

Usage Agreement

for the MediaEval 2016 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 2-3 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 for three tasks (Multimodal Person Discovery, Emotional Impact of Movies, and Placing) additional steps/forms are required, as described below.

Note: Please return one form per team, unless the team is composed of people from more than one organization. In that case, each organization (i.e., university or company) in the team should sign a separate form. We request that these multi-organization teams designate one person to collect and submit all forms from the team in a single email.

Zero Cost Speech Recognition Task (ex QUESST) The goal of this task is to challenge teams to come up and experiment with bootstrapping techniques, which allow to train initial ASR system or speech tokenizers for “free”. By “free” we mean a technique that makes it possible to train a speech recognition system on public resource data without the need of buying (expensive, or ideally any) datasets. This year’s language will be Vietnamese.

Multimodal Person Discovery in Broadcast TV 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 composite multilingual corpus from INA (French), DW (German & English) and UPC (Catalan) using standard information retrieval metrics based on a posteriori collaborative annotation of the corpus by the participants themselves.

C@merata: Querying Musical Scores with English Noun Phrases Task 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.

Emotional Impact of Movies Task For this task, participating teams develop approaches to automatically infer the emotional impact of movies. Specifically, induced valence and arousal scores must be predicted in two scenarios: (1) global prediction for short video excerpts, and (2) continuous prediction for longer videos. The training data consists of Creative Commons-licensed movies (professional and amateur) together with human annotations valence-arousal ratings. The results will be evaluated using standard evaluation metrics.

Predicting Media Interestingness Task (New in 2016!) This task requires participants to automatically select frames or portions of movies which are the most interesting for a common viewer. To solve the task, participants can make use of the provided visual, audio and text content. System performance is to be evaluated using standard Mean Average Precision.

Retrieving Diverse Social Images Task This task requires participants to refine a ranked list of Flickr photos retrieved with general purpose multi-topic queries 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 Task** 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.

[] **Verifying Multimedia Use** Participants are required to design a system that verifies whether a tweet and its accompanying multimedia content item (image or video) truthfully reflect a real-world event. The system can make use of any available tweet characteristics and/or features derived from the multimedia content. The data used are posts shared during high impact events on social media. The F1 score is the metric that will be used for evaluation.

[] **Context of Multimedia Experience Task** 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.

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 <teamname>_ME2016UA.pdf (add your organization name at the end of the filename for multi-organization teams)*
- ❖ *Return the form as an attachment to martha.larson+me16agree@gmail.com (do not use this email for any other purpose)*
- ❖ *Please give your email the subject line: <teamname> ME2016U*

Team name used in MediaEval 2016 (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 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 2016 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 2016 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 2016 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 2016 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): _____

Position/Organizational Role: _____

E-mail _____

(if different from contact person above)

MediaEval 2016 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>

Zero Cost Speech Recognition Task (ex QUESST)

(Data: Recordings of Vietnamese spoken audio and corresponding transcription)

The data set for this task is composed of audio recording and associated transcripts obtained from different sources. The data sets are obtained, among others, from the following sources:

1. Downloaded from Rhinospike.com and Forvo.com – these data sets belong to a Creative Commons Attribution 3.0 Unported License and as such could be freely used by participants for whatever they would like, during the evaluation and after the evaluation is finished. Please refer to the terms of service in the respective websites for more information.
2. Recorded by ELSA Corp. (www.elsanow.io) – this data set is free to be used for research purposes and after the evaluation, provided that proper attribution is given.

In the case that you provide and share additional data with the task participants, you acknowledge that this data can be freely used for research purposes within or outside of this evaluation. You also acknowledge, that either (a) you have rights for your data or (b) you accompany the data with the original license and reference the source.

Note that you are kindly requested to cite the task overview paper, which will appear in the MediaEval 2016 Working Notes proceedings, in any publications you write concerning this dataset.

Signature _____

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

Multimodal Person Discovery in Broadcast TV

(Data: Broadcast TV content)

Please note: The Multimodal Person Discovery Task data sets also require separate additional usage agreements, which will be made available by the task organizers (check the website).

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 with English Noun Phrases Task

(Data: Natural language questions and classical music scores)

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

Emotional Impact of Movies Task

(Data: short Creative Commons-licensed movie scenes together with human annotations of valence-arousal ratings.)

The data set for this task is the LIRIS-ACCEDE data set (liris-accede.ec-lyon.fr). It consists of Creative Commons-licensed movies (professional and amateur) together with human annotations of valence-arousal ratings. It 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.

The use of this data set also requires an additional usage agreement, available on the website (liris-accede.ec-lyon.fr) or at http://multimediaeval.org/docs/ME2016_EULA.pdf

Signature _____

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

Predicting Media Interestingness Task

(Data: Creative Commons-licensed trailers together with human annotations of interestingness.)

Features: The data set for this task might be 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 **Predicting Media Interestingness Task** to indicate you have read and accepted the task specific conditions)

Retrieving Diverse Social Images Task

(Data: Creative Commons Flickr 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 Task

(Data: A list of Creative Commons licensed Flickr images and videos and auxiliary information.)

The data set for this task is principally formed by a list of identifiers pointing to metadata of Flickr photos and videos that are part of the publicly available YFCC100M collection. Obtaining the YFCC100M is required in order to make sense of the list of identifiers we supply for this task, and it can be requested here: <https://bit.ly/yfcc100md>.

Additionally using the publicly available YLI corpus is optional but recommended as it contains supplemental data (pixels, features, annotations) for the YFCC100M, which can help you better solve the task, and it can be accessed here: <http://mmcommons.org>.

Note that the YFCC100M collection must be used in compliance with the usage conditions set out in its associated usage agreement. The pixel data in the YLI corpus must be used in compliance with the Creative Commons license associated with the corresponding photo or video, while the features and annotations have been placed in the public domain and as such can be freely used. Please cite their respective papers when you use the data offered by the collections. We note that when you use the data from the YFCC100M and the YLI to address the Placing Task as stipulated, you will not violate any of the usage conditions.

Signature _____

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

Verifying Multimedia Use Task

(Data: a. Social media posts' ids, b. Blogs/articles. Basic metadata and 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 2016 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 Multimedia Experience Task

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

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

Please follow these directions to submit this form. These are the same instructions as above, repeated here for convenience (and 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 <teamname>_ME2016UA.pdf (add your organization name at the end of the filename for multi-organization teams)***
- ❖ *Return the form as an attachment to martha.larson+me16agree@gmail.com (do not use this email for any other purpose)*
- ❖ ***Please give your email the subject line: <teamname> ME2016U***

Please return one form per team, unless the team is composed of people from more than one organization. In that case, each organization (i.e., university or company) in the team should sign a separate form. We request that these multi-organization teams designate one person to collect and submit all forms from the team in a single email.

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 2016.

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

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