MediaEval

Multimedia Evaluation Benchmark

https://multimediaeval.github.io/

Second Metascience for Machine Learning Workshop

Hayley Hung

Metascience for Machine Learning

Meta science for ML refers to activities that describe the ***process*** of doing science. What and where is the process? This could span multiple levels of granularity from…

- …research design methodology across one or more projects
- … to how we educate the next generation
- … to how we organise scientific connection and collaboration
- … to reward structures to enable good scientific practice.

Hayley's slides from the First Metascience for Machine Learning Workshop

What is MediaEval?

- MediaEval offers challenges in machine learning related to multimedia data.
- Participants address these challenges by creating algorithms for analyzing, exploring and accessing information in the data.
- This year’s workshop 25-26 October in Dublin (and online) between CBMI and ACM Multimedia 2025.

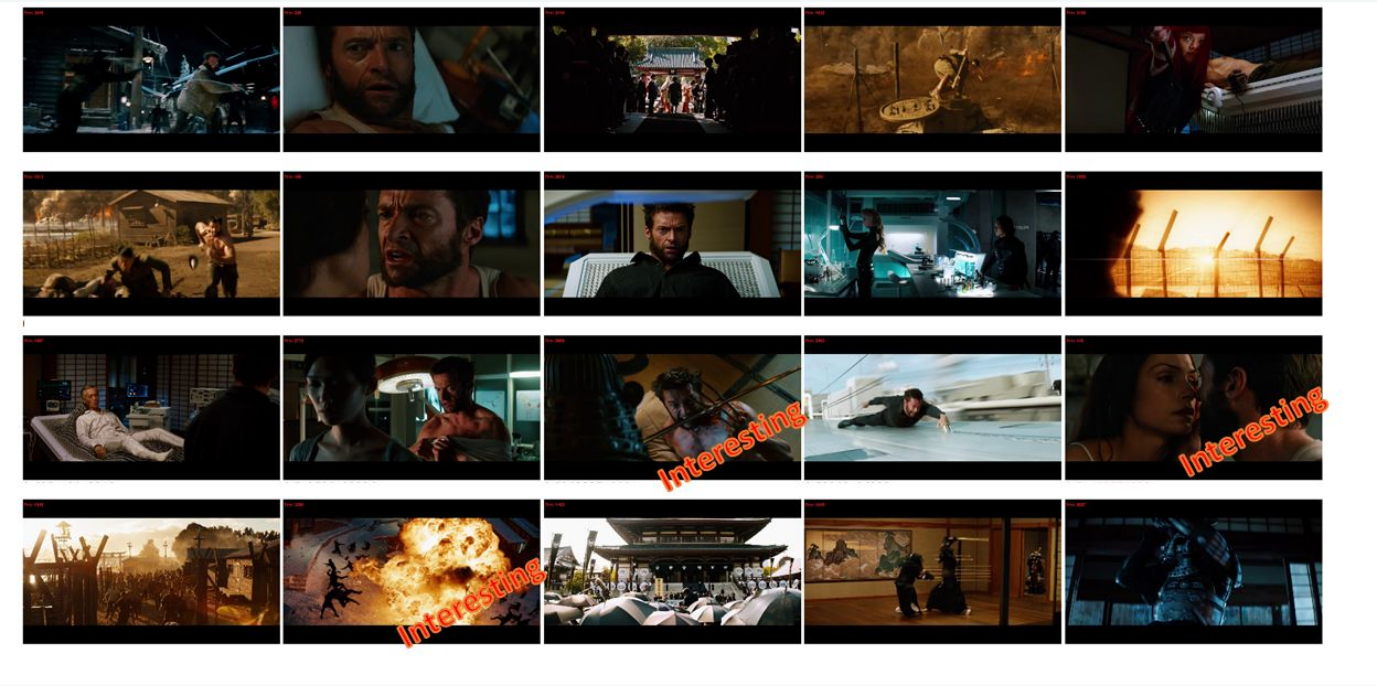
Predicting Media Interestingness Task (2016)

Automatically select frames or portions of movies which are the most interesting for a common viewer.

- Goal: Make use of the visual, audio and text content (features provided).
- Data: consists in ca 100 movie trailers, together with human annotations
- Metric: System performance is to be evaluated using standard Mean Average Precision.

http://multimediaeval.org

Predicting Media Interestingness Task



http://multimediaeval.org

Genesis of a MediaEval Task

- Call for tasks proposals, evaluated by the MediaEval coordination committee.
- Tasks must have:
  - use scenario, people-relevant and ethical
  - research questions
  - accessible, ethical dataset
  - good evaluation design
  - an organization team to run the task

MediaEval Participation

- **Download data**
- **Submit solution**

MediaEval Participation

- Give input on challenge design
- **Download data**
- Interact with organizers
- **Submit solution**
- Submit short working notes paper describing solution
- Present solution at the yearly MediaEval workshop
- Discuss with other task participants
- Aim to write a paper for mainstream venue
- Propose or organize tasks in the next yearly cycle

MetaMediaEval

- Generate new research questions.
- Promote comparability, to move forward the state of the art.
- Innovation evaluation design (metrics, ground truth creation)
- Get people involved in research, by providing a welcoming community.
- Keep on doing these things.

Medico

Visual Question Answering
(with multimodal explanations)
for gastrointestinal imaging

Medico Task Description

The goal is to use Visual Question Answering (VQA) to interpret and answer questions based on gastrointestinal images, aiming to enhance decision support and improve AI-driven medical decision-making. We provide a gastrointestinal dataset containing images and videos with VQA labels and additional metadata.

MultiSumm:

Multimodal summarization of multiple topically related websites

MultiSumm Task Description

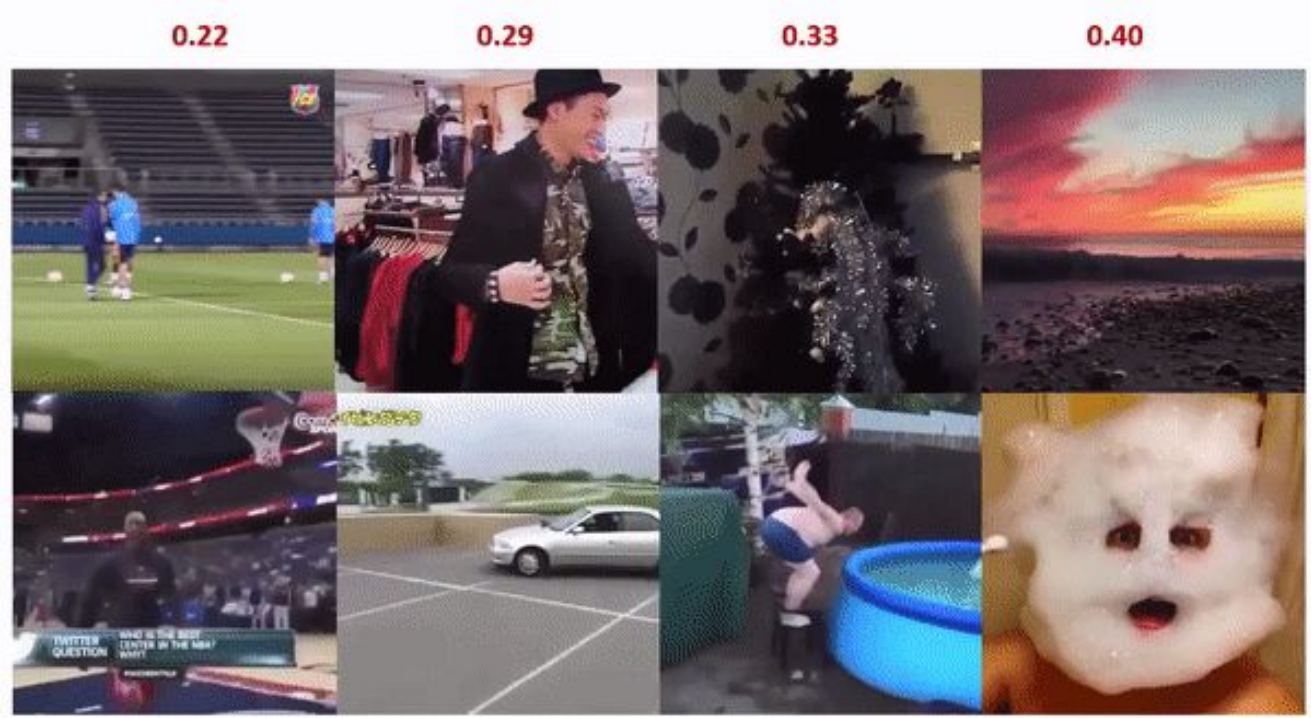
Participants are provided with multimodal web content from several cities listing food sharing initiatives (FSIs) in each city. For each city, participants are tasked with creating a multimodal summary of the FSI activities in the city that satisfy specified criteria. Evaluation will explore the use of emerging LLMs-based methods in automated assessment of multimodal multi-document summarization.

Memorability

Predicting movie and commercial memorability

Memorability Task description

The goal of this task is to study the long-term memory performance when recognizing small movie excerpts or commercial videos. We provide the videos, precomputed features or EEG features for the challenges proposed in the task such as How memorable a video, if a person familiar with a video or if you can predict the brand memorability?

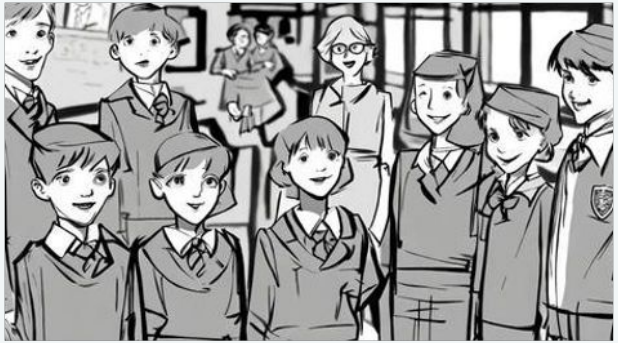

0.22
0.29
0.33
0.40
0.50
1.00
1.00
1.00

NewsImages

Retrieval and generative AI for news thumbnails

News Images Task Description

- We give you a news headline.
- You generate or retrieve a fitting image.
- The image fit is rated by your peers.
- For example…



Title: Bilston school takes a trip down memory lane as staff and pupils celebrate 90th anniversary.

Synthetic Images

Advancing detection of generative AI used in real-world online images

Task Description

The goal of this challenge is to develop AI models capable of detecting synthetic images and identifying the specific regions in the images that have been manipulated or synthesized. Approaches will be tested on images synthesized with state-of-the-art approaches and collected from real-world settings online.

Synthetic or Real?

ACM MM 2025 — Call for Reproducibility Papers

This page gives the instructions for submitting reproducibility companion papers at the 2025 edition of ACM Multimedia. Such a submission has essentially two parts:

Companion paper: The companion paper is 2–3.5 pages in length (with an optional page for references). It contains a high-level description of the experiments carried out in the original paper which are implemented in the archive.

Archive: Contains the artifacts (e.g., code, scripts, data sets, protocols), which are cleanly packaged, ready for download and use to replicate the results from the original paper. It contains detailed readme file(s), examples, and all information needed to successfully carry out the experiments.

General Submission Guidelines for 2025

Authors that have a main-conference paper published at ACM Multimedia 2023 and 2024 are invited to submit a short reproducibility companion paper to the ACM Multimedia Reproducibility track at ACM Multimedia 2025. That companion paper typically focuses on the technical details of what you published at ACM MM 2023/2024.

The companion paper should be submitted as a short paper that is 2–3.5 pages long, excluding references. It must follow the standard ACM style format, double column. It has to involve a majority of the authors of the original paper and provide in the clear their names and affiliations. The original ACM Multimedia

The ACM Conference Series on **Recommender Systems**

Follow us to keep up to date with the latest calls for contributions and news from ACM RecSys!

HOME | RECSYS 2025 | PAST CONFERENCES | HONORS | BLOG | CONTACT | search...

Call for Contributions

Instructions for preparing **camera-ready** versions of accepted papers can be found here.

Full Papers | Short Papers | Reproducibility | Doctoral Symposium | Industry | Tutorials | Workshops | Demos | Late-Breaking Results

Due to widespread power outages that affected several European Countries on April 28, the deadline for reproducibility paper submission is extended to: **May 16, 2025, 23:59 AoE**. The deadline for the abstract submission remains **May 6, 2025, 23:59 AoE**.

Call for Reproducibility Papers

The reproducibility of empirical results is a cornerstone of scientific research and a prerequisite to ensure continuous progress in our field. ACM RecSys 2025 therefore strongly encourages submissions that replicate or analyze prior works in similar or alternative settings or contribute to an increased level of reproducibility in the future.

RECSYS 2025 (PRAGUE)

- About the Conference
- Program at Glance
- Program
- Registration
- Accommodation
- Important Dates
- Call for Contributions
- Accepted Contributions
- Keynotes

Reproducibility Track at RecSys

https://recsys.acm.org/recsys25/call/#content-tab-1-2-tab

MediaEval History

Intro to MediaEval History

- MediaEval was founded in 2008 as part of CLEF.
- It became an independent benchmark in 2010.
- Anyone can propose a task or sign-up to participate in a task.
- Over the years, MediaEval tasks and data have been the basis for a large number of publications:
  - Workshop proceedings
  - Journal papers
  - Book chapters/theses
- Currently, we have counted: 1012 papers.

See: https://multimediaeval.github.io/bib

MediaEval 2010 Workshop

Santa Croce in Fossabanda
Pisa, Italy
24 October 2010
ACM Multimedia 2010

MediaEval 2011 Workshop

Santa Croce in Fossabanda
Pisa, Italy
1-2 September 2011
Interspeech 2011

MediaEval 2012 Workshop

Santa Croce in Fossabanda
Pisa, Italy
4-5 October 2012
ECCV 2012

MediaEval 2013 Workshop

Reial Acadèmia de Bones Lletres de Barcelona
18-19 October 2013
ACM Multimedia 2013

MediaEval 2014 Workshop

Reial Acadèmia de Bones Lletres de Barcelona
16-17 October 2014

MediaEval 2015 Workshop

Schloss Wurzen
Germany
14-15 September 2015
Interspeech 2015

MediaEval 2016 Workshop

Netherlands Institute for Sound and Vision
20-21 October 2016
ACM Multimedia 2013

MediaEval 2017 Workshop

Trinity College Dublin
13-15 September 2017
CLEF 2017

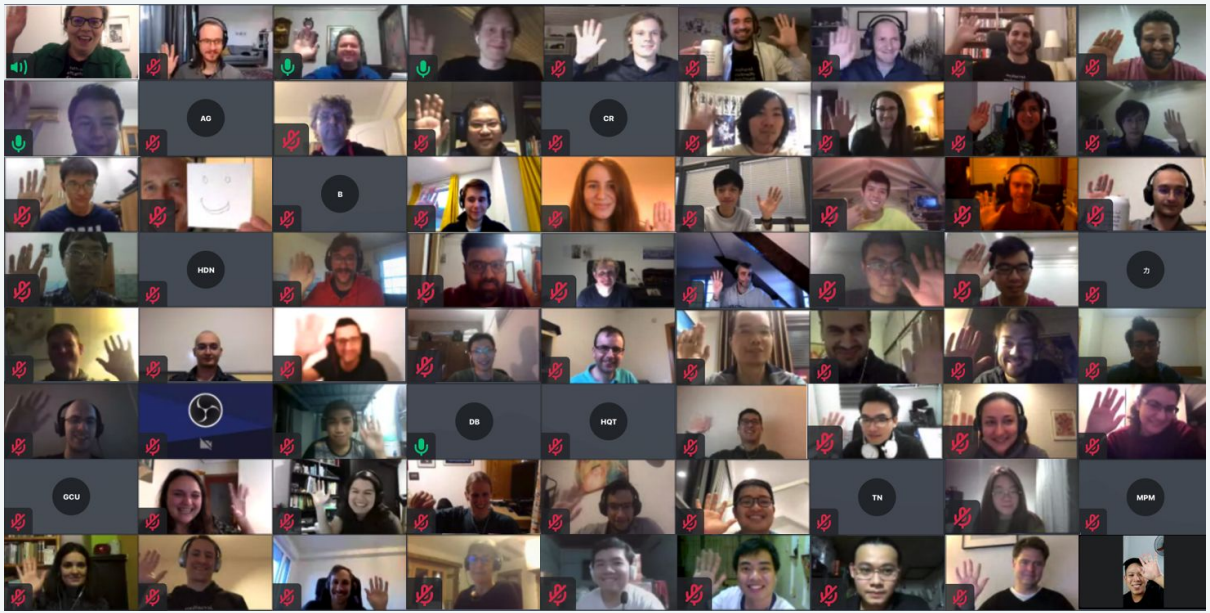
MediaEval 2018 Workshop

EURECOM
Sophia Antipolis, France
29-31 October 2018

MediaEval 2019 Workshop

EURECOM
Sophia Antipolis, France
27-29 October 2019
ACM Multimedia 2019

MediaEval 2020 Workshop



Online
11,14-15
December
2020

MediaEval 2021 Workshop

Online
13-15
December
2021

MediaEval 2022 Workshop

Bergen, Norway and Online
12-13 January 2023

MediaEval 2023 Workshop

Amsterdam, Netherlands and Online
1-2 Feb 2024
MMM 2024

MediaEval Culture

Hayley Hung

Metascience for Machine Learning

Meta science for ML refers to activities that describe the ***process*** of doing science. What and where is the process? This could span multiple levels of granularity from…

- …research design methodology across one or more projects
- … to how we educate the next generation
- … to how we organise scientific connection and collaboration
- … to reward structures to enable good scientific practice.

Hayley's slides from the First Metascience for Machine Learning Workshop

Why do ML researchers ask *Why?* so little?

Marco Loog

Marco's slides from the First Metascience for Machine Learning Workshop

Cynthia Liem

What to do?

- I really think our academic publication model has stretched beyond its limits, and needs considerable revision
- Can we create spaces incentivizing reflection and learning, while still upholding forms of quality control? (can we draw inspiration from open-source software development?)
- Can we more explicitly teach our students awareness of different methodological schools of thought and questions of data/ground truth quality?
- Can we slow down?

Cynthia's slides from the First Metascience for Machine Learning Workshop

MediaEval “Quest for Insight”

Incentivizing people to go beyond the evaluation score to get insight into the data and the task.

- MediaEval Distinctive Mentions
- Questions in Advance
- “Quest for Insight” papers

MediaEval Distinctive Mentions

- Each year, we award “MediaEval Distinctive Mentions” to teams who stand out with respect to risk, insight, creativity, or kick (RICK).
- Such teams develop approaches with potential for the future, but may not score high on the leaderboard.
- MediaEval Distinctive Mentions are part of our larger “Quest for Insight” to deepen our understanding of how and why solutions work.

MediaEval Distinctive Mentions

Risk, innovation, creativity, kick.

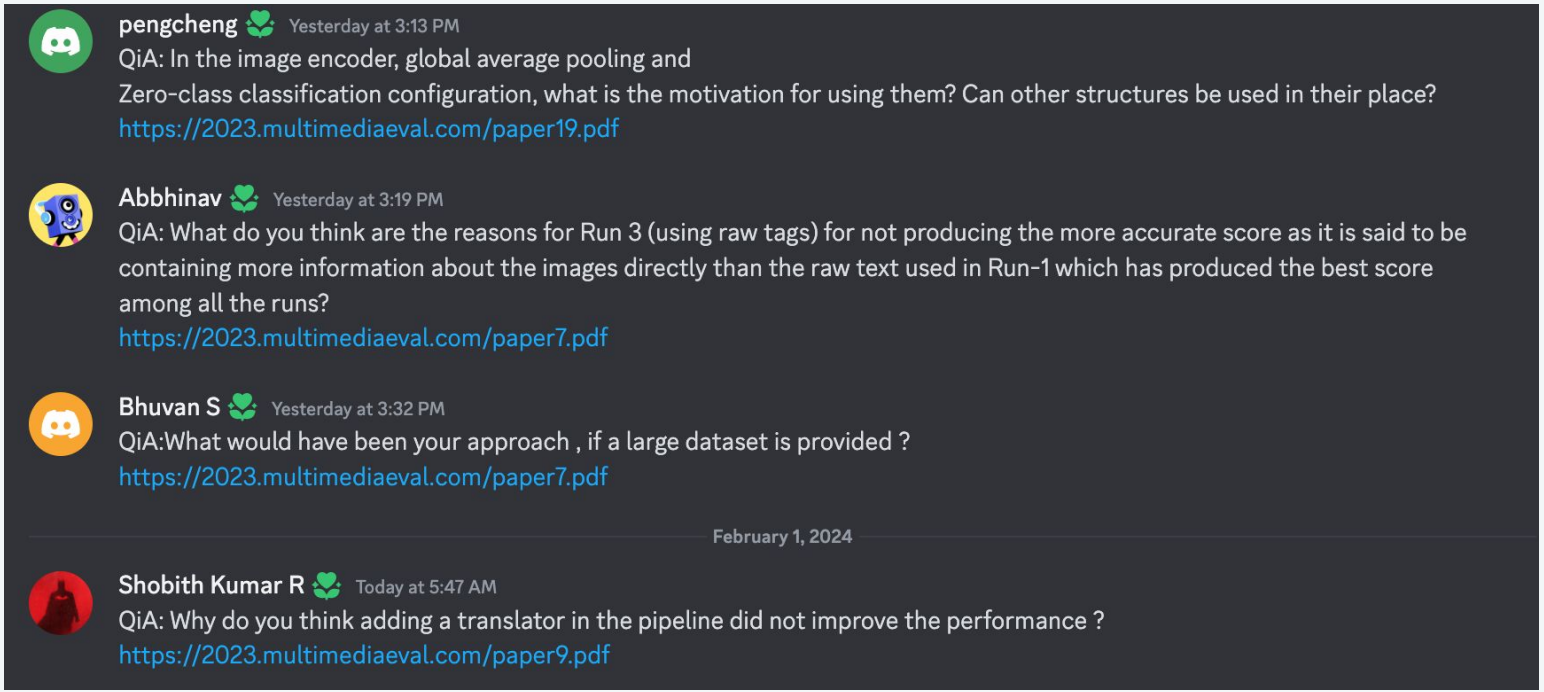
2016

2019

Questions in Advance

- Participants are assigned other papers from other tasks in advance of the workshop.
- They are asked to formulate questions in advance of the workshop in order to promote discussion.

Examples in the #NewsImages channel on Discord



“Quest for Insight” Papers

Papers in the proceedings that analyze the task, the data, and the metric, rather than report performance on the task.

NewsImages

Overview paper:

- The Relation between Texts and Images in News: News Images in MediaEval 2023
  Andreas Lommatzsch, Benjamin Kille, Özlem Özgöbek, Mehdi Elahi, Duc Tien Dang Nguyen

Working notes papers:

- Beyond Keywords: ChatGPT's Semantic Understanding for Enhanced Media Search
  Hoang Chau Truong Vinh, Doan Khai Ta, Duc Duy Nguyen, Le Thanh Nguyen, Quang Vinh Nguyen
- Prompt-based Alignment of Headlines and Images Using OpenCLIP
  Lucien Heitz, Yuin Kwan Chan, Hongji Li, Kerui Zeng, Luca Rossetto, Abraham Bernstein
- Multimodal Fusion in NewsImages 2023: Evaluating Translators, Keyphrase Extraction, and CLIP Pre-Training
  Tien-Huy Nguyen, Hoang-Long Nguyen-Huu, Thien-Doanh Le, Huu-Loc Tran, Quoc-Khanh Le-Tran, Hoang-Bach Ngo, Minh-Hung An, Quang-Vinh Dinh
- Ensemble Pre-trained Multimodal Models for Image-text Retrieval in the NewsImages MediaEval 2023
  Taihang Wang, Jianxiang Tian, Xiangrun Li, Xiaoman Xu, Ye Jiang
- Cross-modal Networks, Fine-Tuning, Data Augmentation and Dual Softmax Operation for MediaEval NewsImages 2023
  Antonios Leventakis, Damianos Galanopoulos, Vasileios Mezaris
- Investigating the Performance of the CLIP Model and Concept Matching in Text-Image Retrieval Systems
  Xiaomeng Wang, Mingliang Liang, Martha Larson
- Integrated Multi-stage Contextual Attention Network for Text-Image Matching
  Jiande Sun, Yi Shao, Yawen Chen, Yang Zhang, Tianlin Zhang, Xuan Zhang, Ye Jiang, Jing Li
- NewsImages Fusion: Bridging Textual Context and Visual Content in Media Representation
  Arvind V, Vettri Chezhiyan, Harish J, Dr. Priyadharshini R, Mohanapriya E
- Optimizing Visual Pairings: A CLIP Framework for Precision News Image Rematching
  Pooja Premnath, Venkatasai Ojus Yenumulapalli, Rajalakshmi Sivanaiah, Angel Deborah Suseelan
- Connecting Text and Images in News Articles using VSE++
  Abbhinav Elliah, Mirunalini P, Keerthick V, Haricharan Bharathi, Anirudh Bhaskar, Vithula S

Quest for insight papers:

- An Empirical Exploration of Perceived Similarity between News Article Texts and Images
  Lucien Heitz, Abraham Bernstein, Luca Rossetto

MediaEval Meditations

Moving towards 2026

- Despite multi-year efforts, we have not arrive at a set of “recommended reading” papers that serve as golden examples of machine learning research practices.
- Despite efforts, we have a low bus factor: Single point of failure.
- Is our legacy lists?

Key Techniques for Challenges (2016)

- **Multiple decision makers:** involve everyone in task and design choices.
- **Multi-dimensional evaluation:** use more than one metric.
- **Multi-institution organization teams:** 3-5 organizers.
- **Multi-faceted products:** data, code, tools.
- **Multi-author papers:** synthesis papers, data set papers, papers for reproducibility tracks.

ABC of Challenges (2016)

- **Analyze**: infinitely.
- **Benchmark**: infallibly.
- **Collaborate**: infectiously
- **Dare**: informedly, especially with *Data*.
- **Explain**: informally.
- **Fail**: informatively.

A List (2025)

- Actual aim
- Analysis
- Accuracy
- Addition
- Assumptions
- Alternatives
- Appropriateness

A closed list of the topics I talk about when I talk about research.

Q&A