

Usage Agreement

for the MediaEval 2013 Research Collections

Please fill out this form and return it following the instructions that are given at the bottom of the last page. On the page 1 (this page) 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 4, and then proceed to also sign any task-specific agreements related to the task or tasks for which you have registered. Note that for the MediaEval 2013 Search and Hyperlinking Task an additional form (that will be provided to you directly by your task organizers) will also be requested.

- Social Event Detection for Social Multimedia** This task requires participants to discover social events and organize the related media items in event-specific clusters, within a collection of Web multimedia. *(Data: URLs of Web multimedia together with their metadata.)*
- Search and Hyperlinking of Television Content** This task requires participants to find video segments relevant to an information need and to provide a list of useful hyperlinks for each of these segments. *(Data: Television data.)*
- Placing: Geo-coordinate Prediction for Social Multimedia** This task requires participants to estimate the geographical coordinates (latitude and longitude) of media items (images and videos), as well as predicting how “placeable” a media item actually is. *(Data: Flickr images and videos that are tagged with Creative Commons licenses as well as their metadata.)*
- Violent Scenes Detection in Film (Affect Task)** This task requires participants to deploy multimodal features to automatically detect portions of movies containing violent material. *(Data: A set of 18 Hollywood movies that must be purchased by the participants.)*
- Visual Privacy: Preserving Privacy in Surveillance Videos** For this task, participants will need to propose methods for protection of privacy sensitive elements. *(Data: About 60 high-resolution video files of an average length of 20 seconds each.)*
- Spoken Web Search: Low Resource Spoken Term Detection** The task involves searching FOR audio content WITHIN audio content USING an audio content query. *(Data: Audio from 9 different languages. Overall about 19 hours of audio in 8kHz with about 500 spoken queries for development and 500 for test.)*
- (New!) Question Answering for the Spoken Web** The problem that we wish to explore in this task is how best to build an information retrieval system in which both the queries and the content are spoken. *(Data: Gujarati spoken audio questions and answers.)*
- (New!) Soundtrack Selection for Commercials (MusiClef Task)** Given a TV commercial, participants are required to predict the most suitable music soundtrack from a list of candidate songs. *(Data: Links to video and to music sources as well as to supporting textual material; low-level content features extracted from the relevant video and music sources.)*
- (New!) Similar Segments of Social Speech** This task involves searching in social multimedia, specifically conversations between students in a university department. *(Data: 2-hour collection of dyadic English-language conversations, each 5-10 minutes in length, by members of semi-cohesive social group.)*
- (New!) Retrieving Diverse Social Images** This task requires participants to automatically refine a ranked list of Internet photos using provided visual and textual information. *(Data: Creative Commons Flickr and Wikipedia data: images and metadata, predominantly English language; together with visual and textual content features.)*
- (New!) Emotion in Music (also an Affect Task)** This task is a new task on emotional characterization of music. *(Data: Collection of songs from Free Music Archive covering different genres of mainstream western music.)*
- (New!) Crowdsourcing for Social Multimedia** This task requires participants to create ground truth from raw labels that have been generated by workers on a commercial crowdsourcing platform. *(Data: Collection of Creative Commons Flickr images and raw crowdsourcing annotations.)*

Please remember to sign both page 4 and also the appropriate task-specific sections (refer to the following pages). Print the form, scan it and sent it as an attachment both to Saskia Peters s.peters@tudelft.nl and to Martha Larson m.a.larson@tudelft.nl.

We ask you to please ensure that your scan is one, single file, in .pdf format.

Team name used in MediaEval 2013 (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: _____

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. Much of the Information has been obtained by crawling the Internet. Every possible measure has been taken to ensure that any audio-visual content that is collected from the Internet for distribution by MediaEval 2013 has been released under a Creative Commons license.¹ The MediaEval 2013 organizers cannot fully guarantee that the collection contains 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.

The following clause applies only to Organizations who are participating in the **Visual Privacy Task**, which is indicated by checking the appropriate box above: The data used for the Visual Privacy comes from various datasets, some of which are allowed to be used for research purposes other than MediaEval 2013. The 2013 Visual Privacy Task organizers are to be contacted prior to use or publication of the data to find out the conditions on use and publication outside of the MediaEval 2013 benchmark.

The following clause applies only to Organizations who are participating in the **Spoken Web Search Task**, which is indicated by checking the appropriate box above: The data used for the Spoken Web Search comes from various datasets in 9 different languages recorded by several institutions.

The following clause applies only to Organizations who are participating in the **Affect Task: Violent Scenes Detection**, which is indicated by checking the appropriate box above: The data used for the Affect Task: Violent Scenes Detection is composed of 18 Hollywood movies, which should be purchased by the Organization. This data is therefore under the

¹ <http://creativecommons.org/>

law of copyright corresponding to the country of the Organization. The MediaEval 2013 organizers decline any responsibility of any kind in the use that the Organization may do with this data.

2. The MediaEval 2013 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 2013 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.

² <http://creativecommons.org/licenses/>

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 2013 organizers in writing of the action taken.

The following clause applies only to Organizations who are participating in the **Affect Task: Violent Scenes Detection**, which is indicated by checking the appropriate box above: Violent Scenes Detection: the agreement is effective as of May 15th, 2013 until December 31st, 2013. The obligation of confidentiality shall survive for an additional period of two (2) years after expiration or termination of this agreement.

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 2013 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/>

Social Event Detection for Social Multimedia

(Data: URLs of Web multimedia together with their metadata)

There are no additional conditions for the Social Event Detection Task.

Search and Hyperlinking of Television Content

(Data: Television data)

The Search and Hyperlinking Task data set also requires a separate, additional usage agreement, which will be made available by the task organizers.

Placing: Geo-coordinate Prediction for Social Multimedia

(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)

Violent Scenes Detection in Film (Affect Task)

(Data: A set of 18 Hollywood movies that must be purchased by the participants.)

Metadata contained in the information: Ground truth and concept annotations are provided with the data set. This is also accompanied by automatically generated shot segmentation (shot boundaries and extracted keyframes).

The provisions of the violent scene detection section shall apply only to Information provided by Technicolor R&D France.

All these Metadata have been produced by Technicolor R&D France and shall remain the exclusive property of Technicolor R&D France. No license whatsoever is implied from this agreement except the license for limited use as set forth in the main usage agreement and the following paragraphs.

They must be used in compliance with the usage conditions of the main usage agreement (above). They shall be kept confidential, shall not be disclosed to third parties and shall be made available only to those of Organization's employees who need to know the Information in connection with MediaEval task; and then, only to the extent of such need to know.

They may be used solely for the purpose of the Organization's participation in the 2013 MediaEval task for non-commercial and in-house evaluation. The Organization certifies that these Metadata will not be used for any other purposes than the MediaEval 2013 participation and related publications without explicit written permission from Technicolor.

As an exception to the section "Copyright", for the avoidance of doubt, the corresponding Metadata contained in Information provided by Technicolor R&D France are not released under a Creative Commons License.

Please acknowledge the use of the Metadata by thanking Technicolor (<http://www.technicolor.com/>).

Signature _____

(sign here if participating in the **Affect Task: Violent Scenes Detection** to indicate you have read and accepted the task specific conditions)

Visual Privacy: Preserving Privacy in Surveillance Videos

(Data: About 60 high-resolution video files of an average length of 20 seconds each.)

The dataset for the 2013 Visual Privacy Task is usable for research and publication within the MediaEval 2013 benchmark under the condition that the relevant scientific paper is cited. The reference for the required citation will be made available in the task readme. The reference for the required citation can also be obtained by contacting the Visual Privacy Task organizers directly. For other uses of the data, the Visual Privacy Task organizers must be contacted, as already stated above.

Signature _____

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

Spoken Web Search: Spoken Term Detection for Low Resource Languages

(Data: Audio from 9 different languages. Overall about 19 hours of audio in 8kHz with about 500 spoken queries for development and 500 for test.)

Speech data: The speech data, the corresponding transcriptions (if available) and the query speech data has been provided by the following institutions:

- Romanian and Albanian data: Politehnica University of Bucarest, Romania.
- Czech and non-native English data: BUT Speech@FIT, Czech Republic.
- African Languages data: CSIR Meraka Institute, South Africa.
- Basque data: University of the Basque Country, Spain.

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 2013 Working Notes proceedings, in any publications you write concerning this dataset.

Signature _____

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

(New!) Question Answering for the Spoken Web

(Data: Audio in Gujarati language, consisting of questions and answers. Each of the ca. 4000 data item is an 8 KHz audio file 4–60 secs in length.)

Speech data: The speech data consisting of questions and answers has been provided by IBM Research India. The data set for this task containing the audio must be used in compliance with the usage conditions of the main usage agreement (above).

The speech data and the provided transcriptions may be used solely for the purpose of the organization's participation in the 2013 MediaEval task. Under no condition can these be used to develop or test any system (or component of any system) for building speech recognition or search systems.

The Organization certifies that the speech data, will not be used for any other purposes than the MediaEval 2013 participation and related publications without explicit written permission from IBM India Research Lab. Further, the access to the data will be restricted to researchers who are working on the MediaEval 2013 Question Answering for the Spoken Web task and will not be distributed further wither within or outside the Organization.

Please cite the following paper while using the data:

Jitendra Ajmera, Anupam Joshi, Sougata Mukherjea, Nitendra Rajput, Shrey Sahay, Mayank Shrivastava, Kundan Srivastava, "Two stream indexing for spoken web search," WWW 2011, Hyderabad, India, March 28 – April 1, 2011.

Signature _____

(sign here if participating in the **Question Answering for the Spoken Web Task** to indicate you have read and accepted the task specific conditions)

(New!) Soundtrack Selection for Commercials (MusiClef Task)

(Data: Links to video and to music sources as well as to supporting textual material; low-level content features extracted from the relevant video and music sources.)

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 **MusiClef Task** to indicate you have read and accepted the task specific conditions)

(New!) Similar Segments of Social Speech

(Data: 2-hour collection of dyadic English-language conversations, each 5-10 minutes in length, by members of semi-cohesive social group.)

There are no additional conditions for the Similar Segments of Social Speech Task.

(New!) Retrieving Diverse Social Images *(Data: Creative Commons Flickr and Wikipedia data: images and metadata, predominantly English language; together with visual and textual 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 guarantee 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)

(New!) Emotion in Music (also 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)

(New!) Crowdsourcing for Social Multimedia

(Data: Collection of Creative Commons Flickr images and raw crowdsourcing annotations.)

The data set contains anonymized worker IDs. We agree not to attempt to crack the anonymization or otherwise try to discover the real-world identify of the workers.

Signature _____

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

Please print this form, fill it in and sign it, and then scan it and send it as an e-mail attachment to both Saskia Peters s.peters@tudelft.nl and Martha Larson m.a.larson@tudelft.nl

We ask you to please ensure that your scan is one single file, in .pdf format.