

## Call for Participation - ACM Multimedia 2022 Grand Challenge:

== MultiMediate: Multi-modal Group Behaviour Analysis for Artificial Mediation ==

<https://multimediate-challenge.org/>

Artificial mediators are a promising approach to support group conversations, but at present, their abilities are limited by insufficient progress in group behaviour sensing and analysis. The MultiMediate challenge is designed to work towards the vision of effective artificial mediators by facilitating and measuring progress on key group behaviour sensing and analysis tasks. This year, the challenge focuses on backchannel detection and agreement estimation from backchannels, but also continues last year's tasks of eye contact detection and next speaker prediction. This year, we take special care to minimize the overhead needed to participate in the challenge.

== Backchannel detection sub-challenge ==

Backchannels serve important meta-conversational purposes like signifying attention or indicating agreement. They can be expressed in a variety of ways - ranging from vocal behaviour ("yes", "ah-ha") to subtle nonverbal cues like head nods or hand movements. The backchannel detection sub-challenge focuses on classifying whether a participant of a group interaction expresses a backchannel at a given point in time. Challenge participants will be required to perform this classification based on a 10-second context window of audiovisual recordings of the whole group.

== Agreement estimation from backchannels sub-challenge ==

A key function of backchannels is the expression of agreement or disagreement towards the current speaker. It is crucial for artificial mediators to have access to this information to understand the group structure and to intervene to avoid potential escalations. In this sub-challenge, participants will address the task of automatically estimating the amount of agreement expressed in a backchannel. In line with the backchannel detection sub-challenge, a 10-second audiovisual context window containing views on all interactants will be provided.

== MultiMediate'21 sub-challenges ==

Last year's eye contact detection and next speaker prediction sub-challenges will be part of MultiMediate'22 again. We define eye contact as a discrete indication of whether a participant is looking at another participant's face, and if so, who this other participant is. Eye contact has to be detected for the last frame of the 10-second context window. In the next speaker prediction sub-challenge, participants need to predict the speaking status of each participant at one second after the end of the context window.

== Dataset & Evaluation Protocol ==

For training and evaluation, MultiMediate makes use of the MPIIGroupInteraction dataset consisting of 22 three- to four-person discussions and of an unpublished test set of six additional discussions. The dataset consists of frame-synchronised video recordings of all participants as well as audio recordings of the interactions. We will provide baseline

implementations along with pre-computed features to minimise the overhead for participants. The test set will be released one week before the challenge deadline. Participants will in turn submit their predictions for evaluation against ground truth on our servers.

== How to Participate ==

Instructions are available at <https://multimediate-challenge.org/>

Paper submission deadline: TBA

== Organisers ==

Philipp Müller (*German Research Center for Artificial Intelligence*)

Dominik Schiller (*Augsburg University*)

Dominike Thomas (*University of Stuttgart*)

Michael Dietz (*Augsburg University*)

Hali Lindsay (*German Research Center for Artificial Intelligence*)

Patrick Gebhard (*German Research Center for Artificial Intelligence*)

Elisabeth André (*Augsburg University*)

Andreas Bulling (*University of Stuttgart*)