

Usage Agreement

for the MediaEval 2015 Research Collections

Please fill out this form and return it following the instructions that are given at the bottom of the last page. On pages 1-2 (this page and the next), mark the box next to the task or tasks that you have registered to participate in. Then fill out page 2 with your team information. Sign on page 5, and then proceed to also sign any task-specific agreements related to the task or tasks for which you have registered. Note that some tasks, i.e., the 2015 Multimodal Person Discovery task and the 2015 Search and Anchoring Task, may require additional paperwork (which will be provided to you directly by your task organizers).

Note: Please return one form per team, unless the team is composed of people from more than one organization. In that case, please return one form per organization.

QUESST: Query by Example Search on Speech Task The task involves searching FOR audio content WITHIN audio content USING an audio content query. This task is particularly interesting for speech researchers in the area of spoken term detection or low-resource/zero-resource speech processing. The primary performance metric will be the normalized cross entropy cost (Cnxe).

Multimodal Person Discovery in Broadcast TV (New in 2015!) Given raw TV broadcasts, each shot must be automatically tagged with the name(s) of people who can be both seen as well as heard in the shot. The list of people is not known a priori and their names must be discovered in an unsupervised way from provided text overlay or speech transcripts. The task will be evaluated on a new French corpus (provided by INA) and the AGORA Catalan corpus, using standard information retrieval metrics based on a posteriori collaborative annotation of the corpus.

C@merata: Querying Musical Scores The input is a natural language phrase referring to a musical feature (e.g., 'consecutive fifths') together with a classical music score, and the required output is a list of passages in the score which contain that feature. Scores are in the MusicXML format, which can capture most aspects of Western music notation. Evaluation is via versions of Precision and Recall relative to a Gold Standard produced by the organisers.

Affective Impact of Movies (including Violent Scenes Detection) In this task participating teams are expected to classify short movie scenes by their affective content according to two use cases: (1) the presence of depicted violence, and (2) their emotional impact (valence, arousal). The training data consists of short Creative Commons-licensed movie scenes (both professional and amateur) together with human annotations of violence and valence-arousal ratings. The results will be evaluated using standard retrieval and classification metrics.

Emotion in Music (An Affect Task) We aim at detecting emotional dynamics of music using its content. Given a set of songs, participants are asked to automatically generate continuous emotional representations in arousal and valence.

Retrieving Diverse Social Images Task This task requires participants to refine a ranked list of Flickr photos with location related information using provided visual, textual and user credibility information. Results are evaluated with respect to their relevance to the query and the diverse representation of it.

Placing: Multimodal Geo-location Prediction The Placing Task requires participants to estimate the locations where multimedia items (photos or videos) were captured solely by inspecting the content and metadata of these items, and optionally exploiting additional

knowledge sources such as gazetteers. Performance is evaluated using the distance to the ground truth coordinates of the multimedia items.

[] **Verifying Multimedia Use (New in 2015!)** For this task, the input is a tweet about an event that has the profile to be of interest in the international news, and the accompanying multimedia item (image or video). Participants must build systems that output a binary decision representing a verification of whether the multimedia item reflects the reality of the event in the way purported by the tweet. The task is evaluated using the F1 score. Participants are also requested to return a short explanation or evidence for the verification decision.

[] **Context of Experience: Identifying Videos Suiting a Watching Situation (New in 2015!)** This task develops multimodal techniques for automatic prediction of multimedia in a particular consumption content. In particular, we focus on the context of predicting movies that are suitable to watch on airplanes. Input to the prediction methods is movie trailers, and metadata from IMDb. Output is evaluated using the Weighted F1 score, with expert labels as ground truth.

[] **Synchronization of multi-user Event Media** This task requires participants to automatically create a chronologically-ordered outline of multiple audio-visual galleries corresponding to the same event, where data collections are synchronized altogether and aligned along parallel lines over the same time axis, or mixed in the correct order.

[] **DroneProtect: Mini-drone Video Privacy Task (New in 2015!)** Recent popularity of mini-drones and their rapidly increasing adoption in various areas, including photography, news reporting, cinema, mail delivery, cartography, agriculture, and military, raises concerns for privacy protection and personal safety. Input to the task is drone video and output is version of the video which protects privacy while retaining key information about the event or situation recorded.

[] **Search and Anchoring in Video Archives** The 2015 Search and Anchoring in Video Archives task consists of two sub-tasks: search for multimedia content and automatic anchor selection. In the “search for multimedia content” sub-task, participants use multimodal textual and visual descriptions of content of interest to retrieve potentially relevant video segments from within a collection. In the “automatic anchor selection” sub-task, participants automatically predict key elements of videos as anchor points for the formation of hyperlinks to relevant content within the collection. The video collection consists of professional broadcasts from BBC. Participant submissions will be assessed using professionally-created anchors, and crowdsourcing-based evaluation.

Please follow these directions to submit this form:

- *Print, sign, and scan the whole form into a single .pdf file*
- *Please remember to sign both page 5 and also the appropriate task-specific sections (following pages).*
- *Please name the file ME2015UA_<teamname>.pdf*
- *Return the form as an attachment to martha.larson+me15agree@gmail.com (do not use this email for any other purpose)*
- *Please give your email the subject line:
<Your organization Name> ME2015UA_<teamname>.pdf*

Team name used in MediaEval 2015 (as specified during registration): _____
Please note that it is important to provide the team name so that we are able to easily identify your team in the registration system. Thank you.

The _____ (the name of your organization, further referred to as "Organization") engages in research and development work in information retrieval, multimedia processing, speech recognition or related areas.

Official mailing address:

Telephone: _____

Fax (optional): _____

Contact person: _____

E-mail: _____

The Organization agrees to use the audio-visual content and associated data including extracted features, automatically generated metadata, manually generated metadata social metadata and speech recognition transcripts (the "Information") under the following understandings, terms and conditions. These understandings, terms and conditions apply equally to all or to part of the Information, including any updates or new versions of the Information supplied under this agreement.

Copyright

1. The following clause applies to tasks that crawl audio-visual content crawled from the Internet that is associated with a Creative Commons license. Every possible measure has been taken to ensure that the association with a Creative Commons license is a valid one. However, the MediaEval 2015 organizers cannot fully guarantee that these collections contain absolutely no audio-visual content without a Creative Commons license. Such content could potentially enter the collection if it was not correctly marked on the site from which it was collected.

2. The MediaEval 2015 organizers declare any social metadata contained in the Information has been at some time made publicly available on the Internet.

3. Owners of copyright for elements contained in the Information may choose to request deletion of these elements from the Information.

4. The limitation on permitted use contained in the following section is intended to reduce the risk of any action being brought by copyright owners, but if this happens the Organization agrees to bear all associated liability.

Permitted Uses

1. The Information may only be used for research and development of multimedia and information retrieval systems.

2. Summaries, analyses and interpretations of the linguistic properties of the Information may be derived and published, provided it is not possible to reconstruct the Information from these summaries.

3. Small excerpts of the Information may be displayed to others or published in a scientific or technical context, solely for the purpose of describing the research and development carried out and related issues. The name of the Information's owner must be clearly identified in writing at the time of disclosure of the Information and/or in publication. In the case of the Creative Commons data, the "licensor"^[2] must be acknowledged.

Own Assessment of Information Use

The Organization must make its own assessment of the suitability of the Information for its research and development purposes under Permitted Uses.

The MediaEval 2015 organizers do not make any warranties or representations, whether expressed or implied or statutory, of any kind with respect to their Information, including without limitation:

1. that the Information is suitable for any particular purpose
2. regarding the results of any use of the whole or part of the Information
3. as to the accuracy, reliability or content of the Information
4. of the absence of any infringement of any proprietary right (including, without limitation, IPRs, trade secret rights and right over confidential information) of third parties by the use of such Information

The Organization shall in any case bear the entire risk of any consequences that may arise from the use to which it, or to which any person that it directly or indirectly permits or allows to use such Information, puts such Information.

The Information's owner shall not have any liability in respect of the infringement of any patent or other right of any third party resulting from any other Organization exercising any of the permitted uses granted under this agreement.

No Information owner makes any representation or warranty, express or implied, other than as expressly stated in this Agreement.

The Organization agrees and acknowledges that the Information's owners shall not be held responsible, alone or jointly and severally, for any loss, damage or injury resulting from the use made by Organization of their respective Information.

Agreement to Delete Data on Request

The Organization undertakes to delete within thirty days of receiving notice all copies of any nominated document that is part of the Information whenever requested to do so by any one of:

1. The MediaEval Organizers
2. the owner of copyright for a particular element

Access to the Information by Individuals:

The Organization:

1. must control access to the Information by individuals and may only grant access to people working under its control, i.e., its own members, consultants to the Organization, or individuals providing service to the Organization
2. remains responsible for any breach of the Individual Application form by individuals under its control

Termination

Either party may terminate the Agreement at any time by notifying the other party in writing. On termination, the Organization must a) destroy all copies of the Information and b) notify the MediaEval 2015 organizers in writing of the action taken.

Applicable Law

This Agreement is governed by the laws of the Netherlands. Signed by the Organization:

Signature: _____

Date: _____

Name (please print): _____

Title: _____

E-mail _____
(if different from contact person above)

MediaEval 2015 Organizers are represented by:

Dr. Martha Larson, Multimedia Information Retrieval Lab, Delft University of Technology
Mekelweg 4, 2628 CD DELFT, Netherlands (m.a.larson@tudelft.nl)

For a complete list of organizers please see the website: <http://www.multimediaeval.org>

QUESST: Query by Example Search on Speech Task

(Data: Recordings of spoken audio; language not pre-specified.)

Speech data: The speech data, the corresponding transcriptions (if available) and the query speech data have been provided by several institutions, who will be announced when the data is released. The data set for this task, containing the audio and the corresponding transcriptions (when available) for these languages is made available for free for research purposes, within or outside of the evaluation. Any other uses of the data (or parts of it) are conditioned to obtaining prior written permission by the institution(s) that are making it available.

Note that you are kindly requested to cite the task overview paper, which will appear in the MediaEval 2015 Working Notes proceedings, in any publications you write concerning this dataset.

Signature _____

(sign here if participating in **QUESST** to indicate you have read and accepted the task specific conditions)

Multimodal Person Discovery in Broadcast TV (New in 2015!)

(Data: Broadcast TV content)

The Multimodal Person Discovery Task data sets also require separate, additional usage agreements, which will be made available by the task organizers.

The data sets for this task are accompanied by source code and automatically extracted features. They must be used in compliance with the usage conditions set out in the main usage agreement (above). They are provided on an as-is basis with no guarantee of being correct.

Signature _____

(sign here if participating in **Multimodal Person Discovery in Broadcast TV** to indicate you have read and accepted the task specific conditions)

C@merata: Querying Musical Scores

(Data: Natural language questions and classical music scores)

There are no additional conditions for the C@merata Task.

Affective Impact of Movies (including Violent Scenes Detection)

(Data: short Creative Commons–licensed movie scenes together with human annotations of violence and valence–arousal rating.)

The data set for this task is accompanied by automatically extracted low–level features. These features must be used in compliance with the usage conditions set out in the main usage agreement (above). Features are provided on an as–is basis with no guaranty of being correct.

Signature_____

(sign here if participating in the **Affective Impact of Movies Task** to indicate you have read and accepted the task specific conditions)

Emotion in Music (An Affect Task)

(Data: Collection of songs from Free Music Archive covering different genres of mainstream western music.)

Features: The data set for this task is accompanied by automatically extracted features. These features must be used in compliance with the usage conditions set out in the main usage agreement (above). Features are provided on an as–is basis with no guarantee of being correct.

Signature_____

(sign here if participating in the **Emotion in Music Task** to indicate you have read and accepted the task specific conditions)

Retrieving Diverse Social Images Task

(Data: Creative Commons Flickr and Wikipedia data: images and metadata in various languages; together with visual and text content features.)

Features: The data set for this task is accompanied by automatically extracted image/text features. These features must be used in compliance with the usage conditions set out in the main usage agreement (above). Features are provided on an as–is basis with no guaranty of being correct.

Signature_____

(sign here if participating in the **Retrieving Diverse Social Images Task** to indicate you have read and accepted the task specific conditions)

Placing: Multimodal Geo-location Prediction

(Data: Flickr images and videos that are tagged with Creative Commons licenses as well as their metadata)

Features: The data set for this task is accompanied by automatically extracted features and screen/video shots. These features must be used in compliance with the usage conditions set out in the main usage agreement (above).

Signature_____

(sign here if participating in the **Placing Task** to indicate you have read and accepted the task specific conditions)

Verifying Multimedia Use (New in 2015!)

(Data: Tweet ids and basic metadata + ground truth annotations for the development set)

The data set for this task is accompanied by utility source code and automatically extracted features. These must be used in compliance with the usage conditions set out in the main usage agreement (above). The source code and features are provided on an as-is basis with no guarantee of being correct.

Note that you are kindly requested to cite the task overview paper, which will appear in the MediaEval 2015 Working Notes proceedings, in any publications you write concerning this dataset.

Signature_____

(sign here if participating in the **Verifying Multimedia Use Task** to indicate you have read and accepted the task specific conditions)

Context of Experience: Identifying Videos Suiting a Watching Situation (New in 2015!)

(Data: URLs of online media files and textual metadata)

There are no additional conditions for the Context of Experience Task.

Synchronization of multi-user Event Media

(Data: Collection of images and videos arranged into multiple user galleries.)

There are no additional conditions for the Synchronization of multi-user Event Media Task.

DroneProtect: Mini-drone Video Privacy Task (New in 2015!)

(Data: About 40 HD video files of an average length 20 sec each with corresponding annotations)

The dataset for this task is available for research and academic publication under the condition that the following scientific paper is cited:

M. Bonetto, P. Korshunov, G. Ramponi and T. Ebrahimi. *Privacy in Mini-drone Based Video Surveillance. Workshop on De-identification for privacy protection in multimedia, Ljubljana, Slovenia, 2015.*

Signature _____

(sign here if participating in the **DroneProtect Task** to indicate you have read and accepted the task specific conditions)

Search and Anchoring in Video Archives

(Data: Television data)

The Search and Hyperlinking Task data set also requires 2 separate, additional usage agreements (for the BBC data and for the ASR transcripts), which will be made available by the task organizers.

Please follow these directions to submit this form (same directions as above, repeated here, because they are important!)

- *Print, sign, and scan the whole form into a single .pdf file*
- *Please remember to sign both page 5 and also the appropriate task-specific sections (following pages).*
- *Please name the file ME2015UA_<teamname>.pdf*
- *Return the form as an attachment to martha.larson+me15agree@gmail.com (do not use this email for any other purpose)*
- *Please give your email the subject line:
<Your organization Name> ME2015UA_<teamname>.pdf*

Thanks for following the directions exactly, it helps us to cut the administrative overhead, and get the data released to you more quickly. Best of success at MediaEval 2015.

^[1] <http://creativecommons.org/>

^[2] <http://creativecommons.org/licenses/>