris, France

© 2023 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-0055-2/23/10.

<https://doi.org/10.1145/3577190.3616883>

1 INTRODUCTION

Modern intelligent systems are expected to support human actions and decisions in complex task environments. Importantly, this demand is pressing in many domains where an application's performance relates to meeting human psycho-social needs, e.g. social robots in therapy [4]. To display adaptive behavior within these settings, systems require plausible estimates of how users think and feel. Research on multimodal analysis of behavioral and physiological data has demonstrated the potential to provide estimates of relevant constructs, e.g., persons' impressions of dependence and power in negotiations [6]. Despite substantial technological advances, important challenges for reliable estimates in real world settings remain unresolved (e.g., context-sensitive analysis [5]), hindering progress toward more successful real-world applications. Crucially, these challenges often arise from a complex interaction of both technological and conceptual issues when modeling the targeted phenomena (e.g., handling subjectivity in machine analysis [3]). This inherent diversity underlines the need for interdisciplinary collaboration in addressing them [1].

Over several iterations, the MSECP-Wild workshop series [7–9, 11] has established itself as an interdisciplinary forum to discuss the state-of-the-art in modeling social and cognitive-affective processes from multimodal signals. It facilitates debates over recent contributions, joint research projects, and critical reflection on current and future efforts.

In the following, we provide a brief overview over the contents of this workshop iteration. Concretely, we will outline submitted paper contributions generally relating to the topic of modeling socio-emotional and cognitive processes, as well as program elements for the thematic focus of this iteration: ethical challenges for multimodal analysis in practice.

2 WORKSHOP CONTENT

2.1 Paper submissions

Submissions to this 5th iteration span efforts relevant to multimodal data collection, modeling, and applications:

Guidelines for designing and building an automated multimodal textual annotation system (Joshua Kim and Kalina Yacef). Presents use-cases for building an automated multimodal annotation system, as well as guidelines for multimodal data collection, feature extraction and user information presentation in the context of automated multimodal annotation systems.

GraphITTI: Attributed Graph-based Dominance Ranking in Social Interaction Videos (Garima Sharma, Shreya Ghosh, Abhinav Dhall, Munawar Hayat, Jianfei Cai and Tom Gedeon). Presents a reformulation of detecting the Most Dominant Person (MDP) by utilizing a graph-based approach, which learns generic and robust person ranking on top of context-level features.

SMYLE: A new multimodal resource of talk-in-interaction including neuro-physiological signal (Auriane Boudin, Roxane Bertrand, Stéphane Rauzy, Matthijs Houles, Thierry Legou, Magalie Ochs and Philippe Blache). Presents the collection of a French multimodal corpus including neuro-physiological data from 60 participants engaged in storytelling and free conversation tasks.

Multiscale Contextual Learning for Speech Emotion Recognition in Emergency Call Center Conversations. (Theo Deschamps-Berger, Lori Lamel and Laurence Y. Devillers) Presents a multi-scale conversational context learning approach for speech emotion recognition, evaluated on a corpus composed of French emergency calls.

A multi-tasking multi-modal approach for predicting discrete and continuous emotions (Alex-Răzvan Ispas and Laurence Y. Devillers). Presents a multi-task multi-modal approach that predicts discrete and continuous emotion representations at the same time, highlighting the importance of cross-regularisation.

Multimodal Entrainment in Bio-Responsive Multi-User VR Interactives (Steve DiPaola). Presents an interactive virtual reality (VR) framework which explores musical and visual forms of entrainment using EEG, heart rate and gestures in real-time.

2.2 Program Focus: Ethical Challenges

Building on previous iterations, the program of this fifth installment of the workshop will focus especially on ethical challenges for estimating socio-emotional and cognitive states. The development and use of technology to estimate these states raises important ethical considerations such as privacy, consent, and bias, especially regarding data collection in the wild (e.g., online data collection has raised substantial questions about consent and impact [2]).

As research's technical and theoretical aspects develop, identifying and considering ethical questions will become more important, especially for application in areas focusing on the socio-emotional needs of vulnerable populations. In particular, research will need to identify ethically viable solutions for multimodal data collection, sharing and storage, as well as (semi)-automatic analysis and evaluation. While some of these may take the form of technological approaches, other elements will require the development and adoption of practices by the community (e.g., formats for documentation to improve accountability [10]). Addressing these pragmatic issues will need to be discussed in line within broader societal debates about what should (not) be done by intelligent systems, e.g. regarding the future of work.

To facilitate a discussion of existing challenges for in-the-wild research and applications of multimodal modeling, we have invited keynote speakers to share insights and experiences from their work. Furthermore, the workshop features several brainstorming sessions and a plenary discussion to reflect on existing practices and debate future research opportunities.

3 CONCLUSIONS

This iteration of MSECP-Wild hosts a variety of submissions covering multimodal data collection, modeling, and interactive adaptation across a diverse range of settings and phenomena. Complementing this variety, we have set a thematic focus on ethical challenges for modeling socio-emotional and cognitive processes through a choice of invited speakers and interactive activities. Overall, we are confident that this program offers an effective platform to stimulate interdisciplinary exchanges and future research.

ACKNOWLEDGMENTS

This research was (partially) funded by the Hybrid Intelligence Center, a 10-year program funded by the Dutch Ministry of Education, Culture, and Science through the Netherlands Organisation for Scientific Research, <https://hybrid-intelligence-centre.nl>, grant number 024.004.022.

REFERENCES

- [1] Patricia Alves-Oliveira, Dennis Küster, Arvid Kappas, and Ana Paiva. 2016. Psychological science in hri: Striving for a more integrated field of research. In *2016 AAAI Fall Symposium Series*.
- [2] Abeba Birhane and Vinay Uday Prabhu. 2021. Large image datasets: A pyrrhic win for computer vision? *2021 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 1536–1546.
- [3] Federico Cabitza, Andrea Campagner, and Valerio Basile. 2023. Toward a Perspectivist Turn in Ground Truthing for Predictive Computing. *Proceedings of the AAAI Conference on Artificial Intelligence* 37 (6 2023), 6860–6868. Issue 6.
- [4] Mohamed Chetouani, Sofiane Boucenna, Laurence Chaby, Monique Plaza, David Cohen, L Chaby, V Luherne-du Boullay, M Chetouani, M Plaza, L Templier, et al. 2017. Social Signal Processing and Socially Assistive Robotics in Developmental Disorders.
- [5] Bernd Dudzik, Joost Broekens, Mark Neerincx, and Hayley Hung. 2020. Exploring Personal Memories and Video Content as Context for Facial Behavior in Predictions of Video-Induced Emotions. *Proceedings of the 2020 International Conference on Multimodal Interaction* 10, 153–162. Issue 20. <https://doi.org/10.1145/3382507.3418814>
- [6] Bernd Dudzik, Simon Columbus, Tiffany Matej Hrkalic, Daniel Balliet, and Hayley Hung. 2021. Recognizing Perceived Interdependence in Face-to-Face Negotiations through Multimodal Analysis of Nonverbal Behavior. *Proceedings of the 2021 International Conference on Multimodal Interaction*, 121–130. Issue 1. <https://doi.org/10.1145/3462244.3479935>
- [7] Bernd Dudzik, Dennis Küster, David St-Onge, and Felix Putze. 2022. The 4th Workshop on Modeling Socio-Emotional and Cognitive Processes from Multimodal Data In-the-Wild (MSECP-Wild). *Association for Computing Machinery (ACM)*, 803–804. <https://doi.org/10.1145/3536221.3564029>
- [8] Dennis Küster, Felix Putze, Patricia Alves-Oliveira, Maike Paetzel, and Tanja Schultz. 2020. Modeling Socio-Emotional and Cognitive Processes from Multimodal Data in the Wild. *ICMI 2020 - Proceedings of the 2020 International Conference on Multimodal Interaction* (2020), 883–885.
- [9] Dennis Küster, Felix Putze, David St-Onge, Pascal E. Fortin, Nerea Urrestilla, and Tanja Schultz. 2021. 3rd Workshop on Modeling Socio-Emotional and Cognitive Processes from Multimodal Data in the Wild. *ICMI 2021 - Proceedings of the 2021 International Conference on Multimodal Interaction* (2021), 860–861.
- [10] Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. 2019. Model Cards for Model Reporting. *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 220–229. Issue Figure 2.
- [11] Felix Putze, Enkelejda Kasneci, Jutta Hild, Erin Solovey, Akane Sano, and Tanja Schultz. 2018. Modeling cognitive processes from multimodal signals. *ICMI 2018 - Proceedings of the 2018 International Conference on Multimodal Interaction* (2018), 663.