Video Hyperlinking TRECVid 2015

Roeland Ordelman, Robin Aly
(University of Twente)
Maria Eskevich, Benoit Huet (EURECOM)
Gareth Jones, David Racca (Dublin City University)

Task history

- VideoCLEF 2009 Linking task (speech as source)
- ME'12 S&HL "*brave new*" task:
  - Search & **Linking** (blip.tv)
- ME'13 S&HL "regular" task
  - Search & **Linking (bbc collection)**
- ME'14 S&HL "regular" task
  - Search & **Linking (bbc collection)**
- ME'15 Search & Anchoring
- **TV'15 Video Hyperlinking**

http://www.multimediaeval.org/

Use scenarios

- Exploration of additional information sources while accessing video content in a linear fashion.
- Exploration of video repositories via a structure of linked video segments.
- Creating narratives on the basis of linked video segments.

Example

Fish and Chips Filet Featuring Jamie Oliver
by **EpicMealTime** • 1 month ago • 1,625,276 views
LIKE/FAVORITE this video!!! EpicMealTime featuring a
HD

FISH 'N' CHIPS - VIDEO RECIPE
by **robjnixon** • 1 year ago • 112,781 view
http://tinyurl.com/NickosFacebook http://tir
http://tinyurl.com
HD

Queen Elizabeth I //

by Animalcrossing95 • 1

Just a tribute to Cate Blan
did

NEW HD

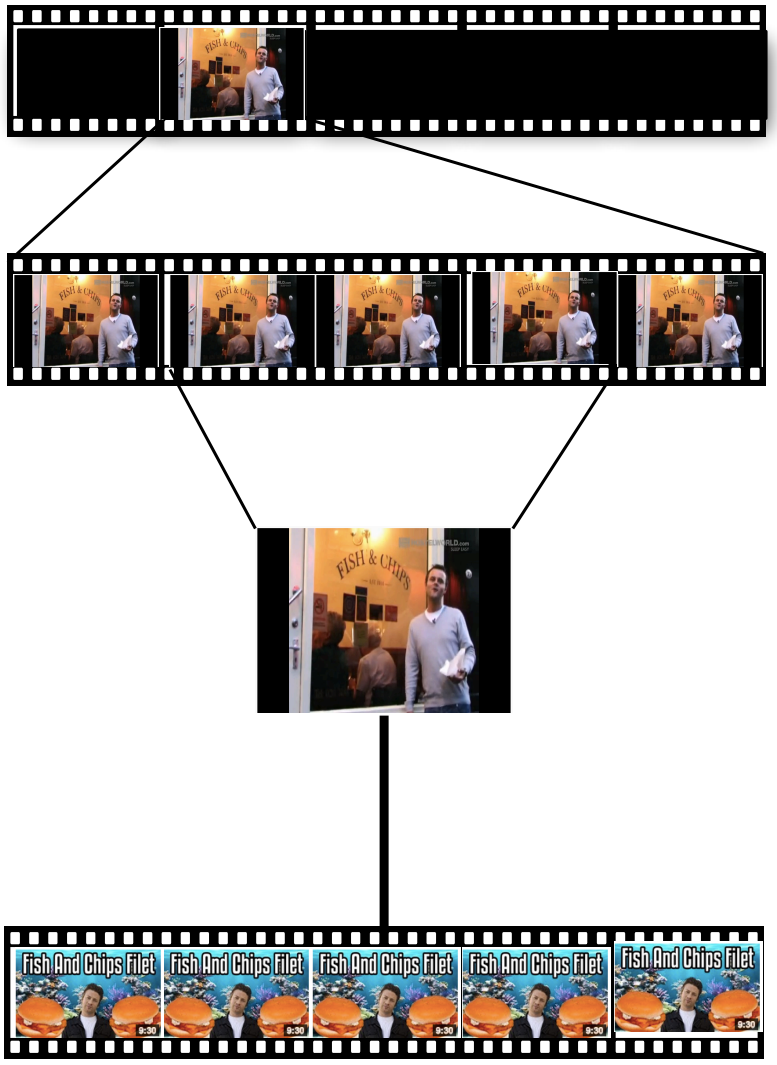
Queen Elizabeth's Historic Move, Discrimination

by ABCNews • 1 week ago • 3,850 views

Queen takes bold step in defending wome

HD

Terminology



- **Video** (e.g., 2 hours)
- **Video clip** (e.g., 10 min)
- **Anchor**: segment (unconstraint) for which a user requests a link (e.g., 1 min) “I want to know more about this”
- **Hyperlink**
- **Target**: relevant segment for given anchor

Target requirements

- Search for *relevant* link targets
- What is relevant?
  - Content **about** what's represented in the anchor (topically related)
  - ~~Content that is **similar**~~

Challenges

- Anchor Identification (not addressed)
  - What segments are useful as anchors? Can we identify these automatically?
- Multimodal Query Extraction
  - Which (multimodal) elements in the anchor define a suitable query to find targets?
- Target Search
  - How to use a (multimodal) query representation for search? How to deal with ambiguity and diversity?
- Target Presentation (not addressed)
  - more complex than merely presenting a list of highest ranked results to the user.

Anchor Creation

Anchor Creation

- Two perspectives:
  - Content creator perspective
    - Selection of anchors adds to better understanding or enhances the experience of users watching the video
    - **A shot of a Rolls-Royce appearing in a video about medieval times**
  - Consumer perspective
    - Selection of anchors is highly personalized and diverse
    - **A shot of a Rolls-Royce appearing in a video about medieval times**

Anchor Creation

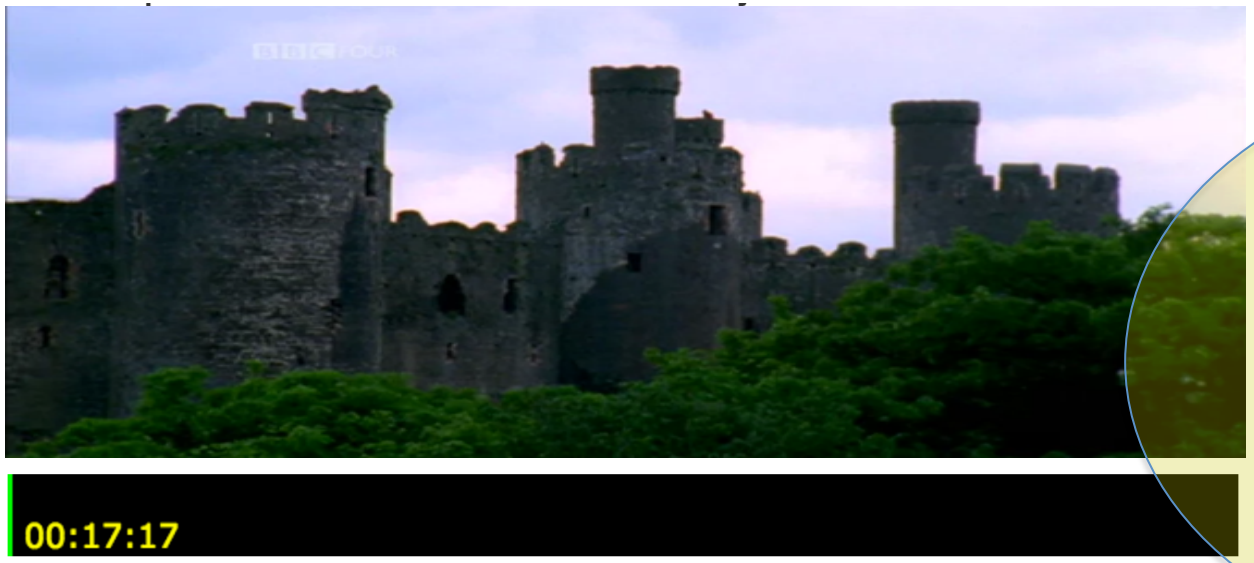
- Creator perspective
  - BBC journalists
  - Employees of British Film Institute
  - Students in Journalism
- Instructions:
  - Face-2-face
  - Teleconference session
- Subjective impression: concept is new, task difficult, but doable.

Preferably residents of UK (familiar with BBC content)

Instructions

- Position yourself in the **role of a producer** wanting to create a new production, e.g., a news item, report or documentary
- S/he is **searching for content in the BBC archive** for this production and selects clips
- Imagine that the producer wants to place hyperlinks in the clips that help the end-user that watches the final program to **understand the program better** or **enrich their watching experience**
- Imagine that these links are provided to end-users for example via a 'second screen' (e.g., iPad)

Interface for Anchor Creation



00:17:17

Start | 00:17:17 | Set | Copy | Go!

End | 00:18:23 | Set | Copy | Go!

When you press this, the player will skip to the defined starting point (SHIFT+i)

Title (new)

Enter title

Description of ideal linked clips (Start with: "*Relevant links have ...*")

Description

Characteristic Visual | Speech | Both

New anchor | Save anchor | End task / select clip

Guidelines | Anchors | Shortcuts

Guidelines

Anchors should be created for one of the following reasons:

- Links may help users to understand the anchor better.
- Links may contain relevant information about the anchor, given what you are currently looking for.
- Links may contain information about occurring objects, places, people, and events that appear in this anchor.

Data

- 100 anchors
  - average length 72sec
  - 11 visual only, 22 speech only, 67 both
- All BBC-owned television broadcasts from the period of 12.05.2008 – 31.07.2008
  - Total length: 2686 hours
  - Removed ~200 videos because of
    - Rebroadcast
    - audio-visual signal was out of sync with subtitles.

Participants

25 registrations
10 submitting participants (40 runs)

CMU-SMU	(Nam+Asia)	Carnegie Mellon Univ. & Singapore Management University
CUNI	(Eur)	Charles University in Prague
VIREO	(Asia)	City University of Hong Kong
DCU	(Eur)	Dublin City University - ADAPT Research Center
EURECOM	(Eur)	EURECOM
IRISA	(Eur)	IRISA Inria Rennes - Bretagne Atlantique
METU_EE	(Eur)	Middle East Technical University
ORAND	(SAm)	ORAND S.A. Chile
TUZ	(Asia)	TUBITAK UZAY
IIP_WHU	(Asia)	Wuhan University - Intelligent Information Processing Lab

EVALUATION

Amazon Mturk Task for Ground Truth Generation

- Task Input:
  - Anchor-Target pairs; from top-10 ranks (max 4 runs per participant)
  - Target description for each anchor
- Task output:
  - Binary relevance judgment : Yes/ No
  - Explanation: why relevant
  - Sanity check: 3-5 meaningful words from anchor + target
- Worker assessment:
  - All fields filled in, sanity check passed -> automatic acceptance
  - Fields missing, sanity check incorrect ->manual inspection

HIT Layout

TASK: WATCH 2 video segments and SAY whether the second video IS RELATED to the first one ACCORDING to the given description.

Please first follow the instructions on the left and then answer the questions on the right side of the screen.

1) Please watch the *first video clip* shown below.

2) Imagine a person watched this *first video clip* on a site like YouTube and wishes to see more video clips that fit the following *description*:

Relevant links have politicans making offensive comments

3) Please watch the following *second video clip* to see whether it satisfies the wish of the person.

4) Based on the *description* given in Section 2, would the person be satisfied watching the *second video clip* after having watched the *first video clip*?

○ **Yes** ○ **No**

***Please be sure that the videos are/can be connected, if you chose the "Yes" option! ***

5) Please give us more details on what connects or disconnets these two video segments? (You can choose more than 1 option from the list below)

☐ Video 2 fits given description	☐ Video 2 does not fit given description,
☐ Video 2 is connected to Video 1	☐ Video 2 is not connected to the Video 1
☐ Same location	☐ Different location
☐ Same objects	☐ Different objects
☐ Same persons	☐ Different persons
☐ Same topic being discussed	☐ Different topic being discussed
☐ Other	☐ Other

6) Please write 1-3 sentences in the box below to explain your decision on the videos relatedness. If you have chosen the option "Other" above, please explain this choice here.

7) Please write 3-5 meaningful words spoken in each of the video clip.

first video clip *second video clip*

NOTES: Please note that in doing this HIT you are taking part in an academic research study. Our review process involves many manual steps. We are also a small team. For this reason, there might be a delay in the approval of your work. We do our best to keep this delay to 2-3 days at the very maximum.

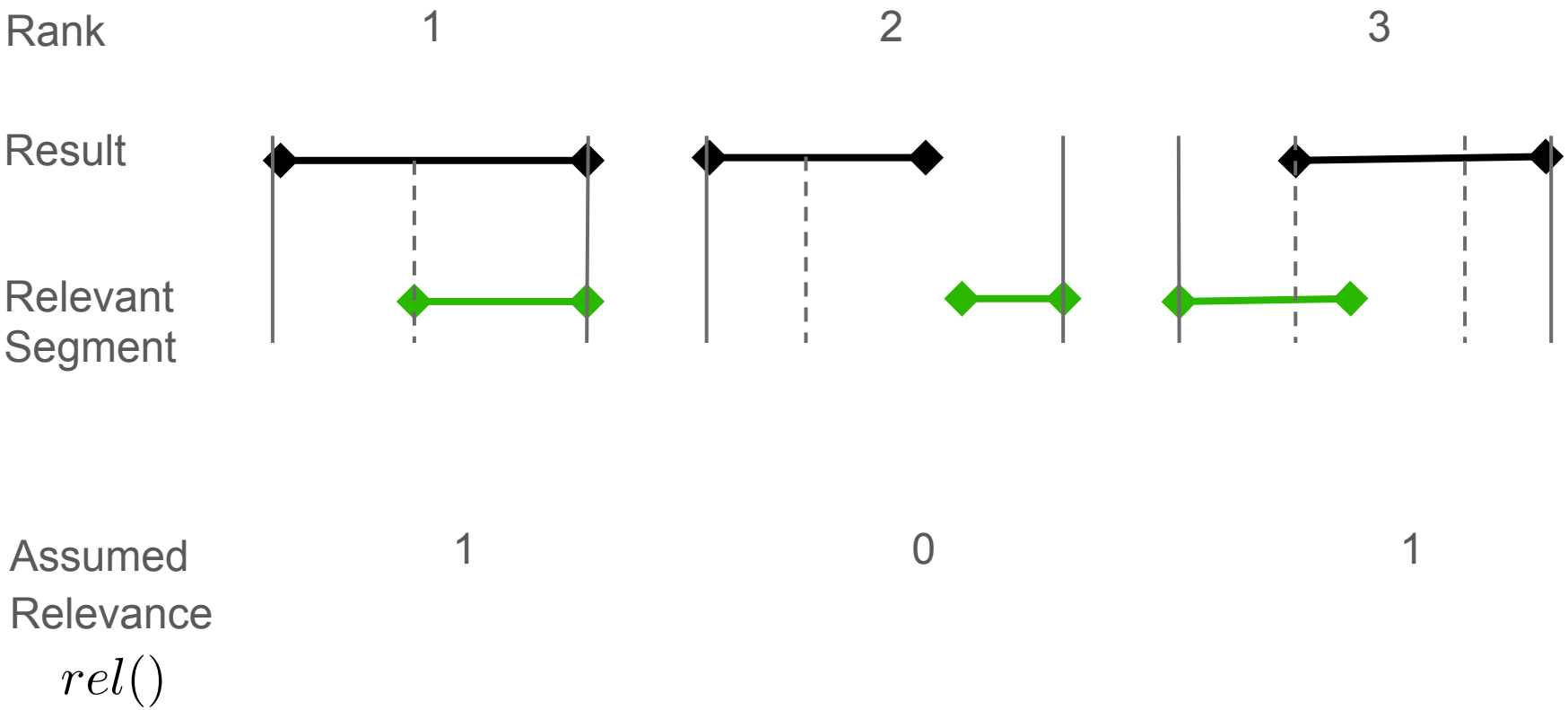
NOTES: It is important that before you submit the HIT you take one more look at the answer that you provided. We ask you to double check that you have written 2-3 complete sentences and that your grammar is OK. We also ask you to check to make sure that the relationship between your sentences and the videos themselves is very clear.

When you are finished with answering the questions, don't forget to click the "Submit" button at the bottom of the page.

Thank you very much for your work!

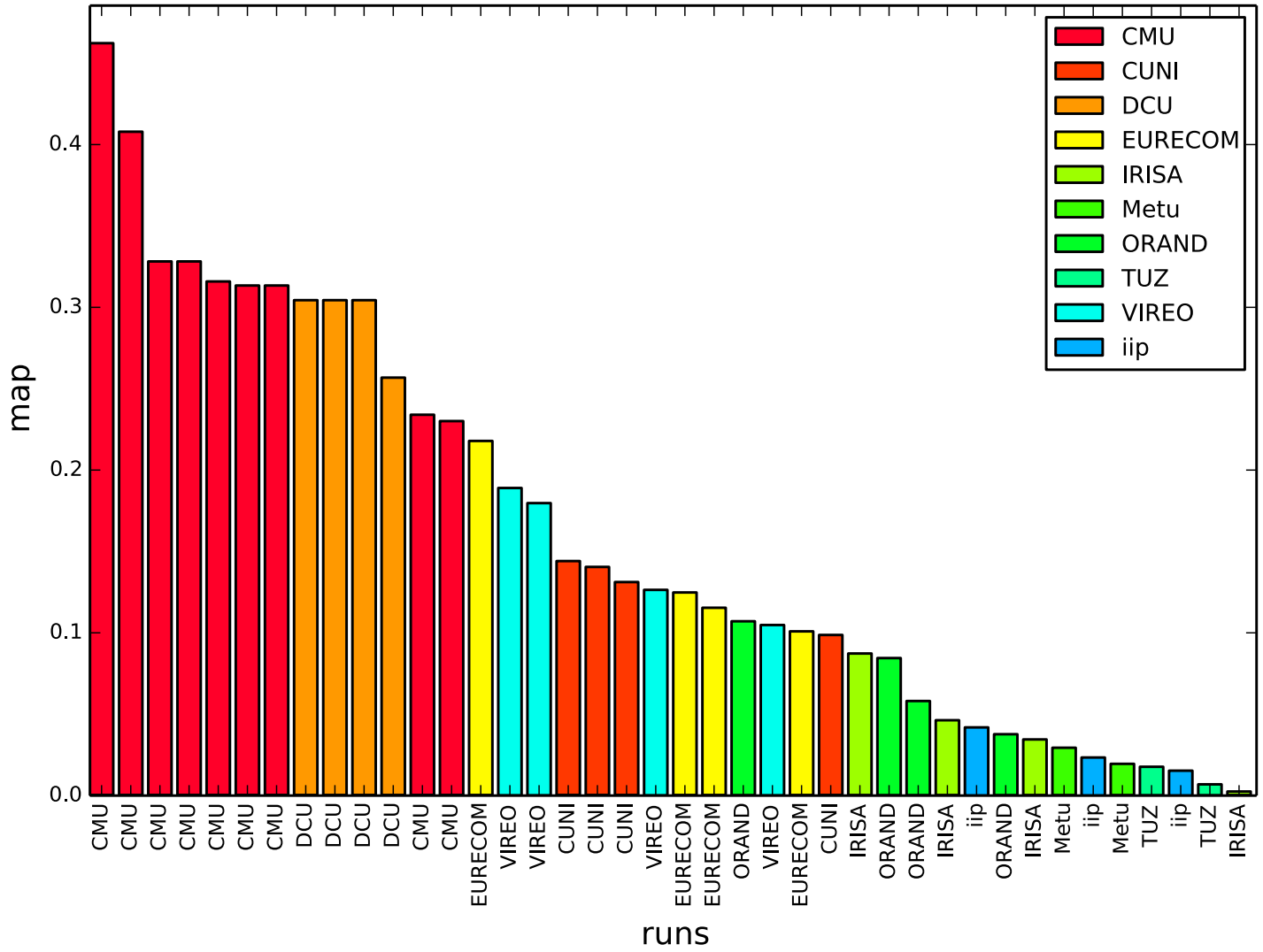
Submit

Adapted MAP

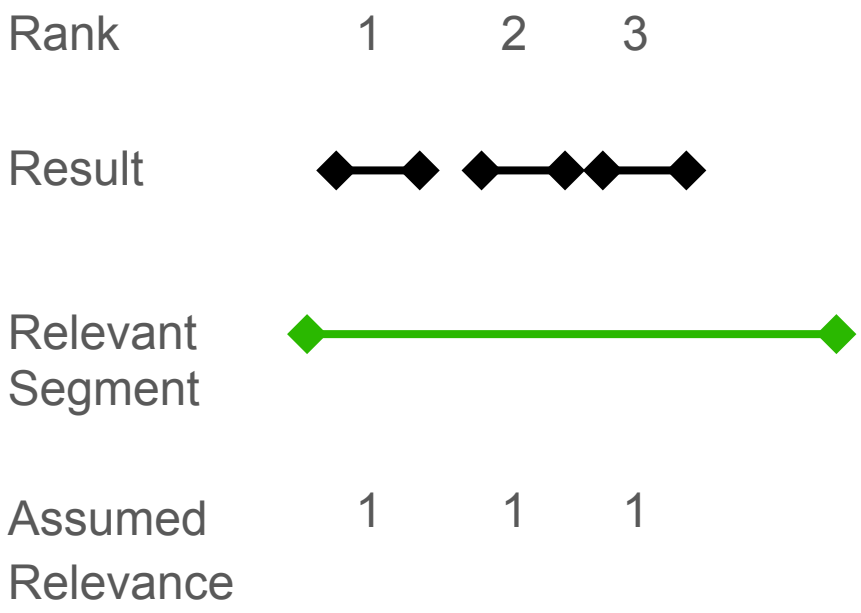


$$ap = \frac{\sum_{r=1}^{n} prec(r) rel(r)}{R}$$

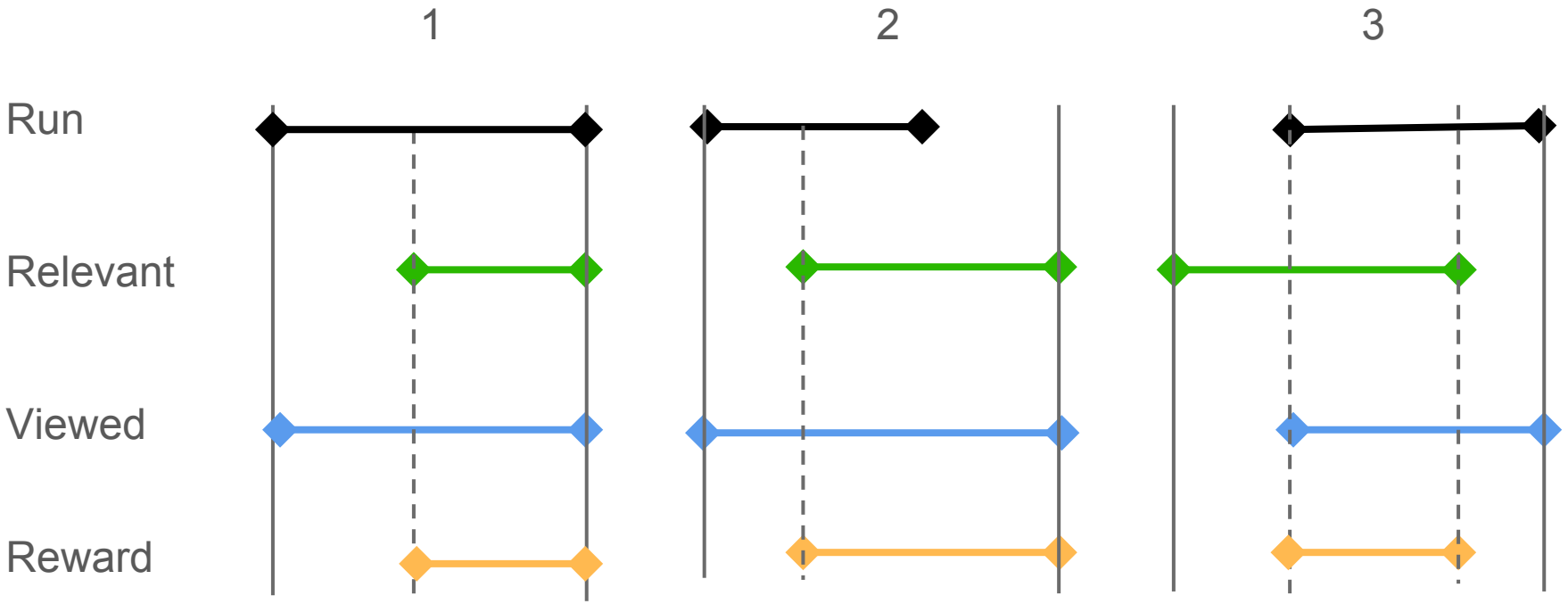
Run Comparison MAP



Adapted MAP



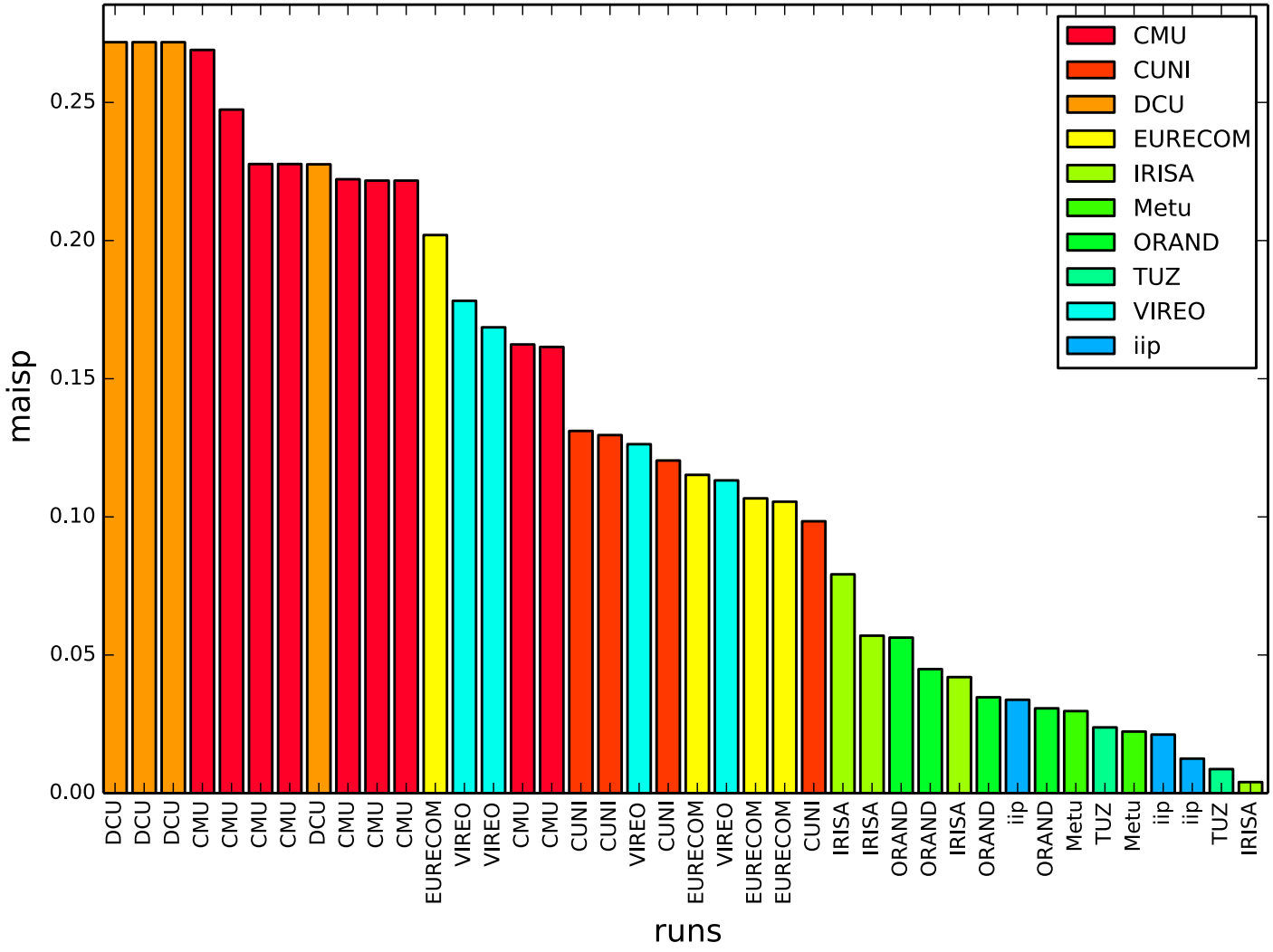
MAiSP Measure



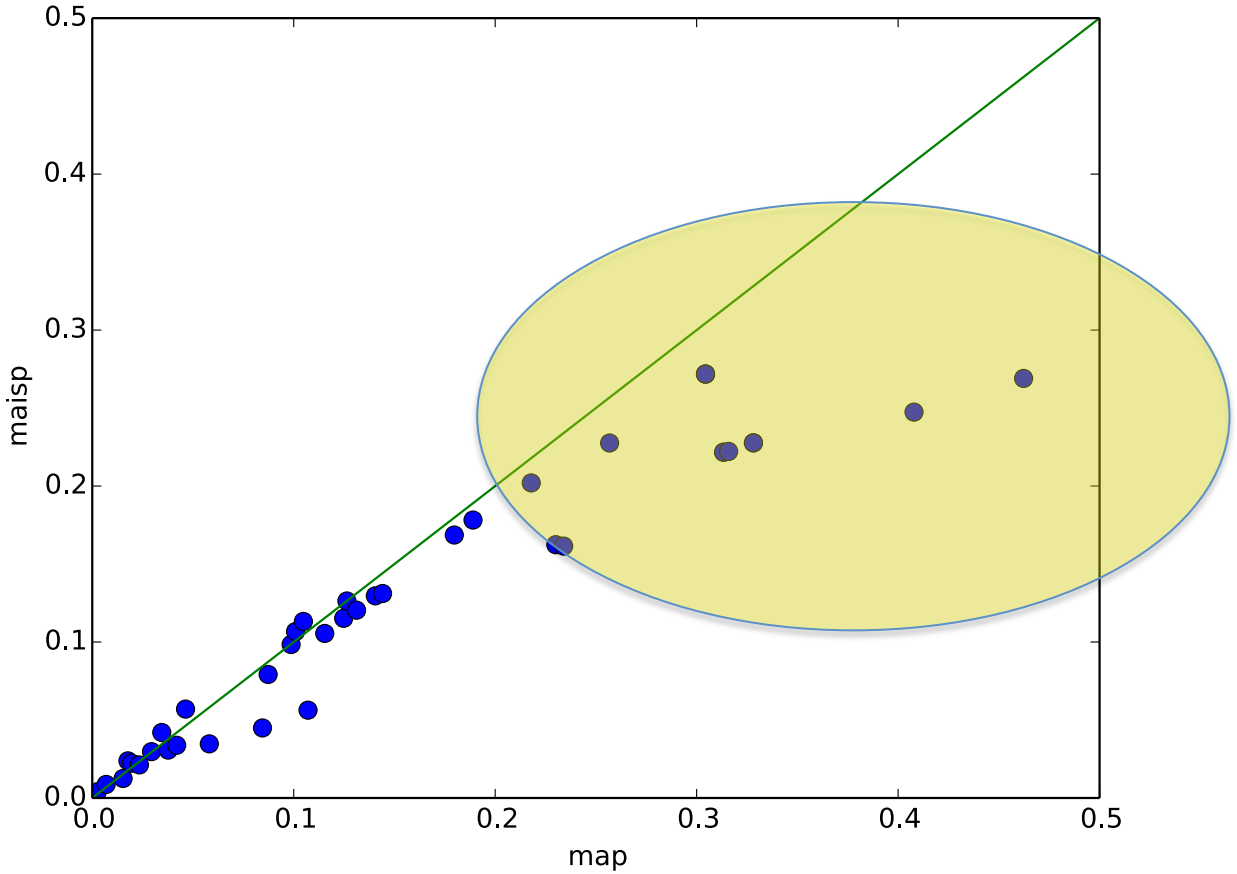
$$avesp = \frac{\int_{t=0}^{\infty} prec(t)rel(t)}{\text{relevant seconds}}$$

https://github.com/robinaly/sh_eval

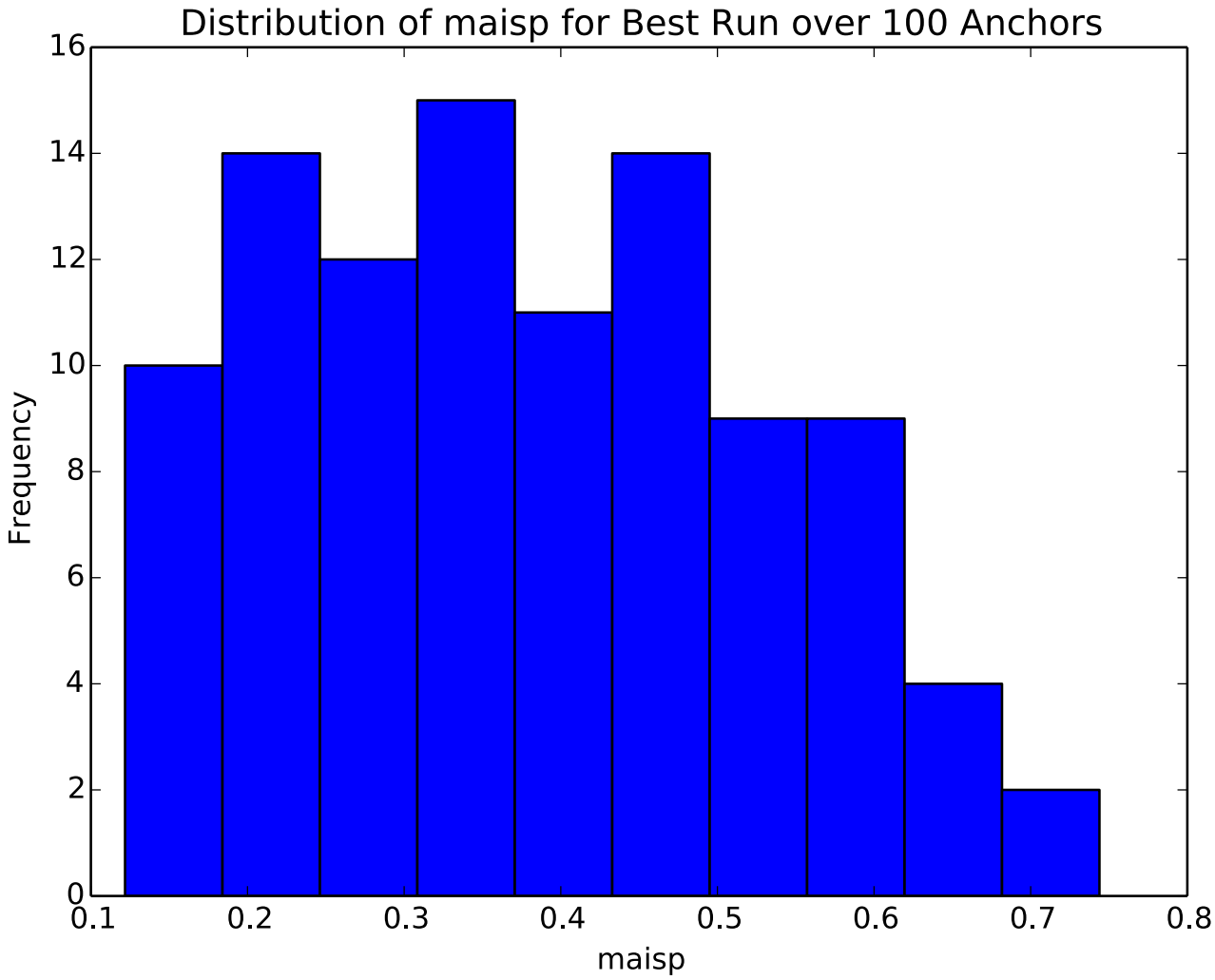
Run Comparison MAISP



Comparison MAP and MAiSP



Task Feasibility



Summary & Conclusions

- Task defined by a practical use scenario
- 100 anchors; 67 multimodal
- 10 participants, 40 runs
- Solution for evaluating free segmentation using MAiSP measure
- Many anchors difficult (max MAiSP < 0.3)
- 0.25 MAiSP reasonable starting point for further exploration and improvement

We are grateful to

Jana Eggink and

Andy O'Dwyer

from the BBC for preparing the collection and hosting the user trials.

The Search and Hyperlinking task was funded by