Multimedia Event Detection and Recounting Task Reports

Time		Presentation
9:00	– 9:20	MED Evaluation Overview (NIST)
9:20	– 9:40	Centre for Research and Technology Hellas (ITI-CERTH)
9:40	– 10:00	Raytheon BBN Technologies (BBN)
10:00	– 10:20	University of Amsterdam (MediaMill)
10:20	– 10:40	**Break**
10:40	– 11:00	City University of Hong Kong (VIREO)
11:00	– 11:20	Carnegie Mellon University (CMU)
11:20	– 11:40	MER Evaluation Overview (NIST)
11:40	– 12:40	MER Panel Discussion (all teams participating in MER)
12:40	– 2:10	**Lunch** on your own (remember closures due to Veteran's Day)
2:10	– 2:40	MED/MER Discussion

2014 TRECVID Workshop Multimedia Event Detection Task

Jonathan Fiscus*, Greg Sanders*, David Joy*, Paul Over*, Vladimir Dreyvitser+

* National Institute of Standards and Technology (NIST)
+ Systems Plus, Inc.

DIGITAL VIDEO RETRIEVAL at NIST

Talk Outline

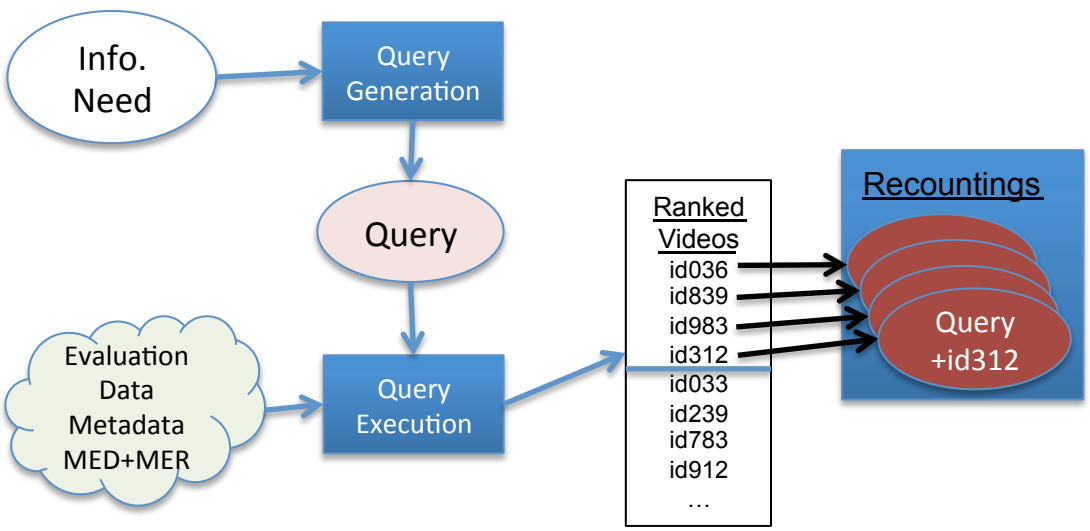
- MED/MER Task Overview
- HAVIC Resources
- MED Results

Multimedia Event Detection (MED)/ Multimedia Event Recounting (MER) Task Definition

Quickly find instances of events in a large collection of search videos and ➔ MED

recount the evidence in a search video for the event. ➔ MER

The query's central role



Extensive System Outputs and Execution Measurements

Detection Outputs:

- Per Event/Video:
  - Rank wrt. the event
  - Confidence score wrt. the event
- Per Event
  - Thresholds for ranks and confidence scores optimizing prescribed metrics

Execution Speed Measurements:

- Per Event
  - Query generation times
  - Event Search execution times
- Per Video Collection
  - Metadata generation

Execution Hardware Reporting

- Per Module
  - Hardware for each module call

Rock Climbing Event Kit Text

Definition:

One or more people climb up or across rock formations or artificial rock walls.

Explication:

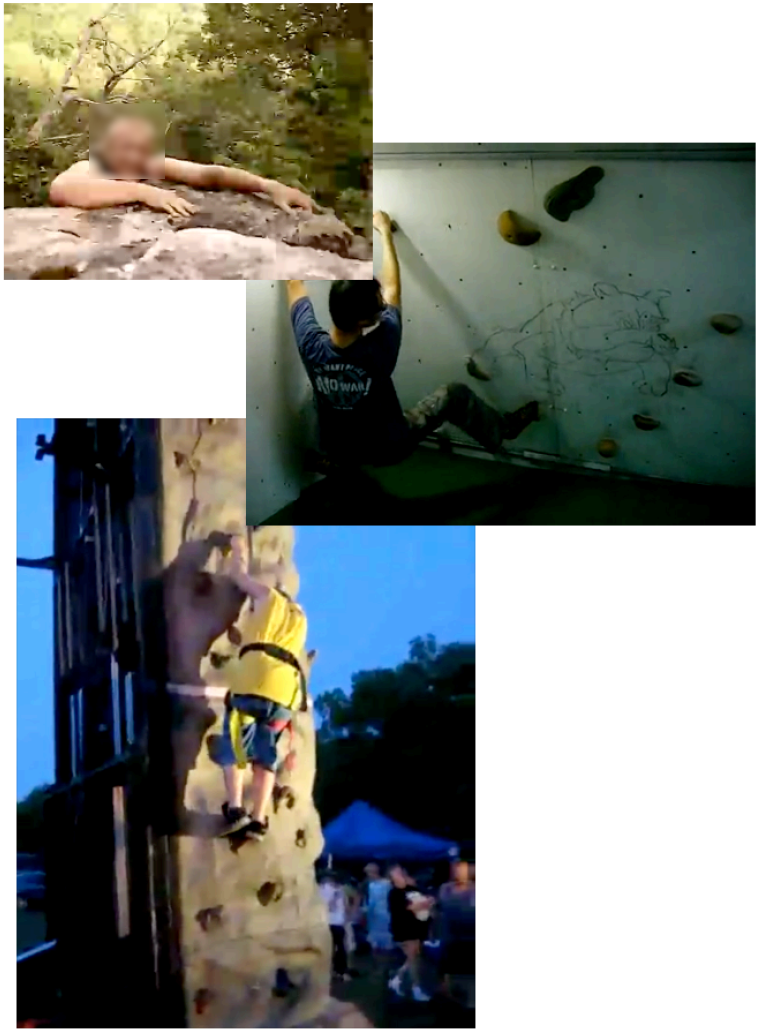
Rock climbing is a physically intense activity, where the goal is to reach the top or endpoint of a pre-defined route on a rock formation or artificial rock wall by finding a grip on the surface using hands and feet, and then pulling up using their arm and leg strength. …

Evidential Description:

- scene: outdoors in natural setting, indoors in rock climbing gym, or outdoors on a specially …
- objects/people: carabiners, rope, helmet, harness, rock formation, artificial rock wall, climbers
- activities: hooking rope to harness, moving hands and feet along side of rock face, grabbing rock …….
- audio: carabiners clinking, climbers making comments on the difficulty of the climb, onlookers cheering on …

Illustrative Examples

- Positive instances of the event
- Non-Positive "miss" clips that do not contain the event



MED Evaluation Conditions

- MED Tasks
  - **Pre-Specified Event (PS)** – MED metadata generation optimized with knowledge of events not using Pseudo Relevance Feedback
  - **Ad-Hoc Event (AH)** – MED metadata generation complete before events revealed not using Pseudo Relevance Feedback
- Event Query Conditions
  - **SQ** – (Semantic Query) Use the event kit texts
  - **0Ex** – Event text **PLUS** 5,000 clip Event Background (Event-BG) collection
  - **10Ex** – Event text, Event BG, **PLUS** 10 positive and 10 miss clips per event
  - **100Ex** – Event text, Event BG, **PLUS** 100 positive and 50 miss clips per event
- Optional System Strategies
  - **PRF** - Pseudo Relevance Feedback
- Search video collection (pick one)
  - **MED14Eval-Full ->** 198,000 videos, 7,580 hours
    - PROGRESS Collection = 98,000 videos, 3,722 hours
    - Novel 1 Collection = 100,000 videos, 3,858 hours
  - **MED14Eval-Sub** -> 32,000 video subset of MED14Eval-Full, 1,238 hours

The TRECVID MED 2014 Events

Pre-Specified Events

MED '13 PS Events	*MED '13 AH Events*
Attempting a bike trick	Beekeeping
Cleaning an appliance	Wedding shower
Dog show	Non-motorized veh. repair
Giving directions to a location	Fixing musical instrument
Marriage proposal	Horse riding competition
Renovating a home	Felling a tree
Rock climbing	Parking a vehicle
Town hall meeting	Playing fetch
Winning a race without a vehicle	Tailgating
Working on a metal crafts project	Tuning musical instrument

Ad Hoc Events

New Events

Baby Shower
Building a fire
Busking
Decorating for a celebration
Extinguishing a fire
Making a purchase
Modeling
Doing a magic trick
Putting on apparel
Teaching dance choreography

17 MED 2014 Finishers and Number of Runs

Years	Team	No PRF								PRF					Organization
		Ad-Hoc				Pre-Specified				Ad-Hoc			Pre-Specified		
		SQ	100Ex	10Ex	0Ex	SQ	100Ex	10Ex	0Ex	SQ	10Ex	0Ex	10Ex	0Ex	
4	BBNVISER	✓	✓	✓	✓	✓	✓	✓	✓						Raytheon BBN Technologies, UMD, Columbia, UCF Team
4	CERTH-ITI		✓	✓			✓	✓							Informatics and Telematics Inst., Centre for Research and Tech.
4	CMU	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	Carnegie Mellon University
4	MediaMill	✓	✓	✓		✓	✓	✓							University of Amsterdam
4	NII		✓	✓			✓	✓							National Institute of Informatics
4	Sesame	✓	✓	✓		✓	✓	✓							SRI International SESAME
4	SRIAURORA	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓			SRI International Sarnoff Aurora
4	TokyoTechCanon		✓	✓			✓	✓							Tokyo Institute of Technology and Canon
3	AXES		✓	✓			✓	✓							Dublin City Univ, Univ Twente, Univ of Oxford, INRIA, Fraunhofer IAIS, Katholieke Universitiet Leuven, Technicolor, Erasmus University Rotterdam, Cassidian, BBC, Deutsche Welle, Beeld et. Geluid (NISV), ERCIM
3	VIREO-TNO	✓		✓	✓	✓	✓	✓	✓						City University of Hong Kong, TNO
2	ORAND		✓	✓			✓	✓							ORAND S.A. Chile
1	ATTLabs		✓	✓			✓	✓							AT&T
1	Fudan		✓	✓			✓	✓							Fudan University
1	INRIA		✓	✓			✓	✓							INRIA-LEAR-LIMSI-Vocapia
1	KU-ISPL		✓	✓			✓	✓							Korea University
1	MCIS			✓				✓							Beijing Institute of Technology
1	MIC		✓	✓	✓		✓	✓	✓						Tongji University
		6	15	17	5	6	15	17	5	1	1	2	1	1	

HAVIC Data Resources

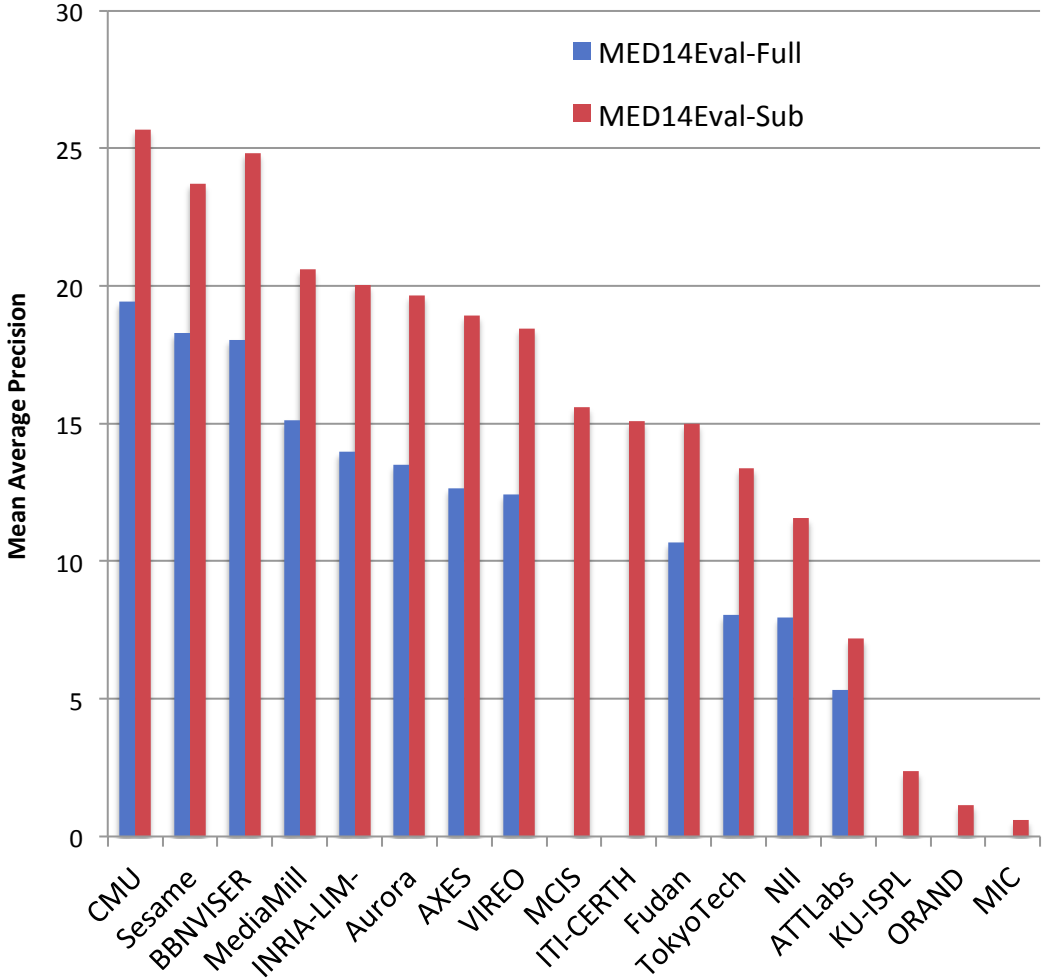
		Video clips	Video duration	2014 Change
Development Data	RESEARCH	10,000	314 hours	-
	10 Event Kits	1,400	74 hours	-
	Transcription	1,500	45 hours	-
Event Training Data	Event Background	5,000	146 hours	-
	40 Event Kits	6,000	270 hours	+53 hours
Test Data	MEDTest	27,000	849 hours	+11 hours
	KindredTest	14,500	687 hours	+11 hours
Evaluation Data	MED14Eval-Full	198,000	7,580 hours	+ 3,858 hours
	MED14Eval-Sub	33,000	1,244 hours	- 7 hours
	Total	244,000	9,911 hours	+3,922 hours

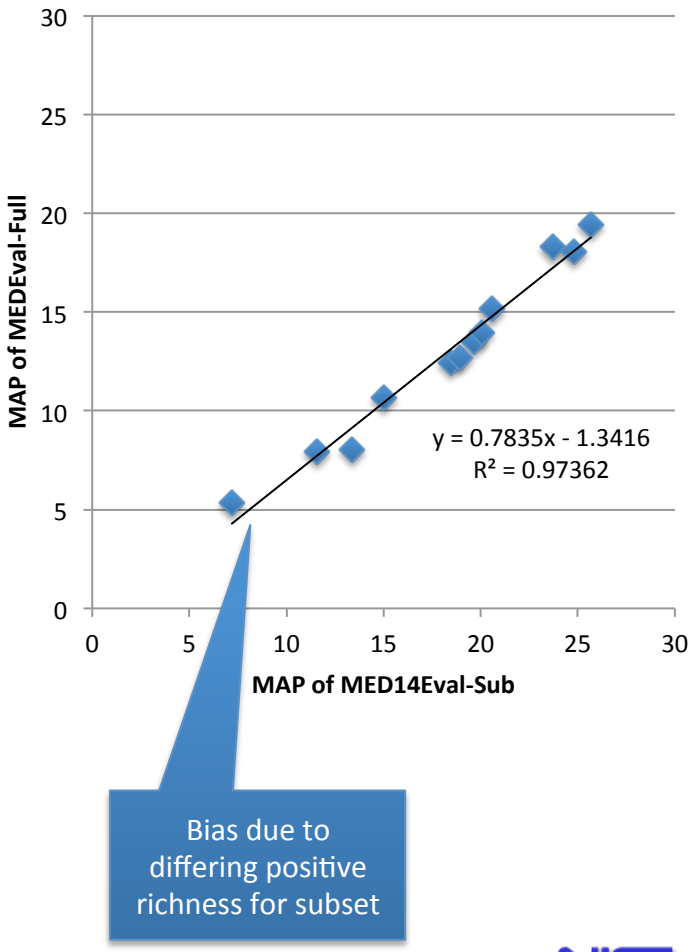
MED '14 Results

- Accuracy metrics
  - Mean Average Precision (MAP) - (Primary)
  - Minimal Acceptable Recall (R_0) - (Diagnostic)
- Computation speed and resource measurements
  - Hardware analysis
  - Metadata generation
  - Event query generation
  - Event search

MED '14 Pre-Specified Event Task

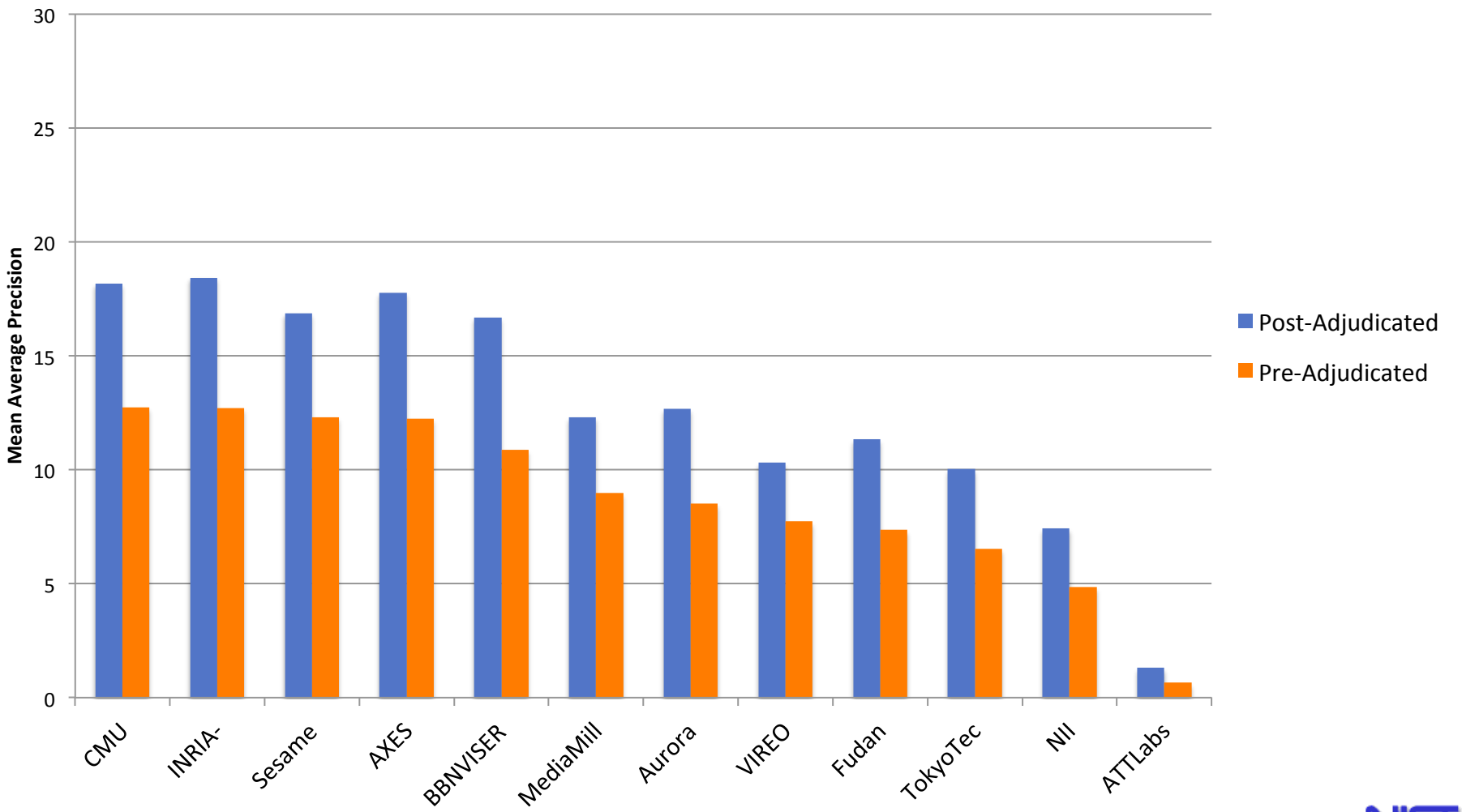
Events E0[21-40], 10Ex, NoPRF, All Teams





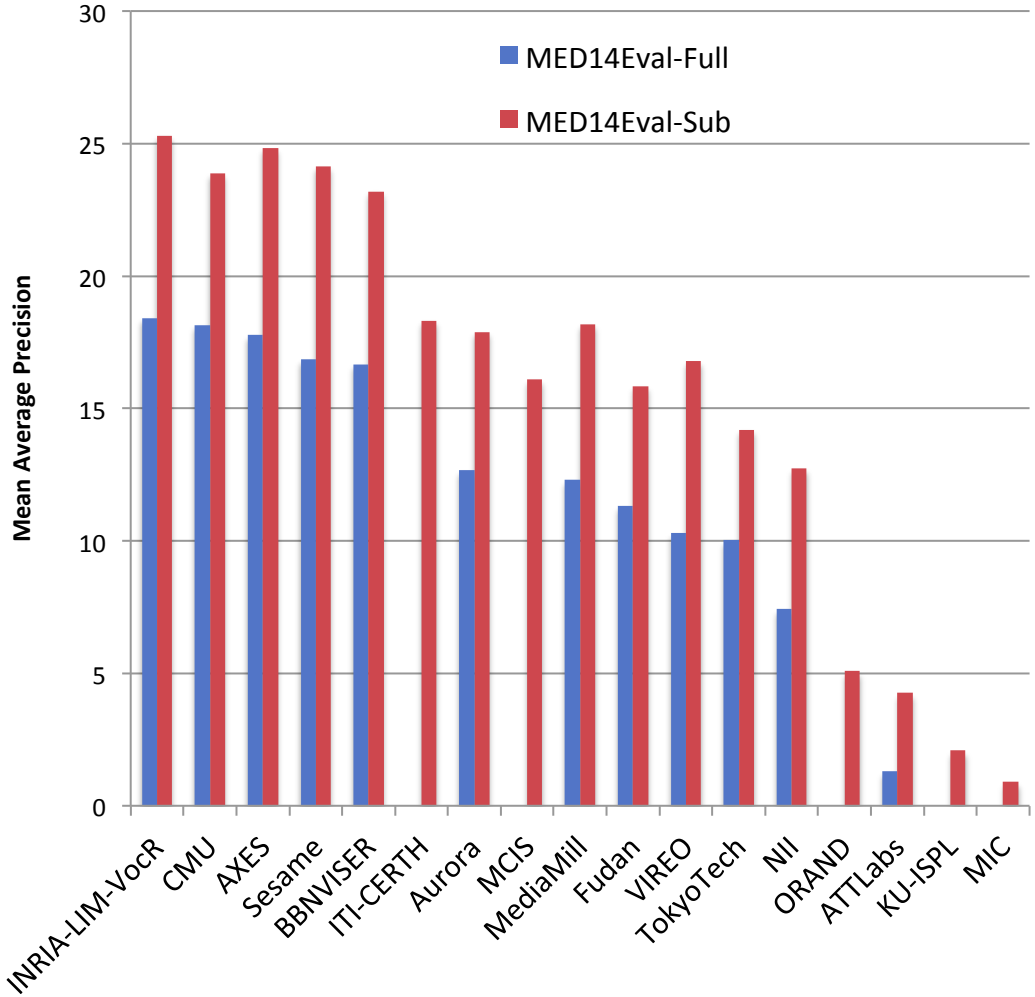
Ad Hoc Pre- and Post-Adjudicated MED '14

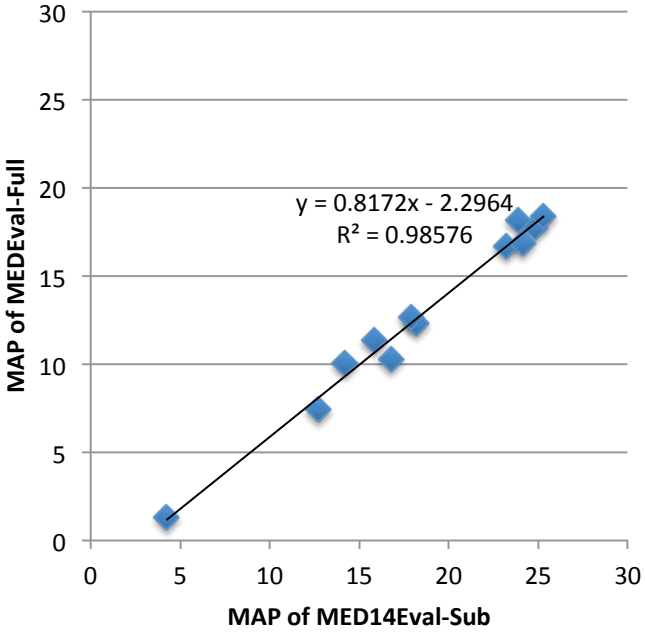
Events E041-E050, 10Ex, NoPRF, MED14Eval-Full Teams



Post-Adjudicated MED '14 Ad-Hoc Event Task

Events E0[41-50], 10Ex, NoPRF, All Teams





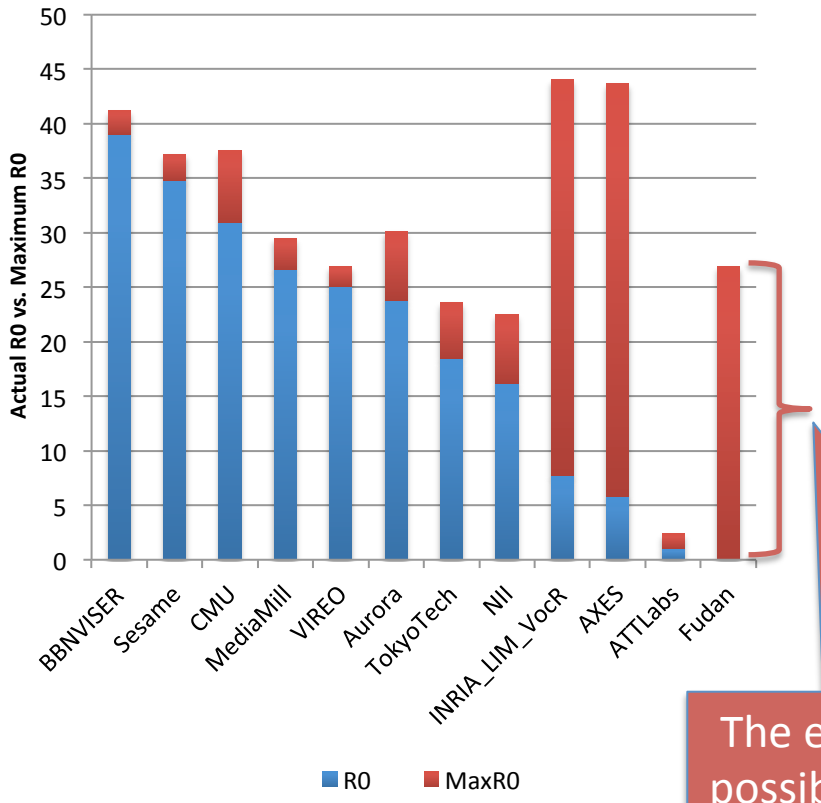
Statistical Comparisons of AH and PS

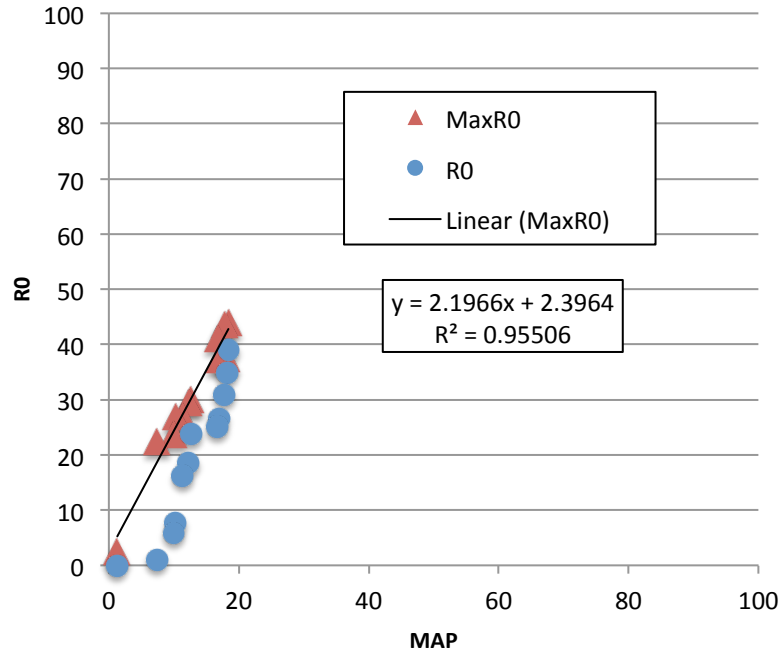
TBD

MAP Vs. Ro

Post-Adjudicated MED ’14 Ad-Hoc Event Task

Events E0[41-50], 10Ex, NoPRF, All Teams

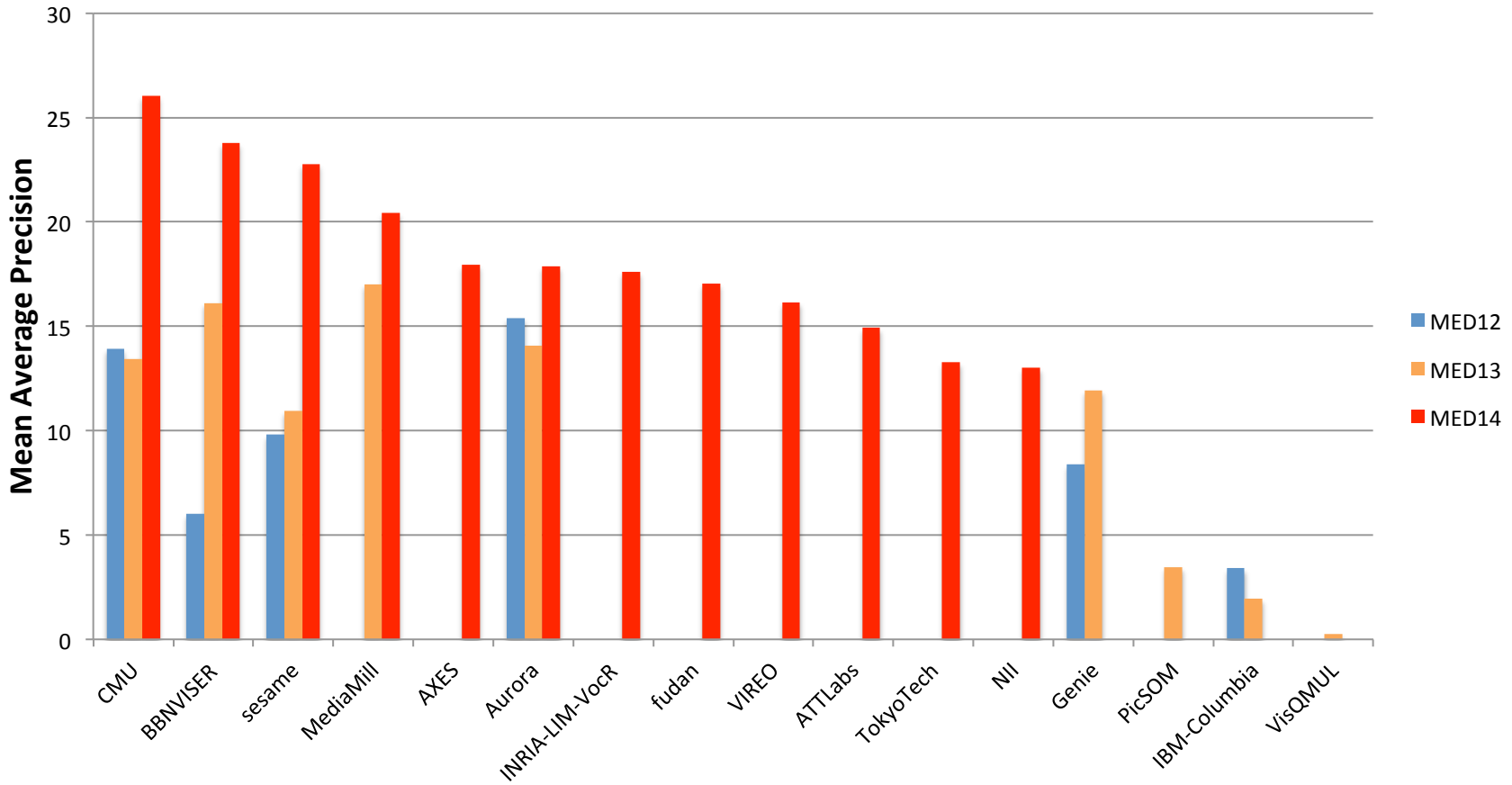




The extent of the red bar indicates the possible improvement with an accurate threshold

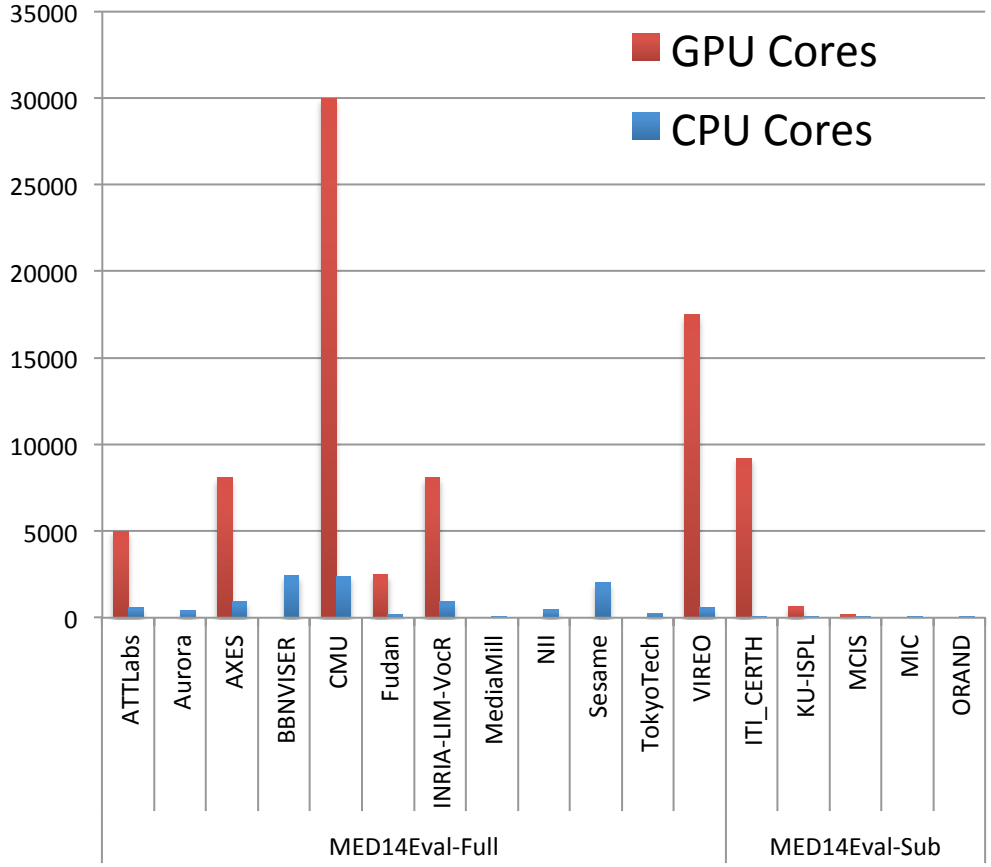
Historical MED MAP Scores

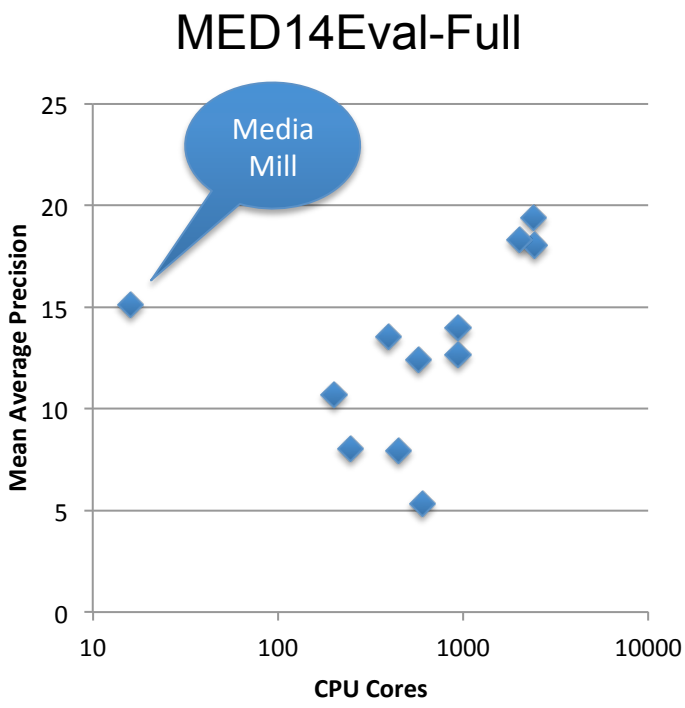
PROGTEST, MED ’12, ’13, and ’14, 10Ex



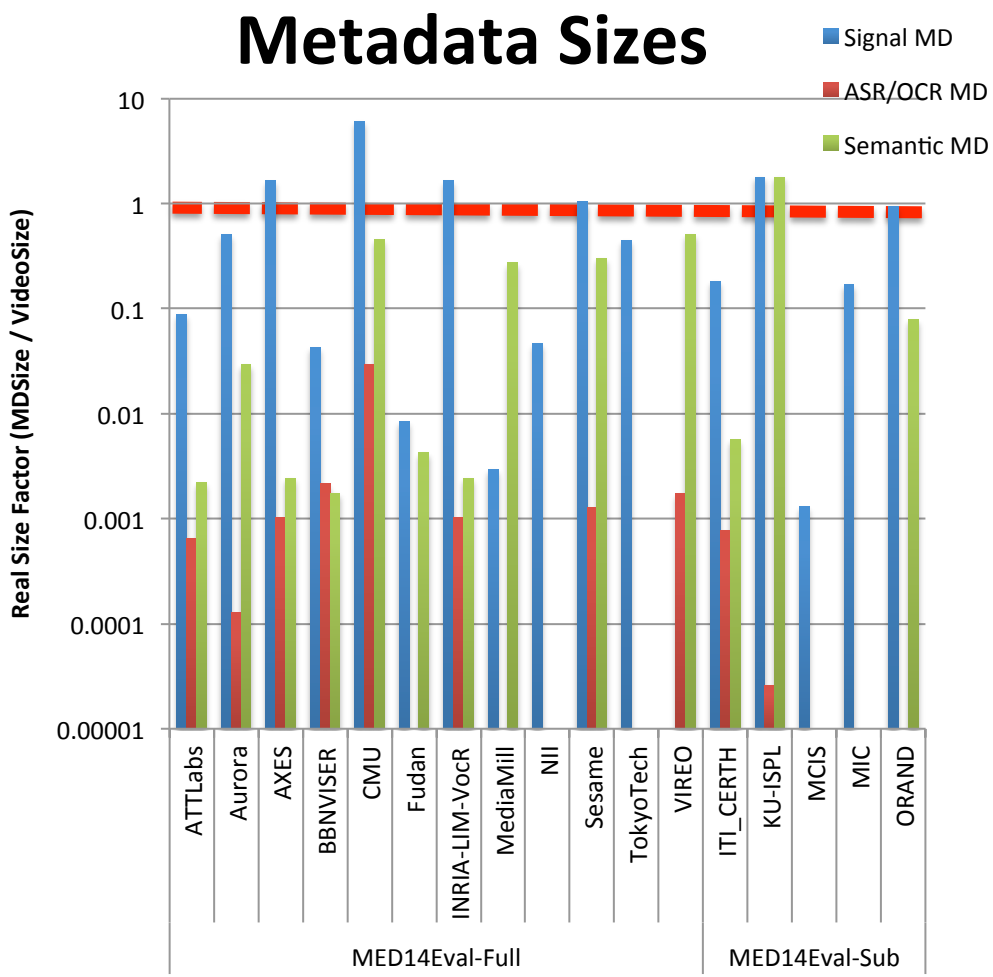
Hardware Summary

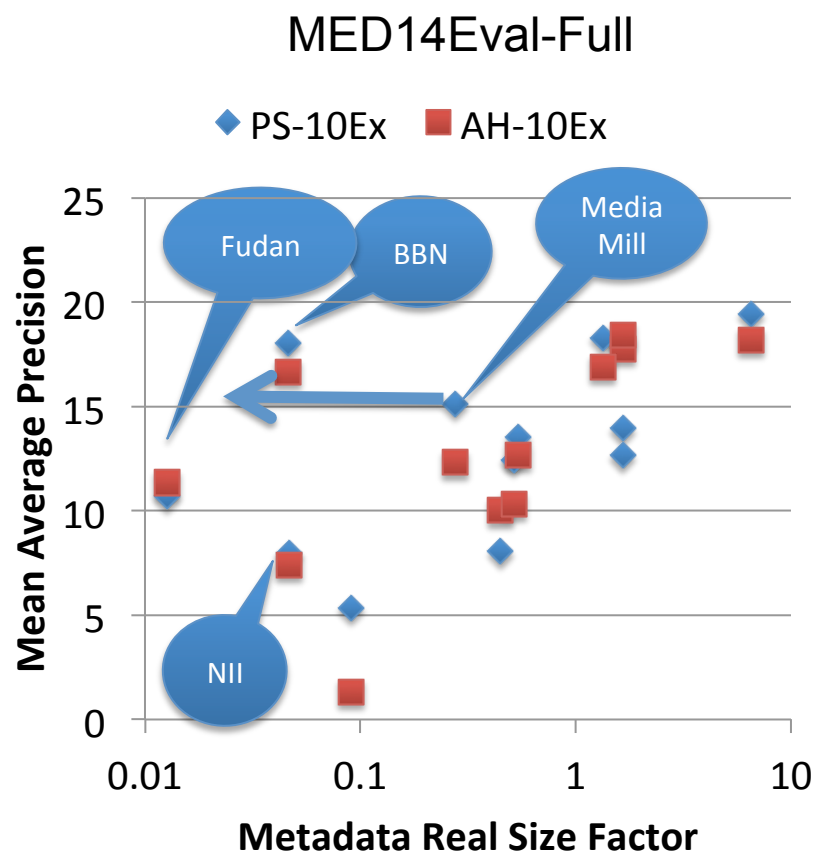
Test Collection Metadata Generation





Metadata Generation

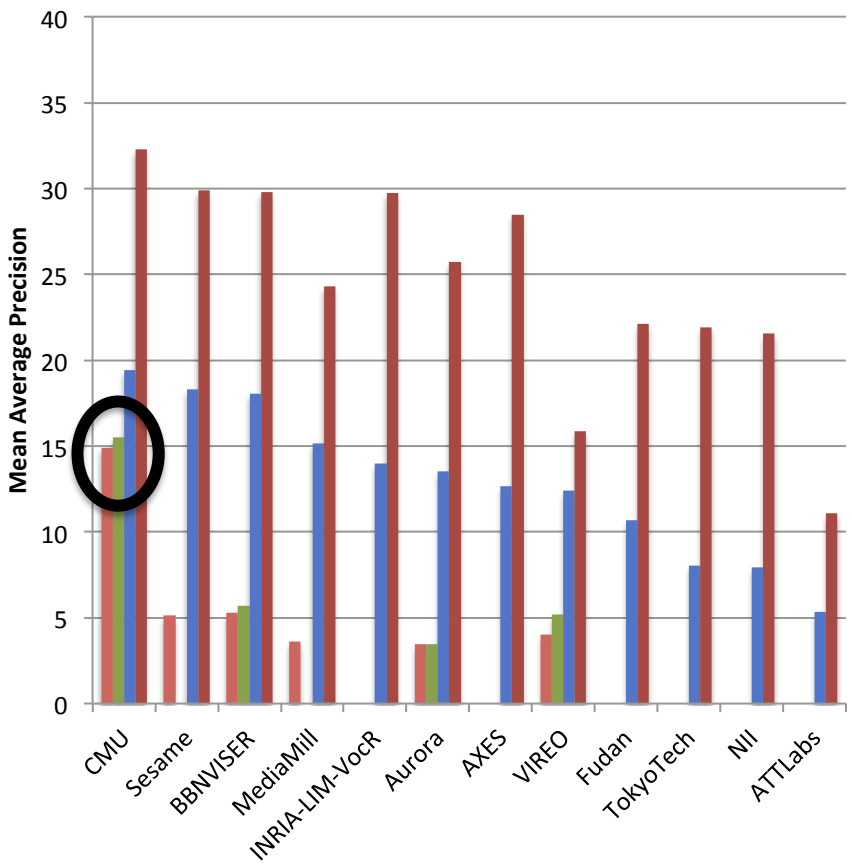




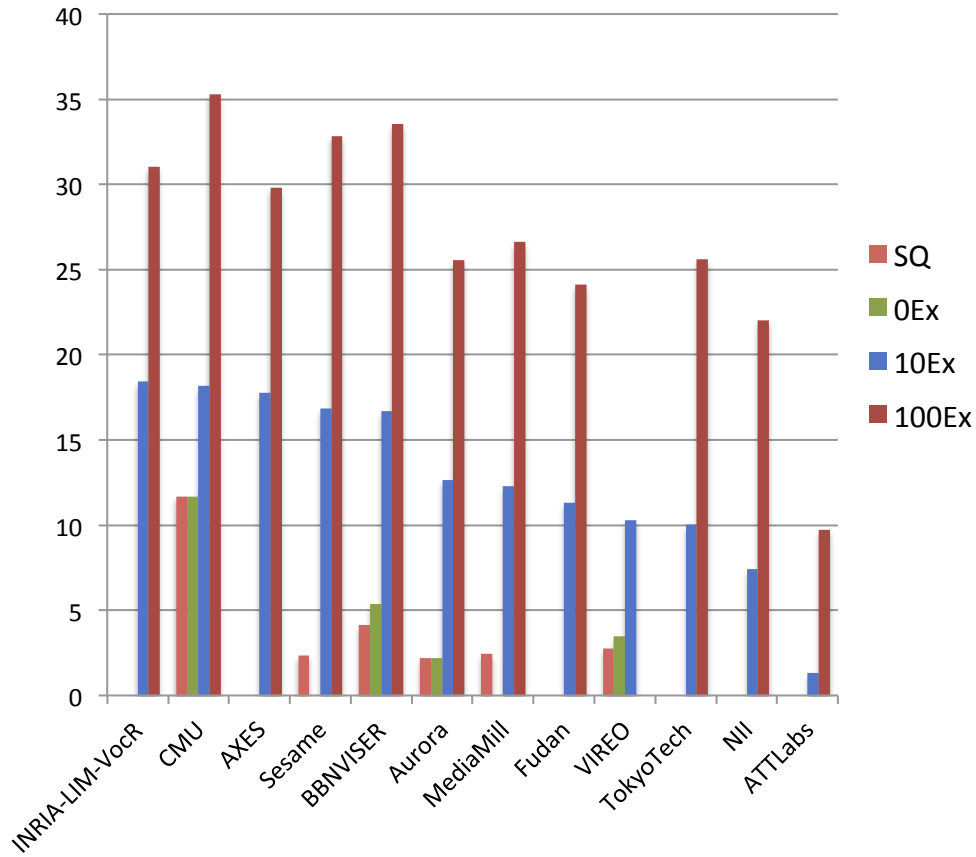
Effect of Query Type

MED14Eval-Full, SQ, 0Ex, 10Ex, 100Ex

Pre-Specified Events



Post-Adjudicated Ad-Hoc Events



Semantic Query Edit Pilot

- Purpose
  - Explore the concept of editable MED semantic queries.
- Pilot Question?
  - Can an un-trained user remove query content to search for a more narrow event?
    - E.g. board trick -> snow board trick
- Procedure
  - Create two new events and reference annotations
  - Group edit queries removing tags to match the new event
  - Teams perform event search on modified queries
  - Provide feedback to teams

Number of Events with Indicated Change in MAP

	Aurora	BBNViser	CMU	VIREO	Sesame
Does the edit reduce performance on original event?	1	1	2	2	1
Does the edit improve performance on the new event?	1	0	2	0	0

Caveat Emptor
No "changes" are significant!
Abs(MAP delta) = 0.5%

Summary/Conclusions

- Results observations
  - Large MAP discrepancy between MED14Eval-Full and MED14Eval-Sub
  - Large change in performance for AH Adjudication
  - The Semantic Edit Query pilot results are not encouraging enough to continue
- Evaluation infrastructure observations
  - I/O server sequencing of modules to automate system execution provided extensive system measurements
- Changes changes for MED 2015
  - A new evaluation collect for MED 2015
  - MED 14 AdHoc events will become PreSpecified events next year
  - Expect new AdHoc Events
  - Changes
    - Adjudication-based reference generation is a viable method for new events
    - Resource control: time, hardware, etc.

Questions?