Usage Agreement
for the MediaEval 2012 Research Collections

Please fill out this form and return it following the instructions that are listed at the bottom of the last page. On the first page (this page) mark the box next to the tasks that you are participating in. Sign at the bottom of page 3 and then proceed to sign any task-specific agreements related to the tasks that you are participating in on the remaining pages. Note that for the Tagging Task and the Visual Privacy Task a second form (that will be provided to you directly by your task organizers) will also be requested. Thank you.

☐ **Social Event Detection Task** This task requires participants to discover events and detect media items that are related to either a specific social event or an event-class of interest. By social events we mean that the events are planned by people, attended by people and that the social media are captured by people. (Data: URLs of images and videos available on Flickr and other Internet archives together with metadata).

☐ **Visual Privacy Task** For this task, participants propose methods whereby human faces occurring in digital imagery can be obscured so as to render them unrecognizable. An optimal balance should be struck between obscuring identity and maintaining the quality of the viewing experience from the user perspective. (Data: about 100 high resolution video files of ca 1m30s each and containing one or more persons in an indoor environment.)

☐ **Placing Task** This task involves automatically assigning geo-coordinates to Flickr videos using one or more of: Flickr metadata, visual content, audio content, social information (Data: Creative Commons Flickr data, predominantly English language, extended from the 2011 data set.)

☐ **Spoken Web Task** This task involves searching FOR audio content WITHIN audio content USING an audio content query. It is particularly interesting for speech researchers in the area of spoken term detection. (Data: Audio from four different Indian languages and four South African languages. Each of the ca. 2000 data item is an 8 KHz audio file 4-30 secs in length.)

☐ **Tagging Task** Given a set of tags and a video collection, participants are required to automatically assign the tags to each video based on a combination of modalities, i.e., speech, metadata, audio and visual. (Data: Creative Commons internet video, nearly exclusively English, extended from the 2011 collection.)

☐ **Affect Task: Violent Scenes Detection** This task requires participants to deploy multimodal features to automatically detect portions of movies containing violent material. Any features automatically extracted from the video, including the subtitles, can be used by participants. (Data: A set of 18 Hollywood movies that must be purchased by the participants.)

Please sign both the main form and the appropriate task-specific sections on the final page. Print the form, scan it and sent it as an attachment to Martha Larson m.a.larson@tudelft.nl.

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 the audio-visual content that is collected from the Internet and distributed by MediaEval 2012 has been released under a Creative Commons license.[1] The MediaEval 2012 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 **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 law of copyright corresponding to the country of the Organization. The MediaEval 2012 organizers decline any responsibility of any kind in the use that the Organization may do with this data.

2. The MediaEval 2012 organizers declare that 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 2012 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

---

[1] http://creativecommons.org/

[2] http://creativecommons.org/licenses/

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 2012 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 2nd, 2012 until December 31st, 2012. 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 2012 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 Task

(Data: URLs of images and videos available on Flickr and other internet archives together with metadata)

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

Visual Privacy Task

(Data: about 100 high resolution video files of ca 1m30s each and containing one or more persons in an indoor environment.)

The Visual Privacy Task data set requires a separate usage agreement, which will be made available by your task organizers.

Placing Task

(Data: Creative Commons Flickr data, predominantly English language, extended from the 2011 data set.)

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)

Spoken Web Search Task

(Data: Audio from four different Indian languages and four South African languages. Each of the ca. 2000 data item is an 8 KHz audio file 4–30 secs in length.)

Speech data: The speech data, the corresponding transcriptions and the query speech data has been provided by IBM Research India and CSIR Meraka Institute South Africa. The data set for this task containing the audio and the corresponding transcriptions 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 2012 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 and the corresponding transcriptions, either together or alone, will not be used for any other purposes than the MediaEval 2012 participation and related publications without explicit written permission from IBM India Research Lab and the CSIR Meraka Institute South Africa. Further, the access to the data will be restricted to researchers who are working on the MediaEval2012 Spoken Web Search task and will not be distributed further wither within or outside the organization. Please cite the following paper while using the data:

Arun Kumar, Nitendra Rajput, Dipanjan Chakraborty, Sheetal K. Agarwal, Amit Anil Nanavati, "WWTW: The World Wide Telecom Web," *NSDR 2007 (SIGCOMM workshop), Kyoto, Japan,* 27 August, 2007.

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

Tagging Task

(Creative Commons Blip.tv data, nearly exclusively English language)

Speech transcripts: The data set for this task is accompanied by automatic speech recognition (ASR) transcripts. ***These transcripts must be used in compliance with the usage conditions of the main usage agreement (above) and also in compliance with a separate ASR transcript usage agreement (which will be made available by your task organizers).*** Please note that you will be required to cite the source papers describing the speech recognition technology in any publications you make concerning the data.

Metadata: The data set for this task is accompanied by automatically generated shot segmentation (shot boundaries and extracted keyframes), which must be used in compliance with the usage conditions (above). Please acknowledge the use of the metadata by citing the following paper:

Kelm, P. Schmiedeke, S. and Sikora, T. 2009. Feature-based video key frame extraction for low quality video sequences. *Proceedings of the 10th Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS '09),* pp.25-28, 6-8 May 2009.

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

Affect Task: Violent Scenes Detection

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

Metadata: 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).

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 2012 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 2012 participation and related publications without explicit written permission from Technicolor.

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)

Please print this form, fill it in and sign it and then scan it (into a single file please) and send it as an e-mail attachment to Martha Larson m.a.larson@tudelft.nl