

# TREC Video Retrieval Evaluation

## TRECVID 2023

Ian Soboroff<sup>1</sup>, George Awad<sup>1</sup>, Asad Butt<sup>1,2</sup>, Keith Curtis<sup>3</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>, Lukas Diduch<sup>4,5</sup>, Eliot Godard<sup>4</sup>, Yvette Graham<sup>6</sup>, Georges Quénot<sup>7</sup>, Deepak Gupta<sup>8</sup>, Dina Demner Fushman<sup>8</sup>

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

<sup>2</sup>John Hopkins University

<sup>3</sup>Technological University of the Shannon, Ireland

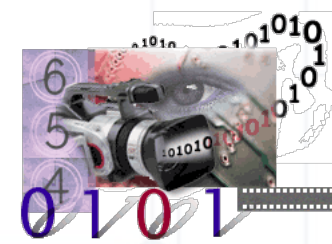
<sup>5</sup>Dakota Consulting, Inc

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

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

<sup>8</sup>National Institute of Health, USA

# What is TRECVID ?



DIGITAL VIDEO  
RETRIEVAL  
at  
NIST

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

2023  
2022  
2021  
2020  
2019  
2018  
2017  
2016  
2015  
2014  
2013  
2012  
2011  
2010  
2009  
2008  
2007  
2006  
2005  
2004  
2003



US TV News



Int. TV News



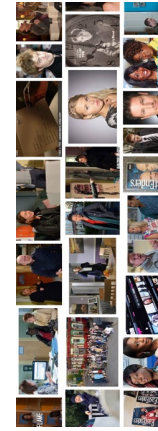
Dutch TV Prog.



Security Cameras



Web Video



Soap Opera

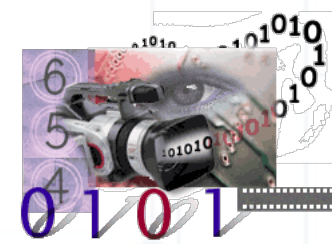


Social Media

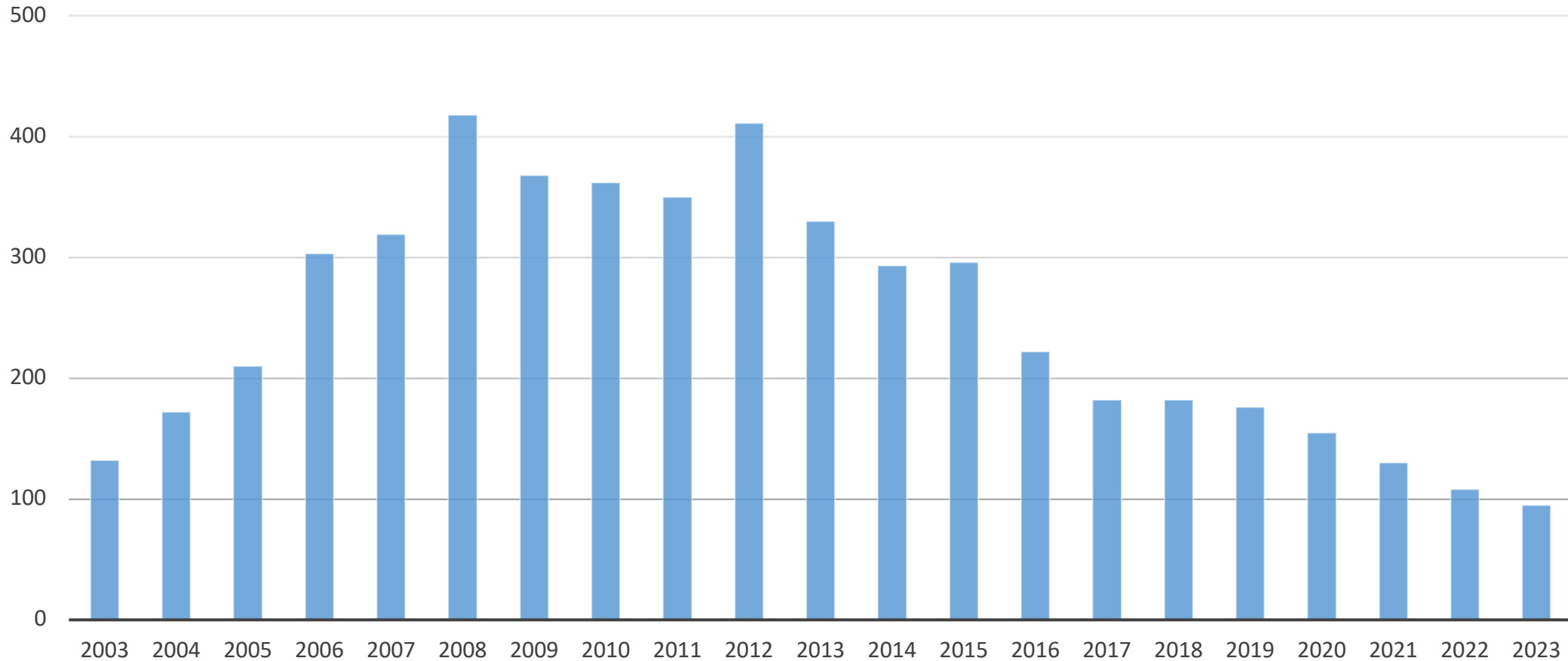
Civil Air Patrol

# Datasets (2003 - 2023)

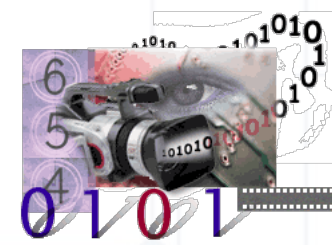
# Annual Workshop Authors



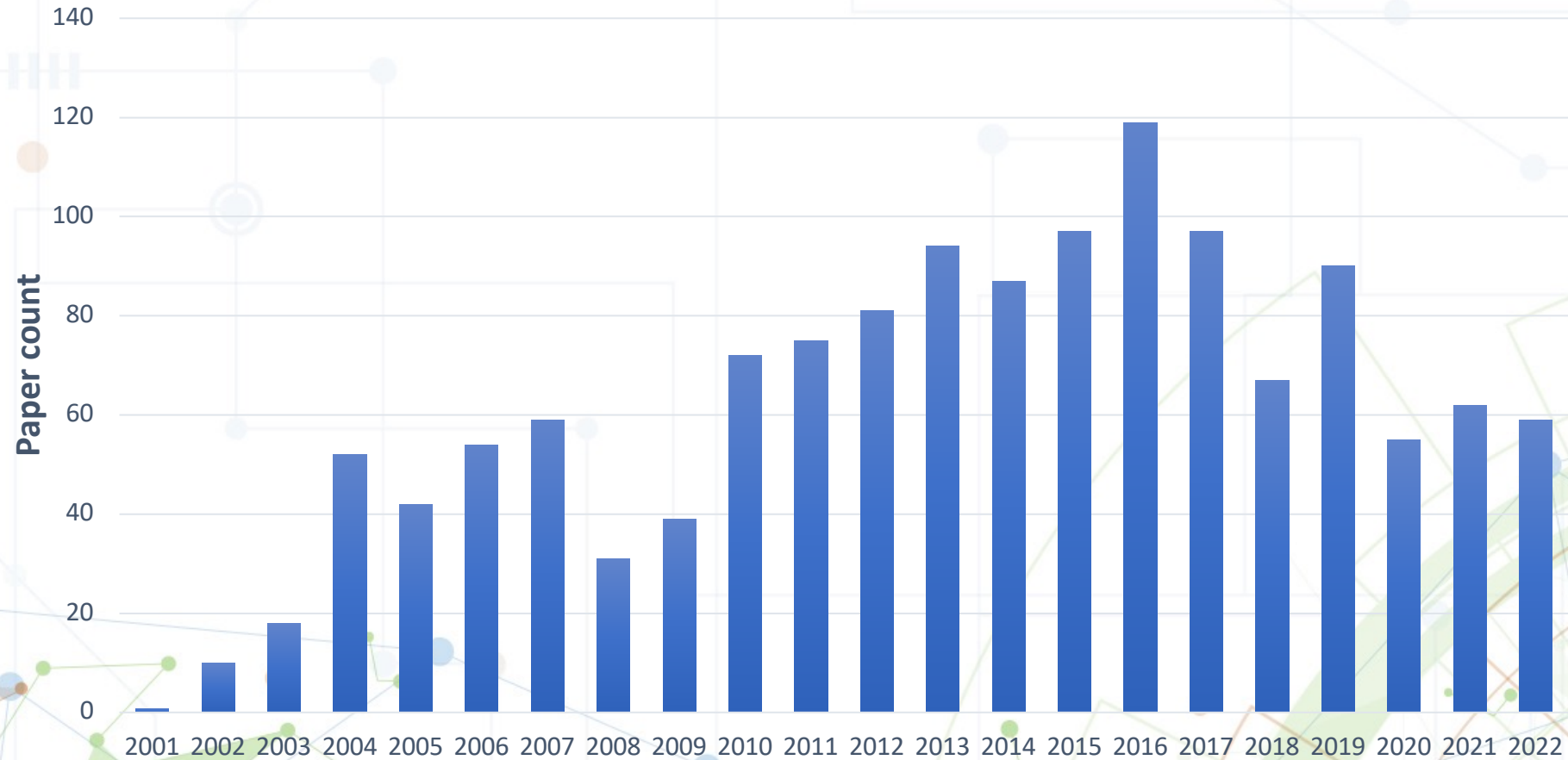
DIGITAL VIDEO  
RETRIEVAL  
at  
NIST



# Bibliography



DIGITAL VIDEO  
RETRIEVAL  
at  
NIST



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

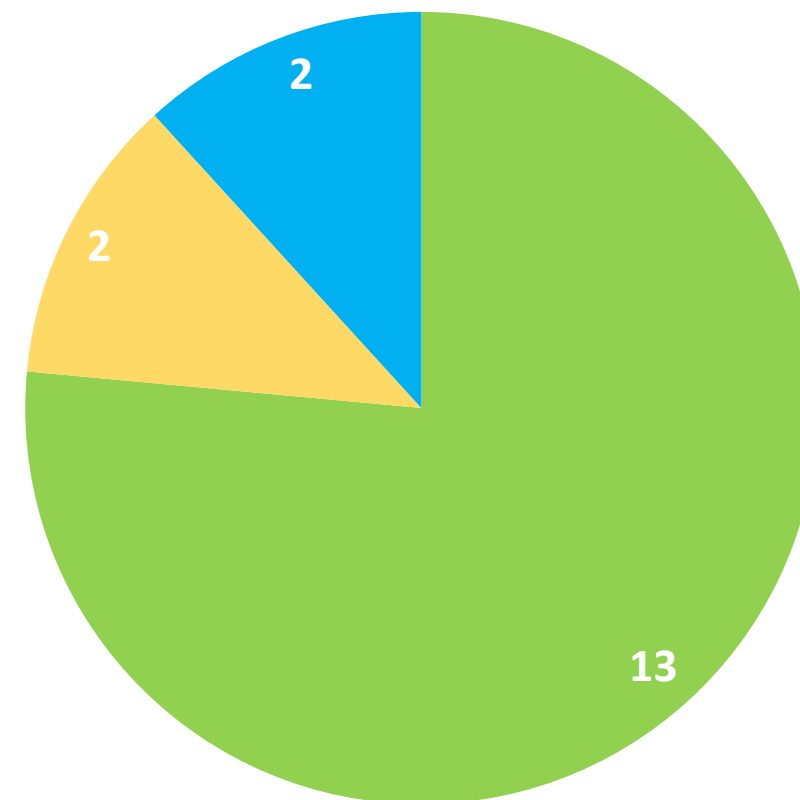
# TRECVID 2023 Tasks & Data

Ad-hoc Video Search (AVS)	Deep Video Understanding (DVU)	Video to Text (VTT)	Medical Video Question Answering (MedVidQA)	Activities in Extended Videos (ACTEV)
<p>Vimeo CC (V3C2)</p> <p>1300 h testing data, ~1.4 M shots, title, keywords, description</p> <p>~3000 h dev data</p> <p>30 textual queries</p> <p>20 progress queries [fixed 2022 – 2024]</p>	<p>KinoLorber</p> <p>19 CC movies (~23 hr) dev set</p> <p>5 licensed movies (7.5 hrs) test set</p> <p>Movie-level queries (relations between people/locations, open domain QA)</p> <p>Scene-level queries (targeting interactions)</p>	<p>Vimeo CC (V3C3)</p> <p>2000 Short Vimeo testing video clips</p> <p>12,870 dev videos and captions from Twitter Vine, Flickr, and V3C</p> <p>Main task:</p> <ul style="list-style-type: none"><li>- Generate automatic video description</li><li>- Progress task (2021-2023)</li></ul>	<p>Medical Instructional Videos</p> <p>Dev-set: 3,010 human-annotated instructional questions + 900 health-related videos</p> <p>Tasks: Visual Answer Localization, Instructional Question Generation</p>	<p>MEVA</p> <p><i>16 h multi-camera video surveillance dataset</i></p> <p><i>Detection and Localization of activities</i></p> <p><i>20 activities evaluated</i></p>

# TRECVID 2023 Finishers (16 / 26)

Groups Finished	Task Code	Task Name
7	AVS	Ad-hoc Video Search
4	ACTEV	Activities in Extended Video
5	VTT	Video to Text Description
2	DVU	Deep Video Understanding
5	MedVidQA	Medical Video Question Answering

Unique finishing teams



■ Asia ■ Europe ■ North America

# TRECVID 2023 Finishers (16 / 26 teams)



Team	Organization
BUPT_MCPRL	Beijing University of Posts and Telecommunications
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
RUC_AIM3	Renmin University of China
RUCMM	Renmin University of China
VIREO	Singapore Management University; City University of Hong Kong
WasedaMeiseiSoftbank	Waseda University; Meisei University ; SoftBank Corporation
UMBCVQA	University of Maryland Baltimore County
MI_TJU	Tianjin University
doshisha_uzl	Doshisha University; Institute of Medical Informatics, University of Lubeck DFKI
VPAI	Hunan University; National Laboratory of Pattern Recognition Institute of Automation, Chinese Academy Sciences
UNCWAI	University of North Carolina Wilmington
FDU_AWS	Fudan University; Amazon Web Service

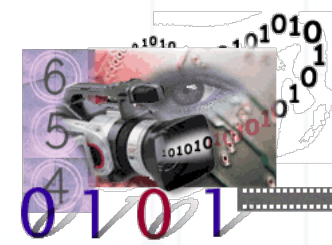


# TRECVID 2023 Non-Finishers (10 / 26)



Team	Organization
CamiloUchile	University of Chile
Umass_BioNLP	Umass Amherst
suvosree	Florida Atlantic University
Delphi	City University of Hong Kong University of Oxford Australian National University
CMU_DVU	Carnegie Mellon University
HSMN	University of Applied Sciences
upvqa	University Politehnica of Bucharest
PKU_WICT	Peking University
TJUMMG	Tianjin University
VRR	Zhongyuan University of Technology

# Support



DIGITAL VIDEO  
RETRIEVAL  
at  
NIST

The running of TRECVID 2023 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

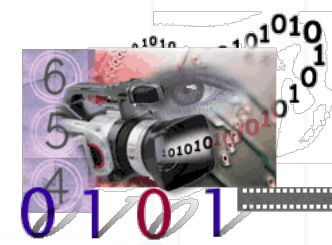
TRECVID is only possible because of the additional efforts of many individuals and groups around the world.



# NIST Assessors

TRECVID 2023

# Resources & Contributions



DIGITAL VIDEO  
RETRIEVAL  
at  
NIST

- 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
- KinoLorberEdu for licensing the movies dataset for the DVU task.
- Deepak Gupta, Dina Demner Fushman, Soumya Gayen, Kush Attal, Rory Mulcahey, Brian Ondov, Gaetano Manzo, Sarvesh Soni, Arun Crispino, NLM Index Section Team – at NIH for supporting the MedVidQA Task

# Agenda at a glance

## 1 Day 1 (Monday)

Ad-hoc Video Search

Deep Video Understanding

## 3 Day 3 (Wednesday)

Medical Video Question Answering

Conclusion & TRECVID 2024

OpenMFC

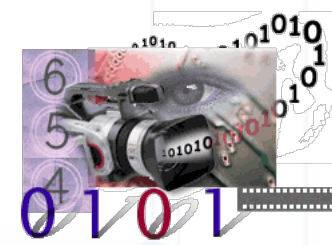
## 2 Day 2 (Tuesday)

Video to Text

Activities in Extended Videos

OpenMFC

# Miscellany



DIGITAL VIDEO  
RETRIEVAL  
at  
NIST

- OpenMFC Sessions running after TRECVID (Tues – Wed)
- Slack
  - 2023 TrecVid Track Channels
  - Agenda/Schedule Channel
  - Random
- Online workshop proceedings:
  - Teams papers, task information, datasets, ground truth, etc) can be found at:
  - <https://www-nlpir.nist.gov/projects/tv2023/workshop/tv23.workshop.notebook/>