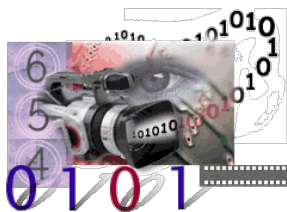


2015 TRECVID Workshop

Multimedia Event Detection Task

Jonathan Fiscus*, **David Joy***, **Martial Michel***
Yooyoung Lee+

* National Institute of Standards and Technology (NIST)
+ Dakota Consulting, Inc.



DIGITAL VIDEO
RETRIEVAL
at
NIST

Talk Outline

- MED Task Overview
- HAVIC Resources
- MED Results
- Future Plans

MED '15 Overview

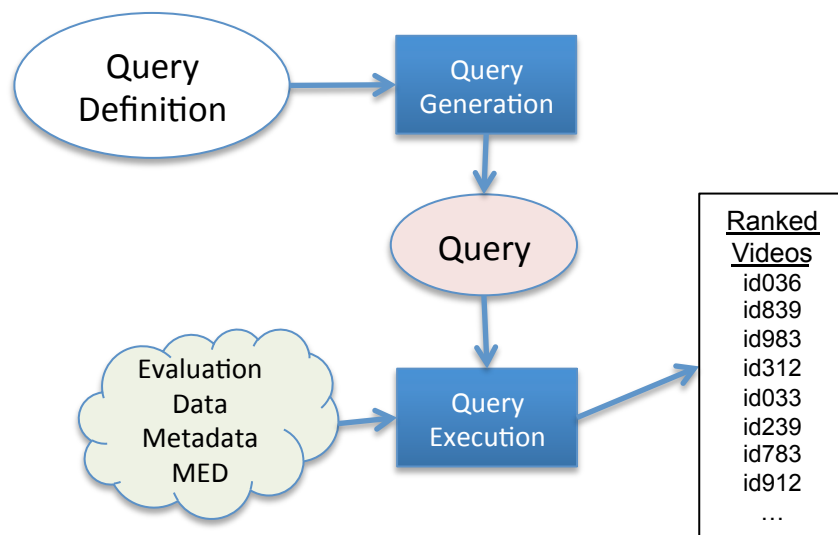
- MED evaluations from 2010-2015
 - Supported by the IARPA Aladdin Program and LDC Collected Data
 - NIST intends to continue support but with reduced costs taking into account lessons learned and addressing un-met challenges
 - Constructed data sets with exhaustive annotation
 - Variable event richness
- 2015 MED task simplifications
 - Multimedia Event Recounting task discontinued
 - 100 Exemplar Ad-Hoc Event training discontinued
 - Hardware/runtime reporting simplified
 - Primary metric changed to Inferred Average Precision with pooled assessment for reference generation.

Multimedia Event Detection Task

Multimedia Event Detection (MED)

Quickly find instances of events in a large collection of search videos

The query's central role



Evaluation Conditions

Execution Hardware Reporting

- 3 Classes of Computing Hardware
- Small: 100 CPU cores, 1,000 GPU cores
- Medium: 1,000 CPU cores, 10,000 GPU cores
- Large: 3,000 CPU cores, 30,000 GPU cores

Query Training Conditions

	Number of Exemplars		
Pre-Specified Events	0	10	100
Ad-Hoc Events		10	
Interactive Ad-Hoc Events		10	

Search Collection

- MED15Eval-Full -> 198K videos, 7,580 hours
- MED15Eval-Sub -> 32K video subset, 1,238 hours

Land Vehicle Accident Event

Definition:

A motorized land vehicle being operated by a human hits or crashes into another vehicle or object.

Explication:

The vehicle initiating the accident must be a motorized land vehicle (e.g. non-motorized bicycles, scooters, etc. are excluded from this event). The vehicles motion must be directed at least in part by a human operator to be considered a positive. Simple loss of control

Evidential Description:

- scene: outdoors, typically on roads or highways and often proximal to other vehicles and/or man-made ...
- objects/people: motorized land vehicle, car, truck, operator, onlookers, man-made structure or barrier
- activities: driving, loss of vehicular control, people observing accident, people and or other vehicles ...
- audio: vehicle impacting other vehicle or object, engine noise, tires screeching, honking, people yelling ...

Illustrative Examples

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



The TRECVID MED 2015 Events

Pre-Specified Events

MED '14 PS Events

Attempting a bike trick
Cleaning an appliance
Dog show
Giving directions to a location
Marriage proposal
Renovating a home
Rock climbing
Town hall meeting
Winning a race without a vehicle
Working on a metal crafts project

MED '14 AH Events

Beekeeping
Wedding shower
Non-motorized veh. repair
Fixing musical instrument
Horse riding competition
Felling a tree
Parking a vehicle
Playing fetch
Tailgating
Tuning musical instrument

Ad-Hoc Events

New Events

Gardeners harvest food
Land vehicle accident
Person jumps into nat. water
Cooking on an outdoor grill
Moving through a flooded st.
Skyscraper window cleaning
Firefighters battle a fire
Climbing a tree
Lecture to an audience
Team scores a touchdown

16 MED 2015 Finishers By Condition

Years	Team	AH		PS						Organization
		10Ex		0Ex		10Ex		100Ex		
		SML	MED	SML	MED	SML	MED	SML	MED	
5	CMU				Full		Full			Carnegie Mellon Univ.
	MediaMill	Full		Full		Full		Full		MediaMill - University of Amsterdam and Qualcomm
	NIIHitachiUIT	Full		Full		Full		Full		National Institute of Informatics
	TokyoTech	Full				Full		Full		Tokyo Institute of Technology
4	VIREO_TNO		Full		Full		Full		Full	City University of Hong Kong & TNO
2	ITICERTH			Full		Sub		Sub		Informatics and Telematics Inst.
	KoreaUnivISPL					Sub		Sub		Korea University
	NTTFudan		Full				Full			NTT Media Intelligence Laboratories and Fudan University
	MCIS					Sub		Sub		Beijing Institute of Technology Mcislab
	UEC					Sub				University of Electro-Communications
1	BCMI-SJTU					Sub		Sub		Center for Brain-like Computing and Machine Intelligence SJTU
	BUPTMCPR					Sub		Sub		Multimedia Communication and Pattern Recognition Labs BUPT
	Etter			Sub		Sub				EtterSolutions
	Fudan		Full				Full			Fudan University
	NEUMITLL					Full				NEU SMILE and MIT Lincoln Labs
	SiegenKobeNict					Sub		Sub		University of Siegen
		3	3	4	2	12	4	9	1	



HAVIC Data Resources

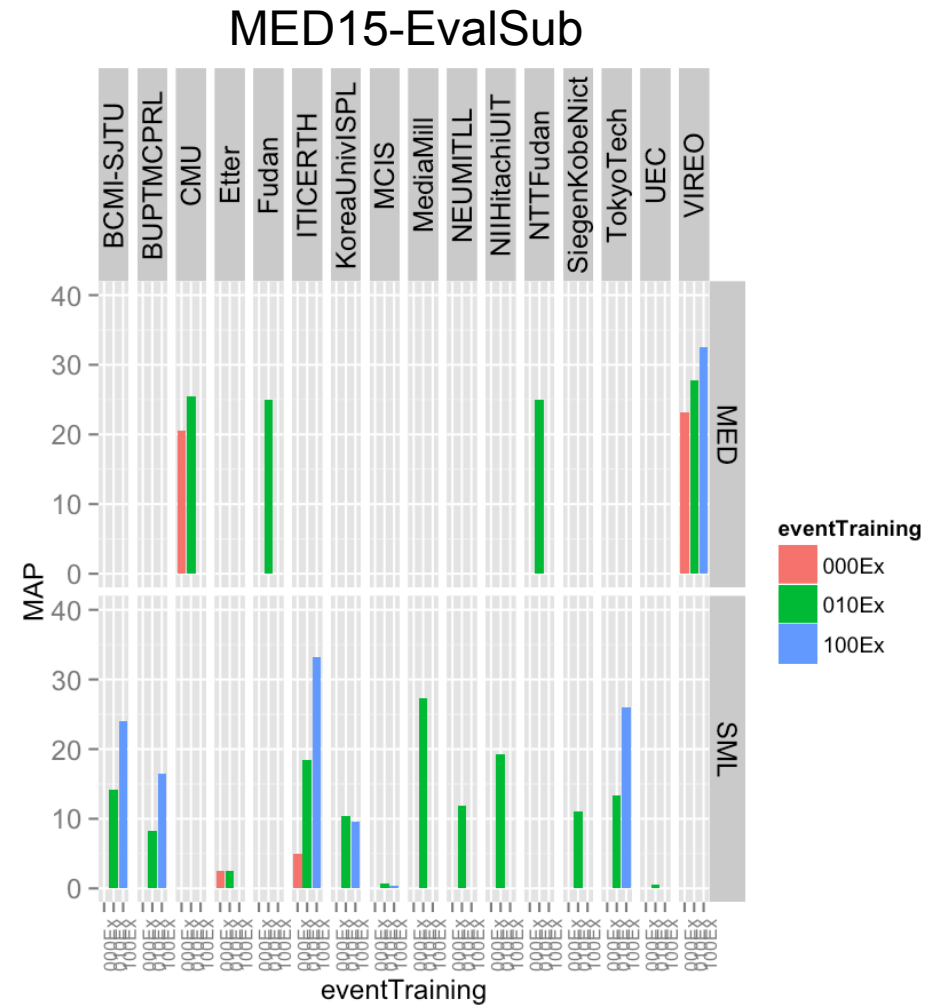
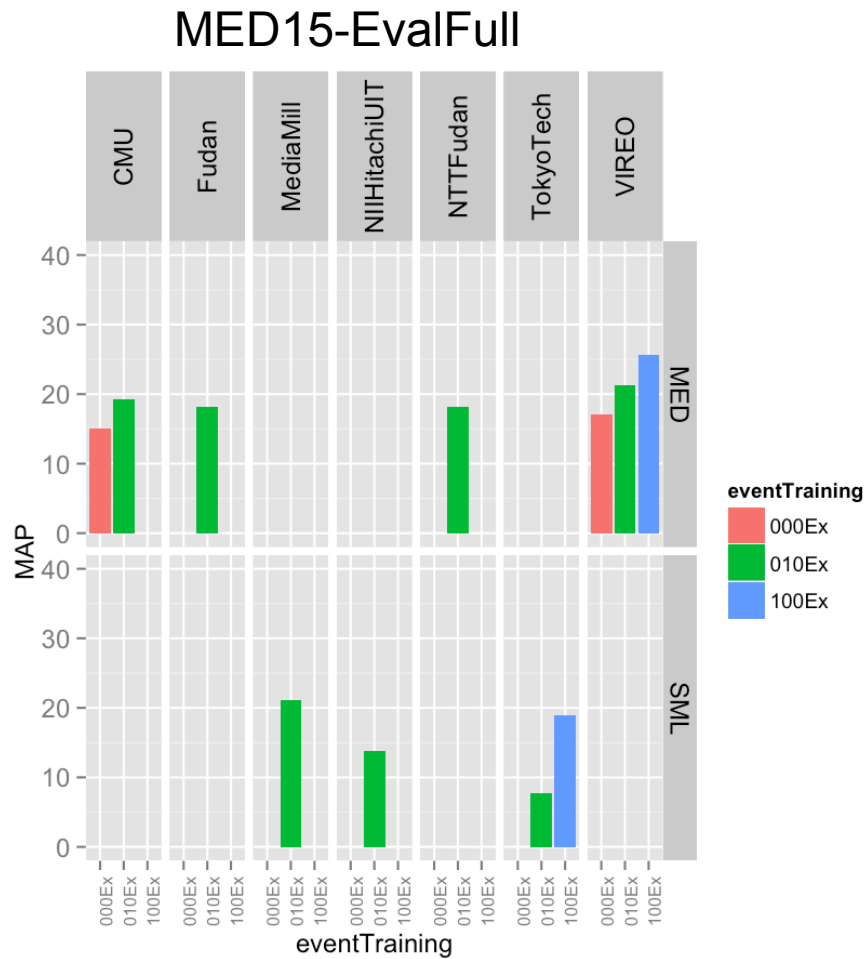
		Video clips	Video duration
Development Data	RESEARCH	10,000	314 hours
	10 Event Kits	1,400	74 hours
	Transcription	1,500	45 hours
Event Training Data	Event Background	5,000	146 hours
	40 Event Kits	6,000	270 hours
Test Data	MEDTest	27,000	849 hours
	KindredTest	14,500	687 hours
Evaluation Data	MED14Eval-Full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

New for 2015:
LIMSI provided
Speech
transcripts

MED '15 Results

- Pre-Specified Event Results
 - Mean Average Precision (MAP)
 - Search set variations Full vs. Sub, Event-Specific AP
 - Inferred Mean Average Precision (InfMAP)
 - Correlation MAP vs. InfMAP
- Ad-Hoc Event Results
 - Inferred Average Precision

Pre-Specified Event Primary Systems

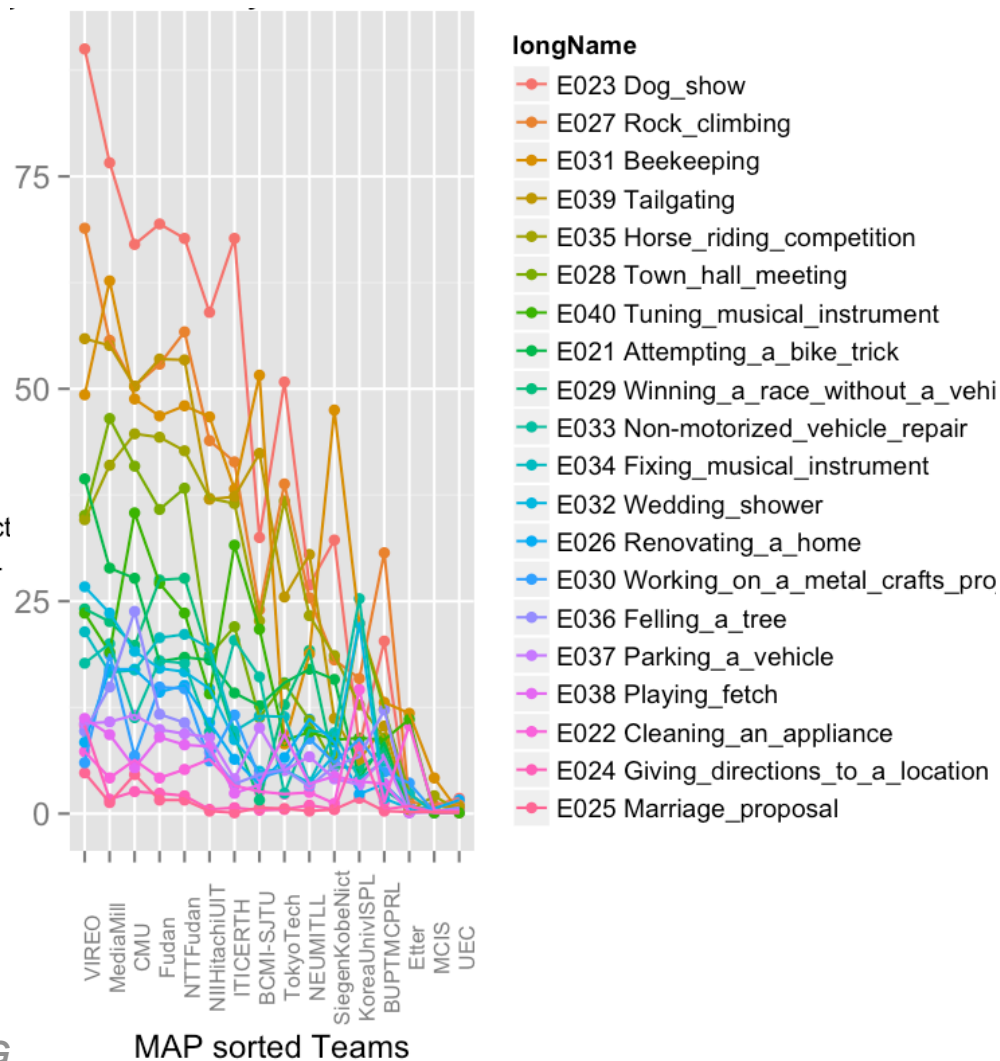
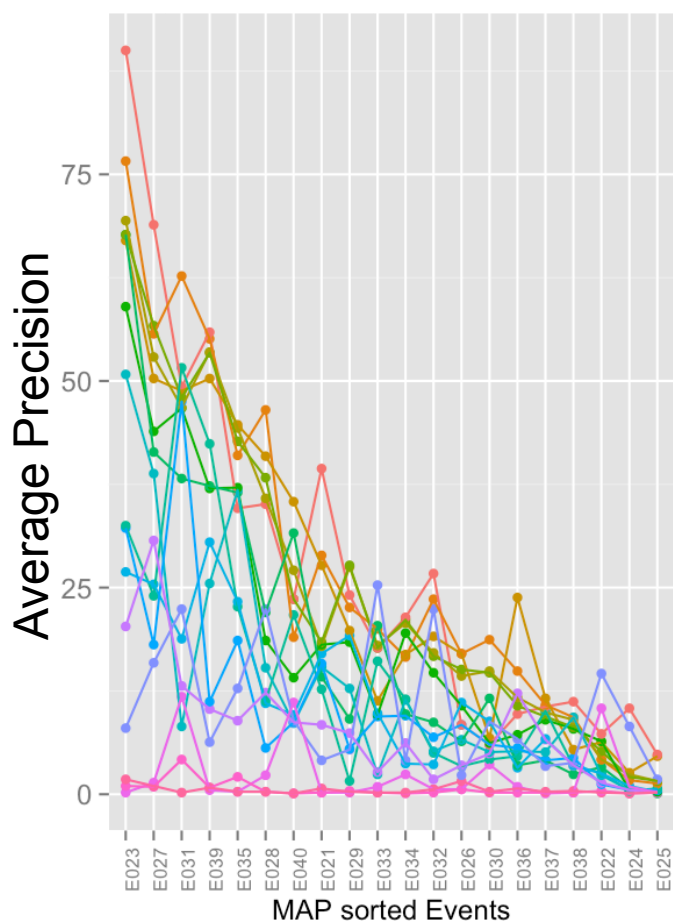


$$\text{MAP}(\text{EvalSub}) = 1.09 * \text{MAP}(\text{EvalFull}) + 4.74$$

$$R^2 = 0.993$$

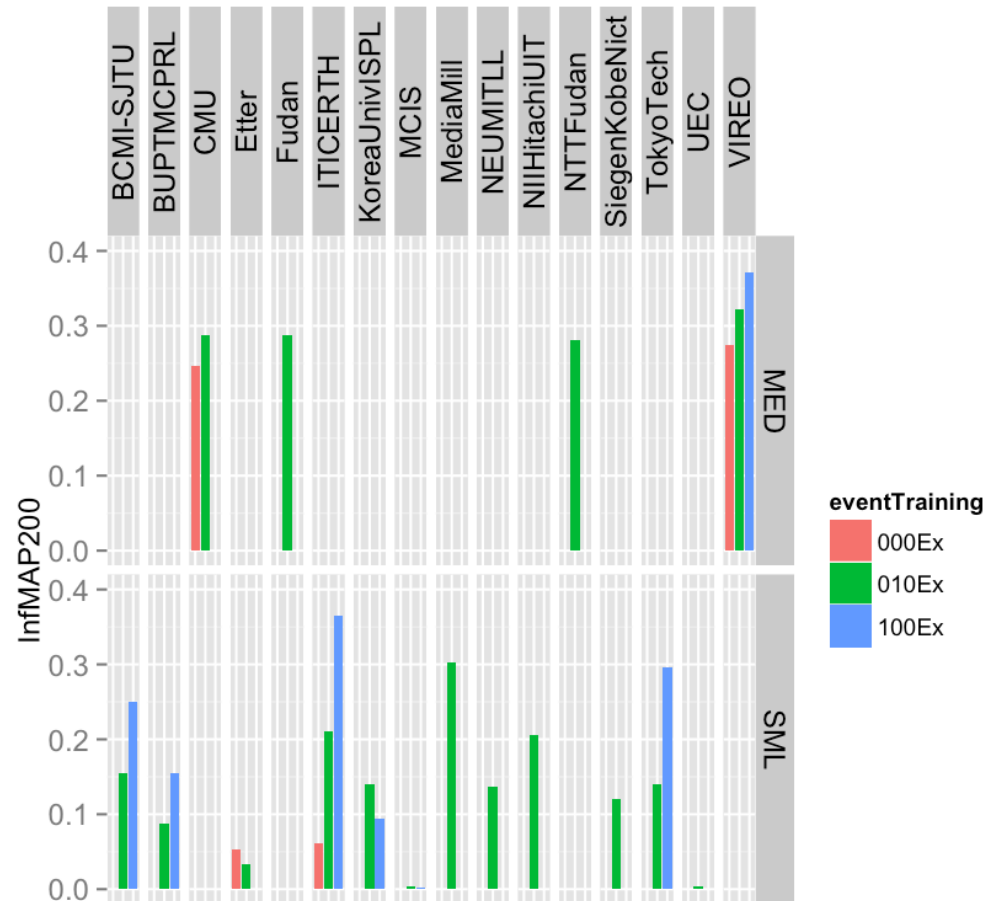
Pre-Specified AP by System and Event

MED15EvalSub, 10Ex, Mixed System Size



MAP → Inferred Mean Average Precision (InfMAP)

- Follows Aslam et. al. procedure, Statistical Method for System Evaluation Using Incomplete Judgments Proceedings of the 29th ACM SIGIR Conference, Seattle, 2006.
 - Stratified, variable density, pooled assessment procedure to approximate MAP
- InfMAP in the 2015 evaluation
 - Scored PS submissions as both MAP and simulated InfMAP with the reference annotation
 - Ad-Hoc Events reference assessment using InfMAP procedures
- NIST ran experiments with 2014 data to optimize the strata sizes and sampling rate
 - Define 2 strata
 - 1-60 → 100 %
 - 61-200 → 20 %

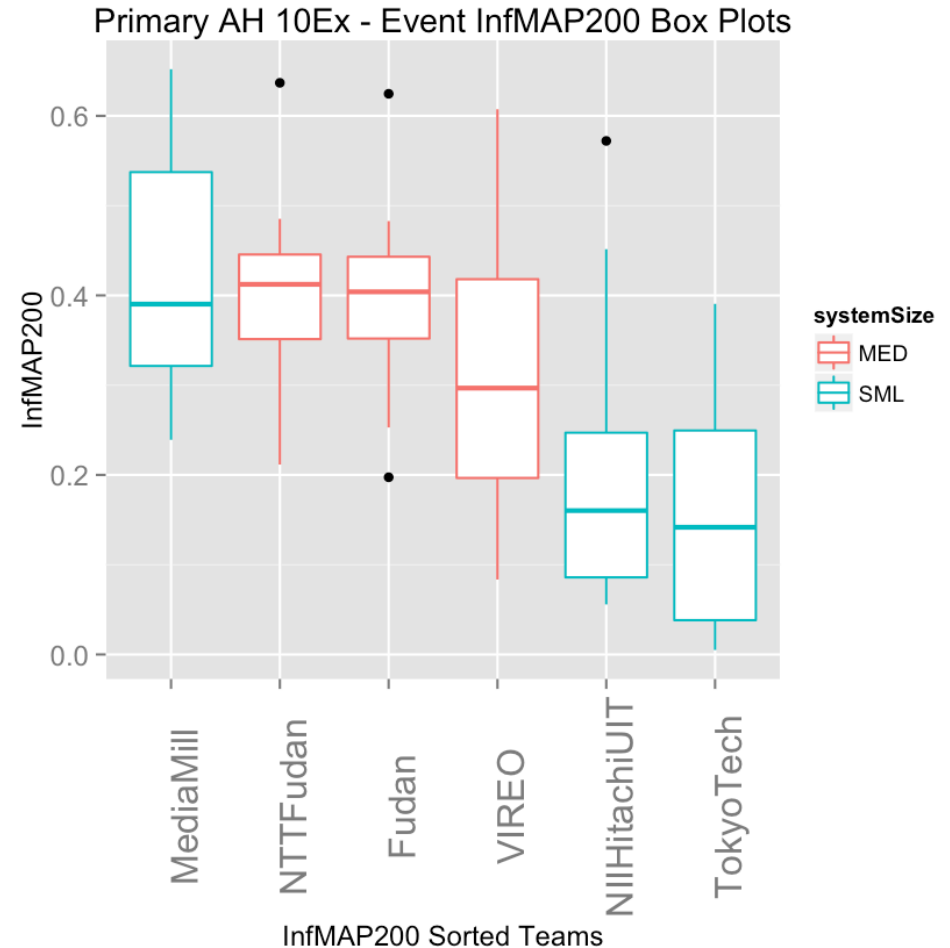


$$M\text{InfAP}200 = 0.011 * \text{MAP} + 0.00203$$

$$R^2 = 0.989$$

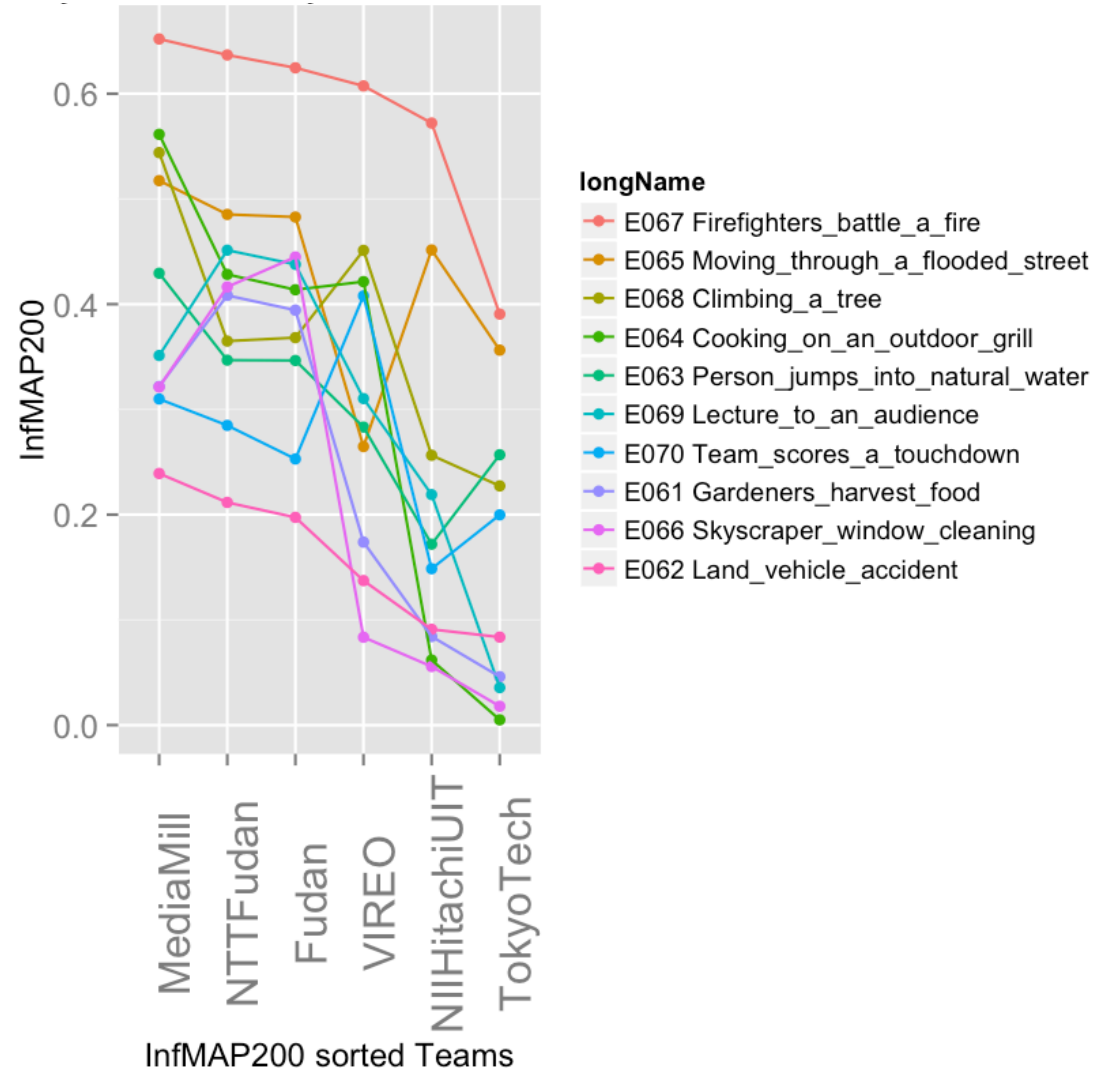
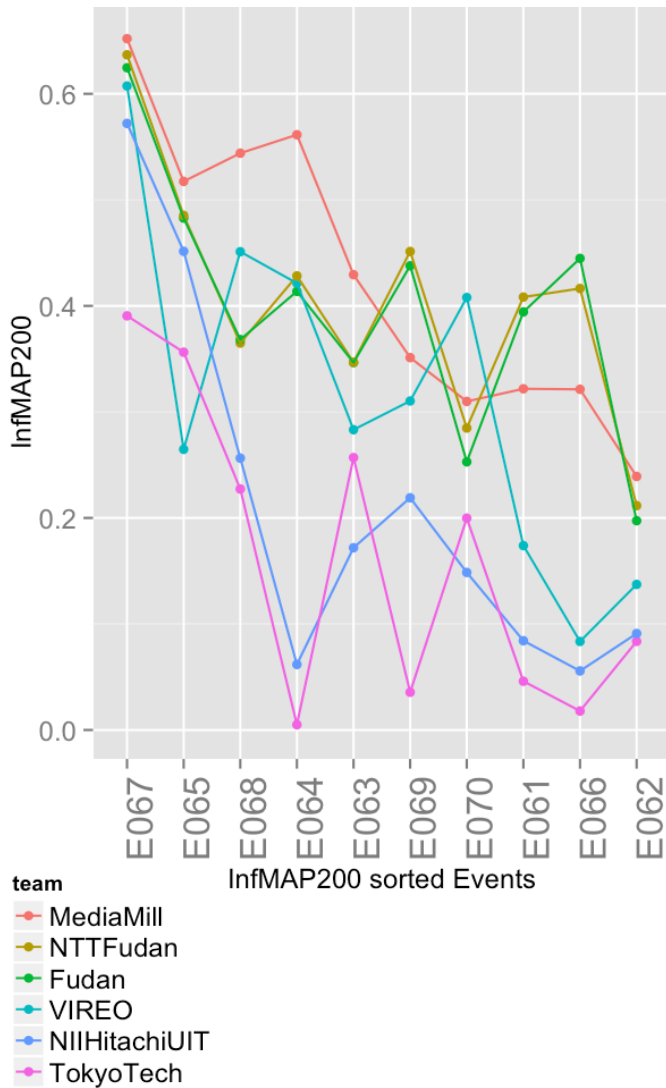
Ad-Hoc Event Results

- 10 new events
 - Exemplars defined using training resources
 - 10 Exemplar training only
- MED15EvalFull the required condition
- Reference Generation
 - Pooled assessment with using all submissions
 - Stats about the pooling/strata
 - 1:60:100%
 - 61:200:20%



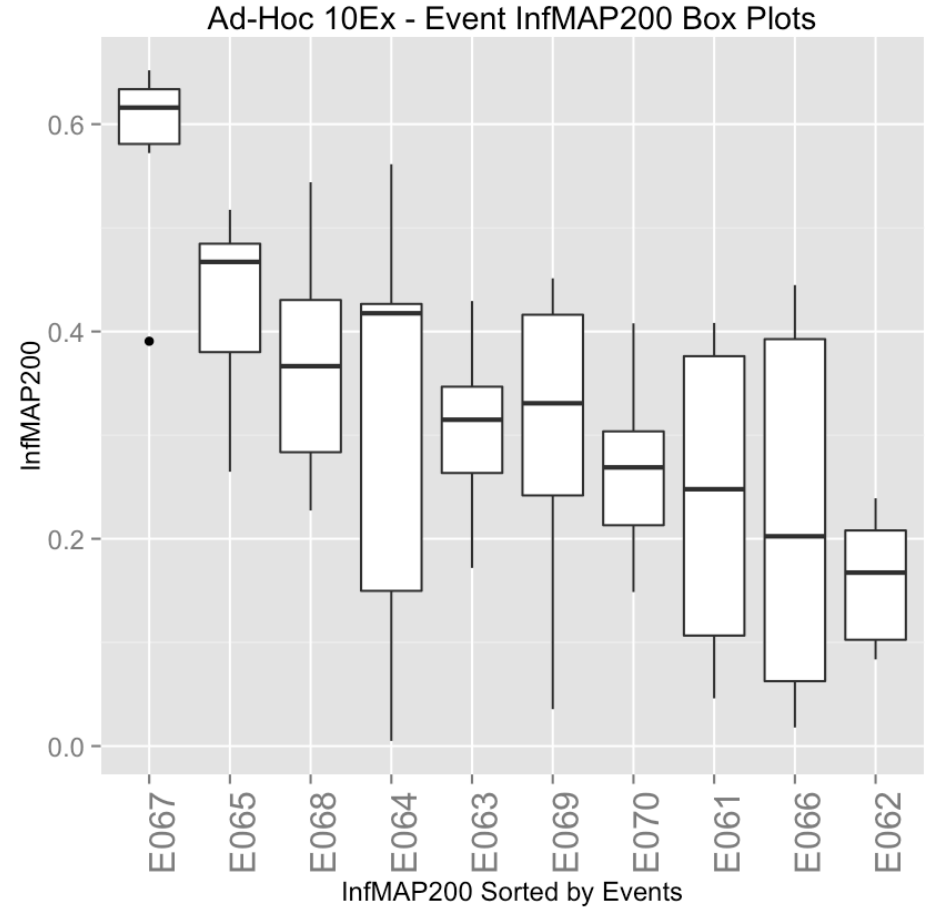
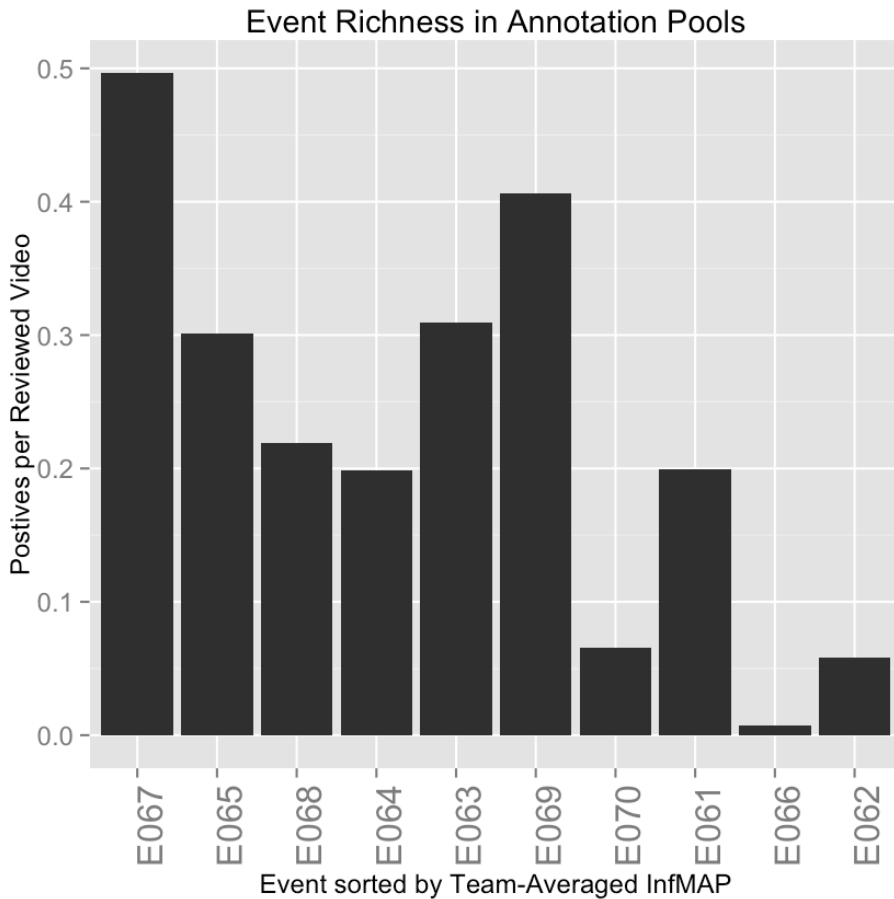
Ad-Hoc InfAP by System and Event

10Ex, Mixed System Size



E067 Firefighters_battle_a_fire
 E065 Moving_through_a_flooded_street
 E068 Climbing_a_tree
 E064 Cooking_on_an_outdoor_grill
 E063 Person_jumps_into_natural_water
 E069 Lecture_to_an_audience
 E070 Team_scores_a_touchdown
 E061 Gardeners_harvest_food
 E066 Skyscraper_window_cleaning
 E062 Land_vehicle_accident

Ad-Hoc Pooled Assessment Event Richness vs. InfAP



MED '15 Summary

- Pre-Specified Results
 - First reporting of event-specific scores
 - Most teams built “Small” hardware systems
 - Most teams processed the subset (1.2K hr.) search collection
- Ad-Hoc Results
 - Limited participation this year – Only teams able to process the full (7.5K hr.) search collection
 - Testing with 10Ex event kits
 - No one participated in the Interactive Event Query test
- The transition from MAP to InfMAP produced correlated results
 - Testing with new collections without exhaustive annotation

Preparation for MED '16

- Ad-Hoc events are a key aspect of MED
 - Current capabilities indicate 10-Exemplar testing is feasible
 - Inexpensive to create/disseminate
 - Does not rely on constructed data
- Pooled assessment and InfMAP produced similar results to MAP
 - Exhaustive annotations not required
- Many new teams built low-resource systems
- Straw proposal for MED '16
 - Use Novel2 as the search collection (Possibility of using the YFCC100M set)
 - PROGRESS data set released as a development collection (date TBD)
 - 10 new Ad-Hoc events
 - Discontinue Pre-Specified Event Testing

Questions?