Usage Agreement
for the MediaEval 2021 Research Collections

Please fill out this form and return it following the instructions that are given at the bottom of the last page. On 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 5, and then proceed to also sign any task-specific agreements related to the task or tasks for which you have registered.

If you are participating in one of the following tasks, additional paperwork is required. You will be asked to fill out a task-specific form for these tasks (more information below):

- Driving Road Safety Forward: Video Data Privacy (each team member registers online on Submittable)
- Insight for Wellbeing: Multimodal personal health lifelog data analysis
- Sports Video Annotation: Detection of Strokes in Table Tennis
- Emerging News

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.

[ ] **Driving Road Safety Forward: Video Data Privacy:** Obscure driver identity while preserving human behavioral information. Successful methods balancing driver privacy with fidelity of relevant information have the potential to broaden researcher access to data, and inform the trajectory of transportation safety research, policy, and education initiatives.

[ ] **Emerging News: Detecting emerging stories from social media and news feeds:** Develop a real-time solution that identifies emerging stories.

[ ] **Emotional Mario: A games analytics Challenge**: Use video and sensor readings capture from video game players to create a believable agent

[ ] **Emotions and Themes in Music:** Recognize emotions and themes conveyed in music recordings (large-scale data set).

[ ] **FakeNews: Corona Virus and Conspiracies Multimedia Analysis Task:** Create a machine-learning approach to automatically detect disinformation and its spreaders in social networks.

[ ] **Insight for Wellbeing: Cross-Data Analytics for (transboundary) Haze Prediction:** Improve prediction and/or find insights for wellbeing based on environmental factors, satellite remote sensing, social/news data, etc. The problems this task tries to tackle are "air pollution" and "transboundary haze", which moves across country borders.

[ ] **Medico: Transparency in Medical Image Segmentation:** Develop transparent algorithms for automatic segmentation of images collected from human endoscopies.

[ ] **NewsImages:** Investigate the link between images and text (headlines, content excerpts) in a large collection of online news articles.

[ ] **Predicting Media Memorability:** Given a data set of multimedia content (videos) and associated memorability annotations, automatically train a system to predict memorability.

[ ] **Sports Video: Fine Grained Action Detection and Classification of Table Tennis Strokes from videos:** Automatically classify and detect strokes in videos of table tennis.

[ ] **Visual Sentiment Analysis: A Natural Disaster Use-case:** Develop algorithms to automatically assign multiple sentiments that could be associated with natural disaster images.

[ ] **WaterMM: Water Quality in Social Multimedia:** Automatically classify bilingual social multimedia as relevant or not to water quality, safety and security.

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 4 and also the appropriate task-specific sections (following pages).*
- ❖ **Please name the file <teamname>_ME2021UA.pdf (add your organization name at the end of the filename for multi-organization teams)**
- ❖ *Return the form as an attachment to agree@mediaeven.org (do not use this email for any other purpose)*
- ❖ **Please give your email the subject line: <teamname> ME2021UA**

Team name used in MediaEval 2021 (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, music analysis, speech recognition or related areas.

Official mailing address: _____

_____

Telephone: _____

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, sensor readings, 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. This clause (points 1-4) applies to tasks that crawl audio-visual content from the Internet that is associated with a Creative Commons (cf. http://creativecommons.org) license. Every possible measure has been taken to ensure that the association with a Creative Commons license is a valid one. However, the MediaEval 2021 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 2021 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" (cf. http://creativecommons.org/licenses) 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 2021 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 the 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 named 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 this access restriction 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 2021 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 2021 contact point:

Martha Larson, Institute for Computing and Information Sciences, Radboud University, Toernooiveld 212 6525 EC Nijmegen, Netherlands mlarson@science.ru.nl

For a complete list of organizers please see the task pages on the website https://multimediaeval.github.io/editions/2021/

Driving Road Safety Forward: Video Data Privacy
(Data: high and low resolution video of drivers, designed to emulate different naturalistic driving environments. The data include footage from the same data acquisition system used in the United States Second Strategic Highway Research Program (SHRP2), a program where drivers across the country provided more than 1 million hours of driver video.)

Each participant must register online to access the data. Task participants agree to the following with respect to the data:
1. *Participant will not use or disclose the data for any purpose other than participation in the task described above.*
2. *Participant will not attempt to learn the identity of the individuals in the data.*
3. *Participant will use appropriate safeguards to prevent the use or disclosure of data for any purpose other than as required for participation in the task described above.*
4. *Participant will, at the conclusion of the task, destroy all copies of the data within its possession or in the possession of a party to whom participant has disclosed the data.*
5. *The data used for this task must be used in accordance with the terms of release of the original datasets. The data owner will be a third-party beneficiary to the agreements made hereunder by participant.*

Signature_____

(sign here if participating in the **Driving Road Safety Forward: Video Data Privacy Task** to indicate you have read and accepted the task specific conditions)

Emerging News: Detecting emerging stories from social media and news feed
(Data: Texts collected from social media and news with automated annotations.)

Each participant team must register to access the data. We understand that the text material used for this task must be used in accordance with the terms of release of the original datasets. We declare any social metadata and text contained in the data have been at some time made publicly available on the Internet. Some datasets contain TweetIDs and URLs of online published news. When downloading tweets by the means of the distributed Tweet IDs, users have to be compliant with Twitter's Developer Agreement and Policy (https://developer.twitter.com/en/developer-terms/agreement-and-policy). When downloading the URLs' content, users have to be compliant with the corresponding provider's copyrights and policies.

Signature_____

(sign here if participating in the **Emerging News Task: Detecting emerging stories from social media and news feeds** to indicate you have read and accepted the task specific conditions)

Emotional Mario: A Games Analytics Challenge
(Data: Toadstool (simula.no/datasets/toadstool))

We provide a multimodal dataset consisting of ten participants playing Super Mario Bros. The dataset contains their game input, demographics, sensor output from a medical-grade device, and videos of their faces while playing. The data is free to use for academic purposes.
Any use of the dataset will cite the following papers:

Henrik Svoren, Vajira Thambawita, Pål Halvorsen, Petter Jakobsen, Enrique Garcia-Ceja, Farzan Majeed Noori, Hugo L. Hammer, Mathias Lux, Michael Alexander Riegler, and Steven Alexander

Hicks. 2021. Toadstool: A Dataset for Training Emotional Intelligent Machines Playing Super Mario Bros. In Proceedings of the 11th ACM Multimedia Systems Conference (MMSys '20).

Signature_____

(sign here if participating in the **Emotional Mario: A Games Analytics Challenge** to indicate you have read and accepted the task specific conditions)

Emotions and Themes in Music
(Data: audio files, spectrograms, features and annotation labels licensed under Creative Commons.)
There are no additional task-specific conditions.

Signature_____
(sign here if participating in **Emotions and Themes in Music**)

FakeNews Detection: Corona Virus and Conspiracies Multimedia Analysis Task

(Data: Tweet IDs with human annotations for relevance to Corona Virus and various related conspiracies posts in English language)

Annotations are released for research under the CC BY 4.0 license. When downloading tweets by the means of the distributed Tweet IDs, users have to be compliant with Twitter's Developer Agreement and Policy (https://developer.twitter.com/en/developer-terms/agreement-and-policy).

Any use of the dataset will be accompanied by the citation of the following paper:

Schroeder, D. T., Pogorelov, K., & Langguth, J. (2019, October). FACT: a Framework for Analysis and Capture of Twitter Graphs. In 2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS) (pp. 134–141). IEEE.

Signature_____

(sign here if participating in the **FakeNews Detection: Corona Virus and Conspiracies Multimedia Analysis Task** to indicate you have read and accepted the task specific conditions)

Insight for Wellbeing: Cross-Data Analytics for (transboundary) Haze Prediction
(Data: The data that is provided will include daily readings of PM10 concentrations recorded by multiple weather stations in ASEAN countries (Brunei, Singapore, and Indonesia), covering periods between 2007 and 2018. The data also includes atmospheric parameters of daily (and in some cases hourly) values of temperature, rainfall, humidity, and wind speed/direction. However, not all parameters are available at every station/country, with some missing values in the data, which add to the challenge of this task. All data will be provided in CSV format, without inclusion of any Personal data. Task participants must sign the agreement for using the data prepared by MediaEval and NICT/UTB, which limit the use of the data for non-commercial research purposes only)

The use of this data set also requires two additional usage agreements (1) Individual Agreement Form, and (2) Organization Agreement Form. You will receive information from the task organizers (contact: dao@nict.go.jp and asem.kasem@utb.edu.bn), who will send you the agreements and ask you to return them separately as scanned .pdf files.

Signature_____

(sign here if participating in the **Insight for Wellbeing: Cross-Data Analytics for (transboundary) Haze Prediction** to indicate you have read and accepted the task specific conditions)

Medico: Transparency in Medical Image Segmentation
(Data: hyper-kvasir dataset (datasets.simula.no/hyper-kvasir), kvasir-Instrument (datasets.simula.no/kvasir-instrument))

The data is free to use for academic purposes. No patient identifiable data is included. All study participants gave consent on using the data for research. Data is anonymized following the Norwegian and European data protection regulations (fully anonymized, no key lists are available).

Any use of the dataset will cite the following papers:

Borgli, H., Thambawita, V., Smedsrud, P.H. et al. HyperKvasir, a comprehensive multi-class image and video dataset for gastrointestinal endoscopy. Sci Data 7, 283 (2020).

Jha, D., Ali, S., Emanuelsen, K., Hicks, S. A., Thambawita, V., Garcia-Ceja, E., Riegler, M. A., de Lange, T., Schmidt, P. T., Johansen, H. D., Johansen, D., & Halvorsen, P. (2021). Kvasir-Instrument: Diagnostic and Therapeutic Tool Segmentation Dataset in Gastrointestinal Endoscopy BT – MultiMedia Modeling. Springer International Publishing.

Signature_____

(sign here if participating in the **Medico: Transparency in Medical Image Segmentation** to indicate you have read and accepted the task specific conditions)

NewsImages
(Data: part of news articles (text snippet, headline), images, user/system behaviour statistics)

Usage agreement for the plista dataset related to the news recommendation subtask:

This agreement between plista GmbH Torstraße 33-35, 10119 Berlin, Germany and the downloading party imposes the following regulations:

1. plista allows certain interested parties to download the "plista dataset". For the usage and handling of this dataset the following rules must be accepted. By accepting this agreement the party warrants and represents that:
   a. Neither the dataset itself, nor single information based or extracted from the dataset are passed to any third party. For any third-party disclosure a prior permission is necessary.
   b. Usage of the dataset is exclusively for scientific purposes.
   c. The party does not infringe any applicable rights of plista or those of third parties e.g., but without limitation, copyrights, rights of trademarks or to a name.
   d. Any publication of results that derive out of scientific analysis of the "plista dataset" will only be of scientific purpose and will not pursue any other objective. Any publication will be made only in compliance with applicable data protection laws and follows the principle of data minimization.
   e. The dataset has to be referred to as "plista dataset" in scientific publications.
   f. If contestants use the server made available by plista, contestants guarantee that they will only use the provided server for purposes of the challenge and will refrain from any usage that is not associated with the challenge. Any action on the server that is not

associated with the challenge or the bare attempt to use it for any other purpose and any misuse will lead to the exclusion of the challenge. plista makes no warranties as to server availability. plista also reserves the right to assert claims of damage.

2. To the maximum extent permitted by law downloading party of the "plista dataset" agree to hold plista, its subsidiaries and affiliates and their respective directors, employees and assigns (collectively the released parties) harmless for any injury or damage caused or claimed to be caused by downloading, usage or publication, except to the extent that any death or personal injury is caused by the negligence of plista.

3. To the maximum extent permitted by law the party agree that plista are is hold harmless at all times from and against any liability, claims, demands, losses, damages, costs and expenses resulting from any act, default or omission of the party and/or breach of any warranty set forth herein.

4. By sending a short answer to the mail that contains this agreement, the downloading party accepts this agreement.

5. Downloading and usage of the "plista dataset" is only allowed if this agreement is accepted.

6. This agreement shall be governed by German law with the exception of the United Nations Convention on Contracts for the International Sale of Goods (CISG) and the International Private Law (IPR). If legally permitted the parties agree on the exclusive jurisdiction of the local courts in Berlin in respect of any disputes arising out of or in connection with this agreement.

Signature_____

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

Predicting Media Memorability
(Data: TRECVid 2019 Video to Text dataset)

The data is publicly available, and no agreement/license needed.

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 guarantee of any kind.

Any use of the dataset will be accompanied by the citation of the following papers:

Awad, G.M., Butt, A.A., Lee, Y., Fiscus, J., Godil, A., Delgado, A., Smeaton, A.F. and Graham, Y., 2019. Trecvid 2019: An evaluation campaign to benchmark video activity detection, video captioning and matching, and video search & retrieval.

We understand that the video media used for this task must be used in accordance with the terms of release of the original dataset.

Signature_____

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

Sports Video: Fine Grained Action Detection and Classification of Table Tennis Strokes from videos:

(Data: The data is the TTStroke dataset comprising videos of table tennis stroke exercises performed by sportsmen and xml files of annotations)

For this task, in addition to the usage agreement for the MediaEval 2021 Research Collections, a specific data usage agreement provided by University of Bordeaux has to be accepted electronically by task participants. You will receive information from the task organizers about this agreement. The data will be made available only after this acceptance. This agreement will be provided by the task organizers once this General Agreement has been signed and returned.

Signature_____

(sign here if participating in the task **Sports Video Annotation: Detection of Strokes in Table Tennis**)

Visual Sentiment Analysis: A Natural Disaster Use-case
(Data: Disaster images collected from social media.)

We understand that the image material used for this task must be used in accordance with the terms of release of the original datasets. Any use of the dataset will need to cite the following paper:

Hassan, Syed Zohaib, et al. "Visual Sentiment Analysis from Disaster Images in Social Media." arXiv preprint arXiv:2009.03051 (2020)

Signature_____

(sign here if participating in the **Visual Sentiment Analysis: A Natural Disaster Use-case** to indicate you have read and accepted the task specific conditions)

WaterMM: Water Quality in Social Multimedia
(Data: Tweet IDs with human annotation for relevance to water quality, safety and security)

Annotations are released for research under the CC BY-NC license. When downloading tweets by the means of the distributed Tweet IDs, users have to be compliant with Twitter's Developer Agreement and Policy (https://developer.twitter.com/en/developer-terms/agreement-and-policy).

Any use of the dataset will be accompanied by the citation of the following paper:

Moumtzidou, A., Andreadis, S., Gialampoukidis, I., Karakostas, A., Vrochidis, S., & Kompatsiaris, I. (2018, April). Flood relevance estimation from visual and textual content in social media streams. In *Companion Proceedings of the The Web Conference 2018* (pp. 1621-1627).

Signature_____

(sign here if participating in the **WaterMM: Water Quality in Social Multimedia** to indicate you have read and accepted the task specific conditions)