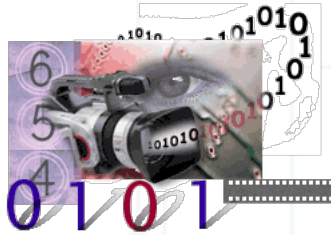
Final Items and Planning for TRECVID 2022

National Institute of Standards and Technology
U.S. Department of Commerce

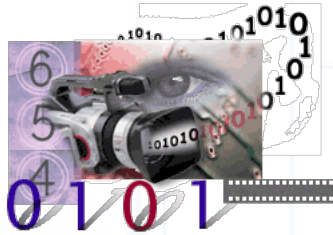
Information Access Division
Information Technology Laboratory

Reminders



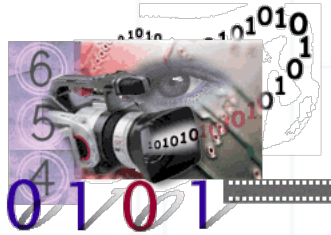
- Please email an **updated final version** of your notebook paper by 28 Feb. 2022 to gawad@nist.gov
- If your draft notebook paper is considered FINAL, **then also let us know!**
- Please use the citations guidelines to give credit for resources used (see trecvid.nist.gov)
- If you gave a **presentation**, please email us a **pdf copy** to put it on the website proceedings page (OR post it on slack channel of the task)
- Please fill in the conference survey sent to you to get feedback about TRECVID 2021

Reminders



- Please send new bibliography of your peer-reviewed work using the TRECVID resources to gawad@nist.gov
- The call for participation will go out early Feb. 2022 with applications due end of May. **Please forward the call to your interested colleagues & network, and post to mailing lists.**
- **Reminder: you must apply for TRECVID call to participate each year (even for regular teams).**
- Please feel free to propose "New & Needed Tasks" - *due Sep, 2022 to start in 2023)*

New Task Proposals Should



- State the goals of the **real world** task being modeled
  - What user task is being modelled; what problem(s) will be addressed; who, besides researchers, cares about it?
  - What is expected to be learned?
  - What valuable infrastructure would be created if the task were run?
- Describe the system task in terms of inputs and outputs
- Define the evaluation procedure and measure(s)
  - Where would assessing/annotation occur (NIST? elsewhere?).
  - Are there are any special qualifications or domain expertise required of the assessors/annotators?
  - How much effort would such annotation/assessing require?
- Identify development and evaluation data (Type, Quantity, Availability)
- Who can serve as the task coordinator (external, NIST, external+NIST)

TRECVID 2022

DIGITAL VIDEO RETRIEVAL at NIST

~Dates: Nov. 16 – 18 (if in-person/hybrid)
Dec. 6 – 9 (if Virtual)

Slack channels for TRECVID 2022 active participants

Tentative 2022 Plans [Track 1]

Ad-hoc Video Search (AVS)

- **Default plan:**
  - Employ V3C2 dataset for new query development effort (90 queries total)
  - Release & eval 30 new ad-hoc Queries in 2022
  - Plan for new round of 20 progress query set (2022 -2024)
    - Evaluate progress on 2 years (2022-2023) in 2023 on 10 progress queries
    - Evaluate progress on 3 years (2022-2024) in 2024 on 10 progress queries
  - Review queries to add a bit of non-trivial cases (e.g. Boolean, preposition, relational, spatial/temporal aspects, etc)
  - Consider new run type (concept-based only)
  - Run type A (trained on iacc.3 -> V3C1)

- Teams are highly encouraged to take advantage of the progress subtask
- **We welcome any further feedback/suggestions**

Tentative 2022 Plans [Track 2]

Video-to-Text (VTT)

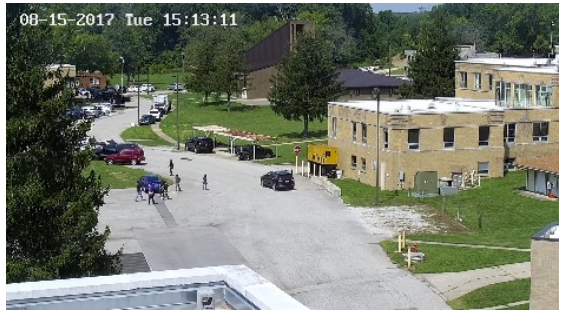
- **Default Plan:**
  - Continue with V3C2 dataset
  - Continue Description Generation
  - Progress task will continue through 2023 with the selected 300 videos
- Decide on the second subtask
  - Do we continue with Fill-in-the-Blanks? Would teams be more interested in another subtask:
    - VQA – to evaluate how well the systems understand a video
    - Dense video captioning – Uses longer videos with possibly multiple events
    - Something else?
- **We welcome any further feedback/suggestions**

Tentative 2022 Plans [Track 3]

TRECVID'21 Activities in Extended Videos (ActEV)

- **Default plan:**
  - Plan to leverage MEVA data (mevadata.org) in addition to VIRAT data (viratdata.org) for the task
  - More MEVA data will be released to the public
  - Support a deeper understanding of system behavior via interactive data analysis
  - Additional activity detection metric will be supported: mean Average Precision (mAP), AUDC based on spatio-temporal overlap in addition to just temporal overlap
  - NIST and the task participants will jointly write a publication

- We welcome any further feedback/suggestion



Tentative 2022 Plans [Track 4]

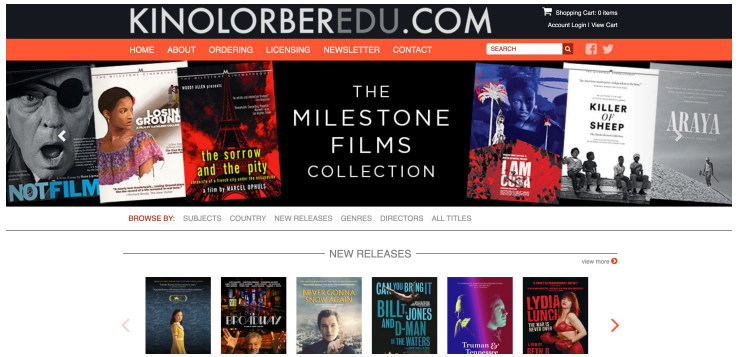
Disaster Scene Description & Indexing (DSDI)

- **Default plan:**
  - Employ new natural disaster event (e.g. Hurricane Ida) as testing dataset
    - Should be more comparable to LADI training dataset
  - Reach out to 2020/2021 teams to share their LADI refined annotations
  - Include source of imagery in testing dataset (drones, CAP)
- **Subject to funding:**
  - Plan for LADI 2.0+ (enhanced labels, additional disaster events, bounding boxes/segmentation, improved documentation/tutorials)
- **LADI-Community**
  - Community branch of LADI will collect user-submitted data CC-BY-4.0, annotations, refinements, and code
  - ladi-dataset-admin@mit.edu
- **We welcome any further feedback/suggestions**

Tentative 2022 Plans [Track 5]

Video Summarization (VSUM) --> Movie Summarization (MSUM)

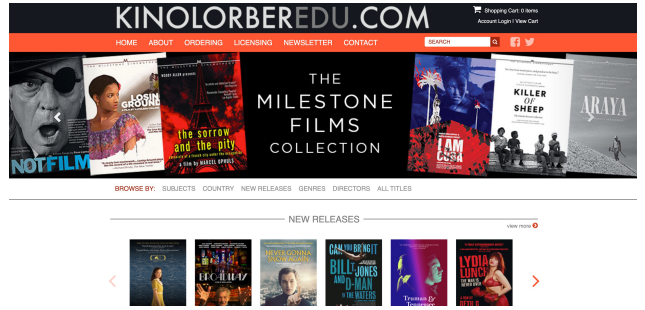
- System Task: Given a Movie and image/video examples of a character:
  - Systems are asked to find key-fact events of the character role & submit:
    - Video Summary of these key-fact events with max X durations
    - Text Summary of these key-fact events with max Y sentences / Z words
    - Run types based on using only movie vs external data sources
  - Evaluations are based on **objective** & **subjective** measures In both modalities.
  - Evaluation of separate submitted scenes should be supported.
- Dataset (2 available sources):
  - Self collected 14 CC movies (17.5 hr) from vimeo and IA.
  - Licensed set of 10 movies from kinolerberedu.
  - Division of training/testing is yet to be planned.
- We welcome any further feedback/suggestions.



Tentative 2022 Plans [Track 6]

Deep Video Understanding (DVU) [NEW TASK]

- Follow the same annotation framework as done in 2021 ACM MM Grand Challenges.
- Training dataset : 14 movies used in 2020 and 2021 ACM MM GCs including their annotations.
- Testing dataset : 10 new movies.
- Two main subtasks:
  - **Movie-level Track**:
    - **Required Query Type: QA** on part of knowledge graphs for selected movies.
    - **Optional Query Types:** Given a partial graph, systems will be asked to fill in the graph space.

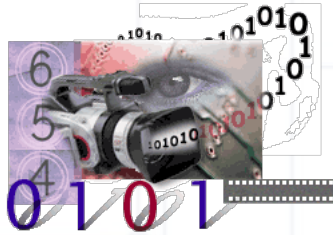


  - **Scene-level Track**:
    - **Required Query Type:** Find next/previous **interaction**, **given two people**, **specific scene**, & interaction between them.
    - **Optional Query Types:** Find a unique scene given a set of interactions and a scene list.

- After annotations are generated, there is a chance to share a subset of the ontology classes that annotators actually used (thus, simplifying teams efforts to focus on the important subset of classes)

General Suggestions/Questions

- *How can we encourage more collaboration across teams?*
  - *Sharing baseline systems, model scores, annotated data, etc*
  - *Please send suggestions for new software development resources (e.g. gitHub repos, etc) to be added to the tools & resources link at our website: http://wwwnlpir.nist.gov/projects/trecvid/tv.resources.html*
- *How can we attract more teams to join TRECVID ? Specifically to get more teams to submit runs?*
- *Any interested volunteers to be task coordinators along with NIST personnel?*
- *Any general comments or feedback to the organizers ?*



DIGITAL VIDEO RETRIEVAL at NIST

Thank You!
See you again in TRECVID 2022

In-person ?
Hybrid ?
Virtual ?