

# TREC Video Retrieval Evaluation

## TRECVID 2022

Ian Soboroff<sup>1</sup>, George Awad<sup>1</sup>, Asad Butt<sup>13</sup>, Keith Curtis<sup>1</sup>, Angela Ellis<sup>1</sup>, Darrin Dimmick<sup>1</sup>, Jonathan Fiscus<sup>4</sup>, Afzal Godil<sup>4</sup>

Yooyoung Lee<sup>4</sup>, Andrew Delgado<sup>4</sup>, Jesse Zhang<sup>4</sup>, Lukas Diduch<sup>45</sup>, Eliot Godard<sup>4</sup>, Baptiste Chocot<sup>4</sup>, Yvette Graham<sup>6</sup>, Georges Quénot<sup>7</sup>, Jeffrey Liu<sup>8</sup>

<sup>1</sup>Retrieval Group / <sup>4</sup>Multimodal Information Group  
Information Access Division, Information Technology Laboratory, NIST

<sup>3</sup>John Hopkins University

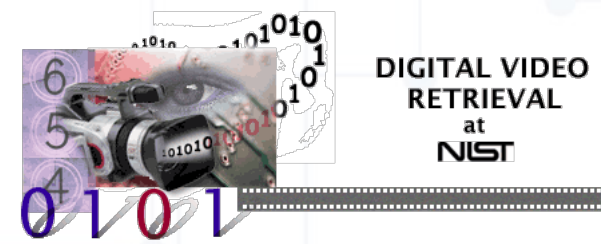
<sup>5</sup>Dakota Consulting, Inc

<sup>6</sup>Trinity College Dublin, Ireland

<sup>7</sup>Laboratoire d'Informatique de Grenoble, France

<sup>8</sup>MIT Lincoln Laboratory

# What is TRECVID ?



Workshop series  
(2001 – present)

Started as a track in the TREC (Text REtrieval Conference) evaluation benchmark

Became an independent evaluation benchmark since 2003

Focus: **content-based** video analysis, retrieval, detection, etc

Provides data, tasks, and uniform, appropriate scoring procedures

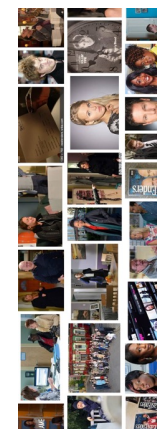
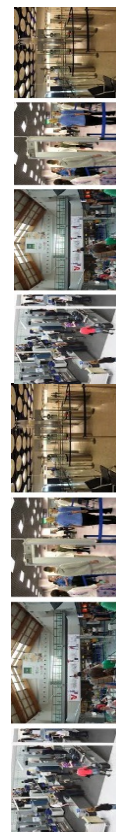
Aims for realistic system tasks and test collections using:

- Unfiltered data
- Focus on relatively high-level functionality (e.g. interactive search)
- Measurement against human abilities

Forum for:

- Exchange of research ideas
- The discussion of research methodology – what works, what doesn't, and why

2022  
2021  
2020  
2019  
2018  
2017  
2016  
2015  
2014  
2013  
2012  
2011  
2010  
2009  
2008  
2007  
2006  
2005  
2004  
2003



# Datasets (2003 – 2022)

US TV  
News

Int. TV  
News

Dutch  
TV Prog.

Security  
Cameras

Web  
Video

Soap  
Opera

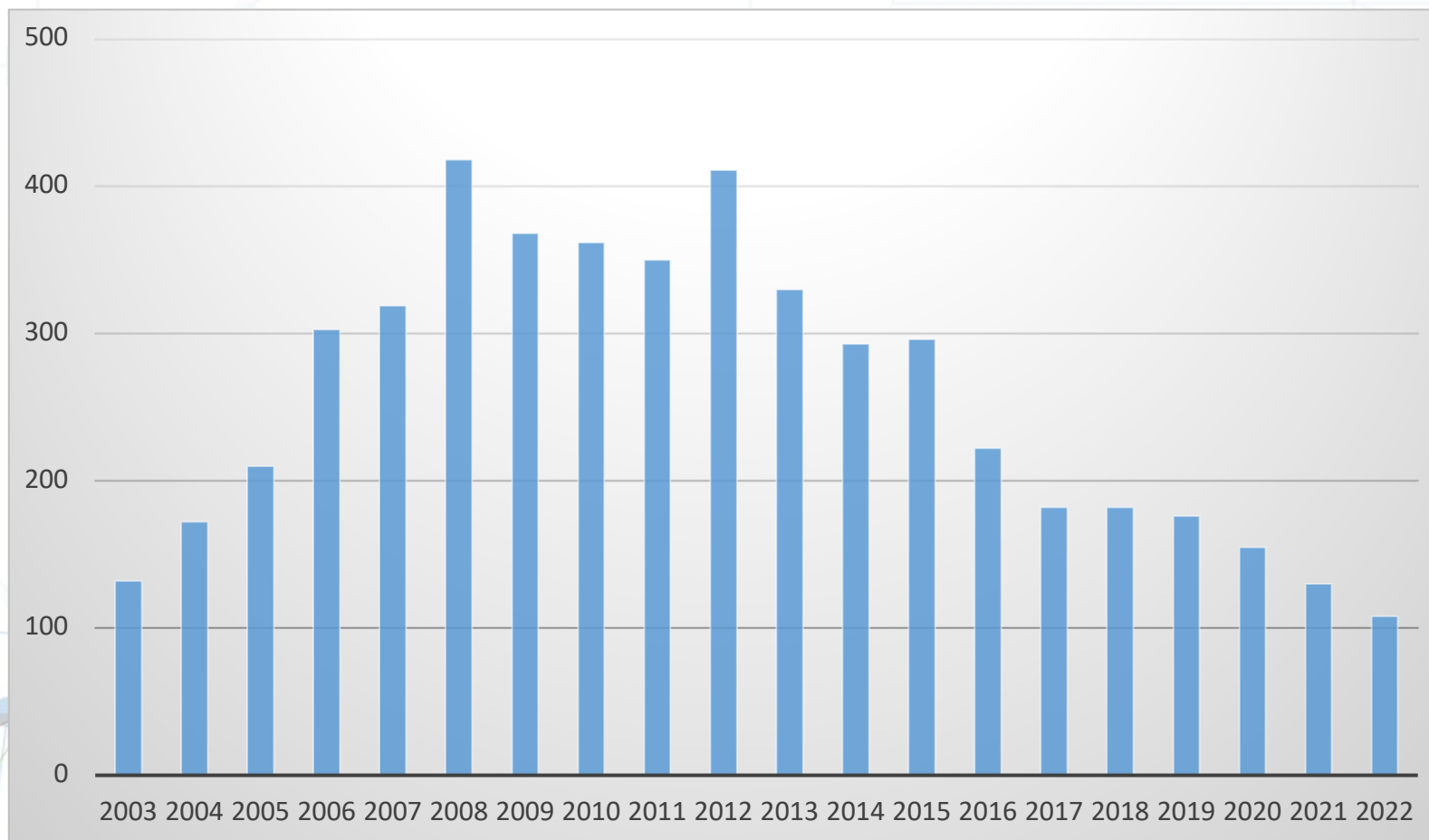
Social  
Media

Civil Air  
Patrol

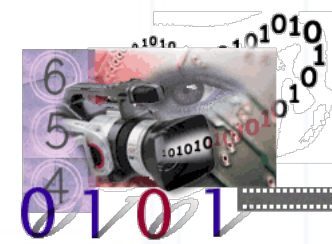
# Annual Workshop Authors



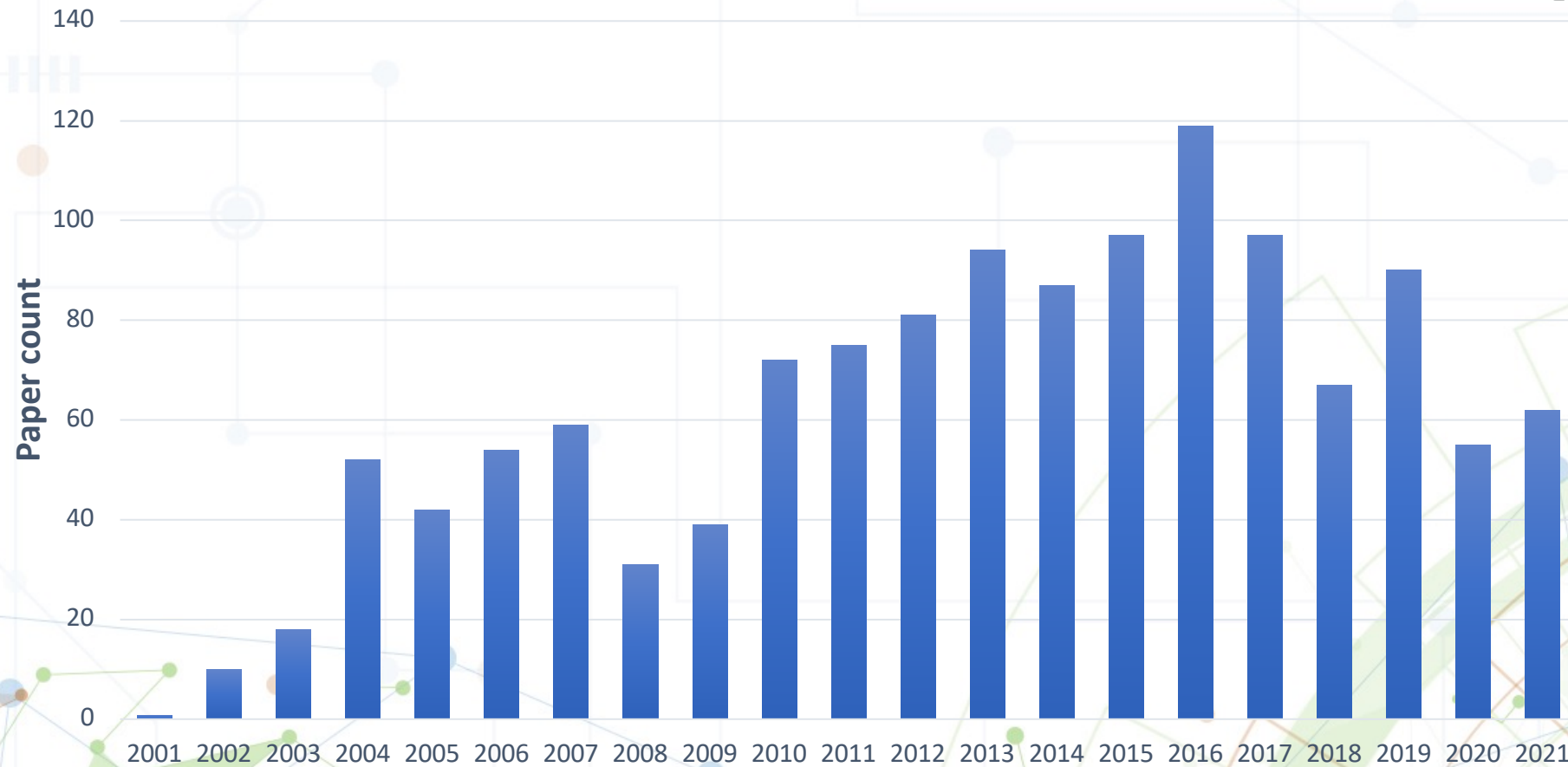
DIGITAL VIDEO  
RETRIEVAL  
at  
NIST



# Bibliography



DIGITAL VIDEO  
RETRIEVAL  
at  
NIST



<http://www-nlpir.nist.gov/projects/trecvid/trecvid.bibliography.html>



# TRECVID 2022 Tasks & Data



Ad-hoc Video  
Search (AVS)

Deep Video  
Understanding  
(DVU)

Video to Text  
(VTT)

Movie  
Summarization  
(MSUM)

Activities in  
Extended  
Videos (ACTEV)

Disaster Scene  
Description &  
Indexing (DSDI)

Vimeo CC  
(V3C2)

1300 h testing data,  
~1.4 M shots, title,  
keywords,  
description

~3000 h dev data

30 textual queries

20 progress queries  
[fixed 2022 – 2024]

KinoLorber

14 CC movies (~17 hr)  
dev set

6 licensed movies (8.5  
hrs) test set

Movie-level queries  
(targeting relations  
between  
people/locations)

Scene-level queries  
(targeting interactions)

Vimeo CC  
(V3C1)

2008 Short Vimeo  
testing video clips

10,862 dev videos and  
captions from Twitter  
Vine and Flickr

Main task:  
- Generate automatic  
video description  
- Progress task (2021-  
2023)

KinoLorber

10 licensed movies  
(5 testing, 5 training).

Task: Summarize movie  
character key-fact  
events in text & video  
summaries.

Testing queries:  
image samples of 8  
surprise characters  
from 5 movies.

MEVA

*16 h multi-camera  
video surveillance  
dataset*

*Detection and  
Localization of  
activities*

*20 activities evaluated*

LADI

20,000+ dev images  
from airborne low  
altitude disaster  
events + 11 hrs DSDI  
test data (2020-2021)

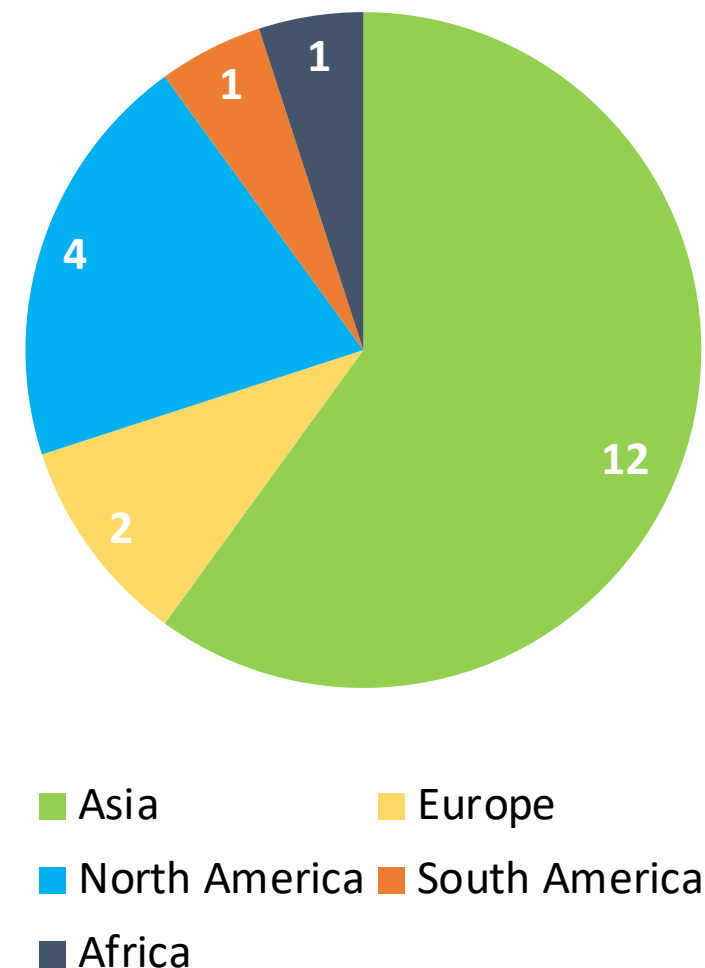
~6 h testing data from  
different natural  
disaster events

32 known features  
identified by search  
& rescue personnel

# TRECVID 2022 Finishers (20 / 35)

Groups Finished	Task Code	Task Name
7	AVS	Ad-hoc Video Search
6	ACTEV	Activities in Extended Video
6	VTT	Video to Text Description
3	DVU	Deep Video Understanding
2	DSDI	Disaster Scene Description & Indexing
1	MSUM	Movie Summarization

Unique finishing teams



# TRECVID 2022 Finishers (20 / 35 teams)



Team	Organization
Adapt	ADAPT Research Centre, Dublin City University
BUPT_MCPRL	Beijing University of Posts and Telecommunications
VIDION	Carnegie Mellon University
ELT_01	Elyadata
kindai_ogu_osaka	Kindai University; Osaka Gakuin University; Osaka University
columbia_graphen	Graphen,Inc Columbia University
MLVC_HDU	Hangzhou Dianzi University
ITI_CERTH	Information Technologies Institute, Centre for Research and Technology Hellas
kslab	Nagaoka University of Technology
NII_UIT	National Institute of Informatics, Japan; University of Information Technology, VNU-HCMC, Vietnam
WHU_NERCMS	Wuhan University
PKU_WICT	Peking University
RUCAIM3-Tencent	Renmin University of China
RUCMM	Renmin University of China
VIREO	Singapore Management University City University of Hong Kong
CamiloUchile	University of Chile
UMKC	University of Missouri-Kansas City; Florida International University
WasedaMeiseiSoftbank	Waseda University; Meisei University ; SoftBank Corporation
UMD	University of Maryland; John Hopkins University
TokyoTech	Tokyo Institute of Technology

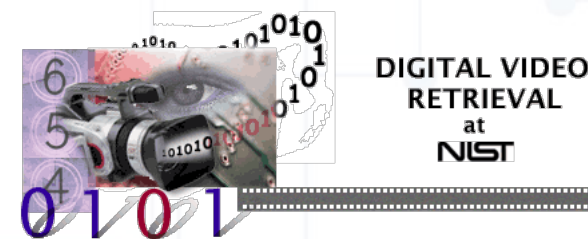


# TRECVID 2022 Non-Finishers (15 / 35)



Team	Organization
alisc_video	Alibaba group
Arete	Arete_Associates
drylwn	drylwn_visual_intelligence
EURECOM	EURECOM
spacetime_memory	Facebook AI Research
AIV4	Hitachi, Ltd.
INF	Language Technologies Institute - Carnegie Mellon University
ActiVisionLinks	Links Foundation
LIG	Multimedia Information Modeling and Retrieval group of LIG Explainable and Responsible Artificial Intelligence Chair of the MIAI Institute.
OzuCod	Ozyegin University
SEUGraphDy	Southeast University
MMGofLDS	University of Science and Technology of China
MiguelUA	University of Alicante, Alicante, SPAIN.
USF_EE	University of South Florida
yorku22	York University

# Support



The running of TRECVID 2022 has been funded directly by:

- National Institute of Standards and Technology (NIST)
- Intelligence Advanced Research Projects Activity (IARPA)
- NIST Information Technology Laboratory Building The Future program

TRECVID is only possible because of the additional efforts of many individuals and groups around the world.

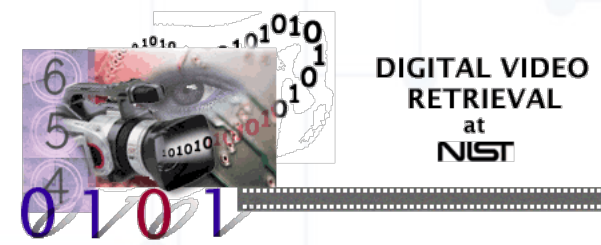


# NIST Assessors

TRECVID 2022



# Resources & Contributions



- Luca Rossetto and Heiko Schuldt (University of Basel) who Provided the V3C dataset, master shot reference and metadata
- Georges Quénot at LIG who provided the master shot reference for the IACC.3 video
- The LIMSI Spoken Language Processing Group and Vocapia Research provided ASR for the IACC.2-3 videos and HAVIC data
- Koichi Shinoda of the TokyoTechCanon team agreed to host a copy of the IACC.2 data
- Yvette Graham (DCU, Ir) for supporting the video-to-text Direct Assessment annotations
- Andrew Weinert and Jeffrey Liu (MIT Lincoln Laboratory) for supporting the LADI dataset
- KinoLorberEdu for licensing the movies dataset for the DVU task.

# Agenda at a glance



## 1 Day 1 (Tuesday)

Ad-hoc Video Search

Keynote Talk

OpenMFC

## 3 Day 3 (Thursday)

Activities in Extended Videos

Disaster Scene Description & Indexing

## 2 Day 2 (Wednesday)

Movie Summarization

Video to Text

Keynote Talk

OpenMFC

## 4 Day 4 (Friday)

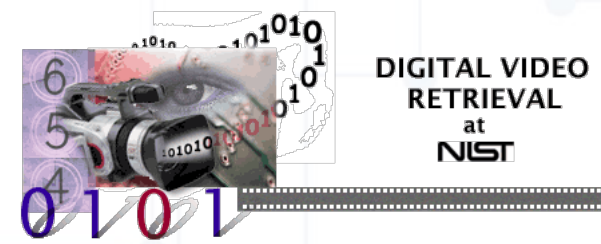
Keynote Talk

Deep Video Understanding

New Task Proposals

Workshop Conclusion

# Miscellany



- OpenMFC Sessions running after TRECVID (Tues – Wed)
- Slack
  - 2022 TrecVid Track Channels
  - Agenda/Schedule Channel
  - BlueJeans & Google (breaks & after workshop hours) Meeting links
  - Random
- Online workshop proceedings:
  - Teams papers, task information, datasets, ground truth, etc) can be found at:
  - <https://www-nlpir.nist.gov/projects/tv2022/workshop/tv22.workshop.notebook/>