



# Overview of ImageCLEF lifelog 2017: Lifelog Retrieval and Summarisation

Duc-Tien Dang-Nguyen, Luca Piras, Michael Riegler,  
Giulia Boato, Liting Zhou and Cathal Gurrin

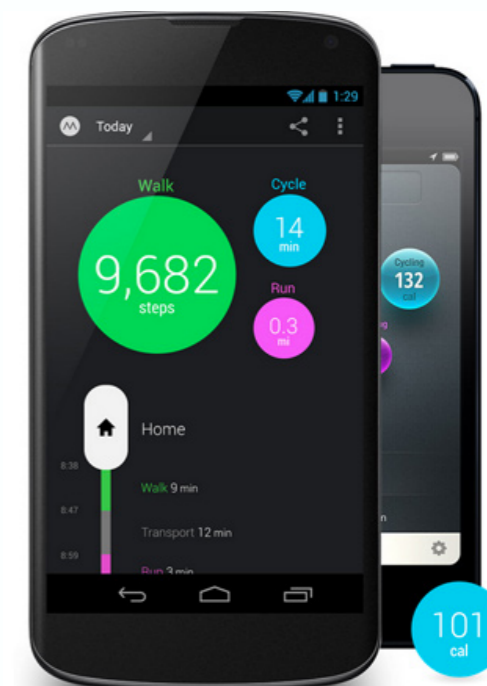


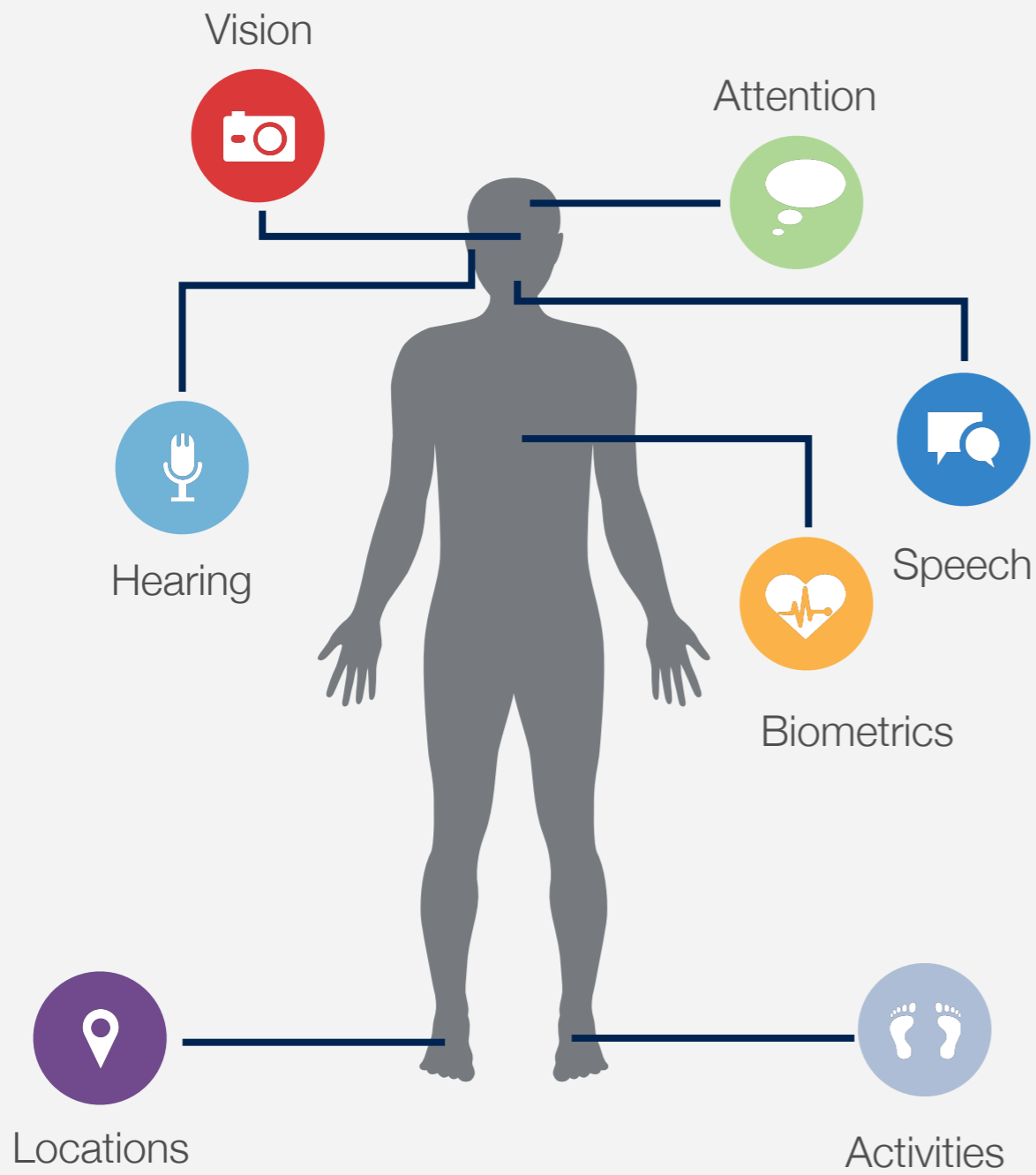




## All-day data gathering

*Narrative Clip 2, OMG Autographers and the Moves smartphone app.*





**Capturing everyday-life-experience**



# Holistic Solutions Require a Multidisciplinary Approach

## Data Processing

A variety of data, different timings, different accuracies, needing different tools.

MULTIMEDIA  
ANALYTICS

## User Experience

Develop fixed and ubiquitous capture & access methods for all stakeholders.

HUMAN  
COMPUTER  
INTERACTION

## Personal Data

The ethics of how to use rich personal data & doing so in a privacy-aware manner.

ETHICS &  
PRIVACY

ETHNOGRAPHY

INFORMATION  
RETRIEVAL

## Search & Retrieval

Scalable & efficient indexing with contextual querying and no defined unit of retrieval.

PERVASIVE  
COMPUTING

## Anywhere, Anytime

Use-cases need pervasive access and contextual querying.

MEMORY



let's face real practical challenges ...

**Lacking of common benchmarking  
datasets and tasks!**

## Kvarsir

A Multi-Class Image-Dataset for Computer Aided Gastrointestinal Disease Detection

## Hard to find

A crowdsourced data set of edited images online

## RAISE

A Raw Images Dataset for Digital Image Forensics

## Heimdallr

A dataset for sport analysis




We have been working on ...

*lifelog is very different*

*however...*

# Dataset Statistics

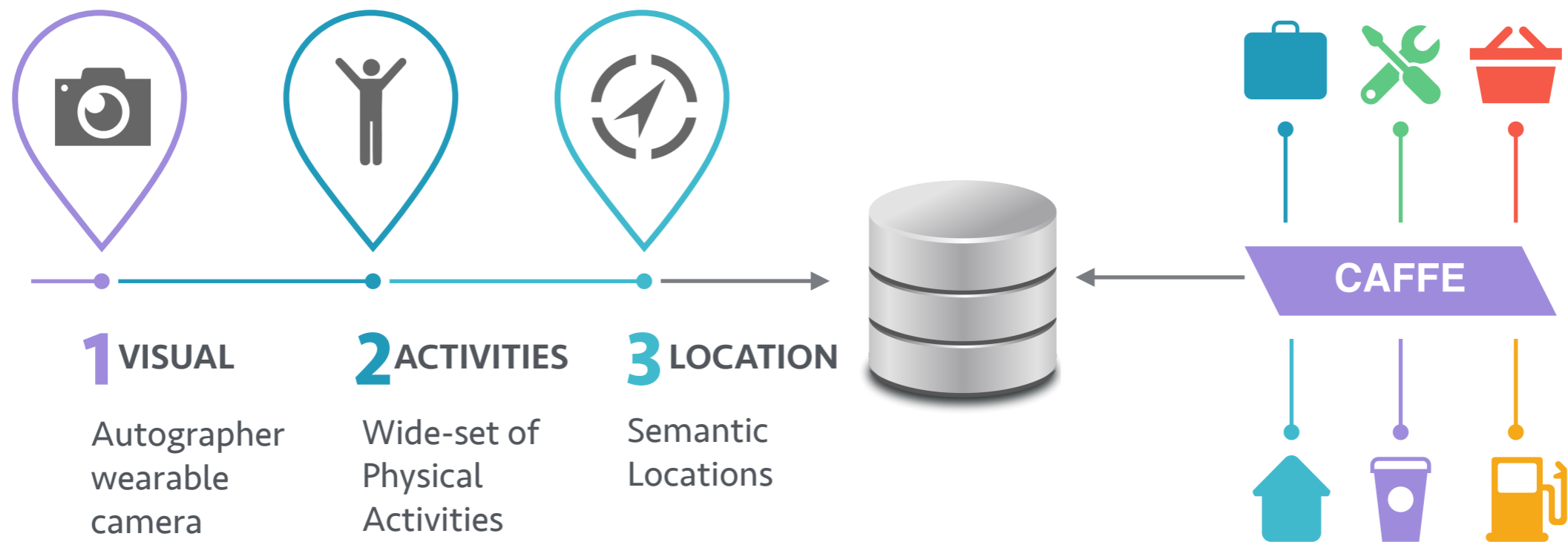
Number of Lifeloggers	
Size of the Collection (GB)	18.18 GB
Number of Images	88,124
Number of Locations	130
Number of Visual Concepts	1,000
Number of LRT Topics	36 (16 devset, 20 testset)
Number of LST Topics	15 (5 devset, 10 testset)







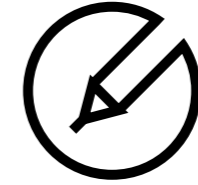
# Features





## Lifelog Retrieval Task

Analyse the lifelog data and according to several specific queries **return** the correct answers.



## Lifelog Summarization Task

Analyse the lifelog data and **summarize** them according to specific requirements.

The summary should be represented by 50 images, and it is required to be both relevant and diverse.



# An example of a LRT topic

Topic: Having a Drink

Query: Find the moment(s) when user u1 was having a drink in a bar with someone.

Description: Any moment in which the user is clearly seen having a beer or other drink in a bar venue is considered relevant. Having a drink in any other location (e.g. a cafe), or without another person present is not considered relevant. The type of drink is not relevant once it is presumed alcoholic in nature and not tea/coffee.

# An example of a LST topic

Topic: Eating Habits

Query: Summarize the moment(s) when user u1 was eating or drinking.

Description: User u1 wants to know insights of his eating/drinking habits. He would like to have a summary of what, when, where, and whom together he was eating or drinking. To be relevant, the images must show entirely or partially visible food/drink. Blurred or out of focus images are not relevant. Images that are covered (mostly by the lifelogger's arm) are not relevant, even if they are recorded while the user was eating.

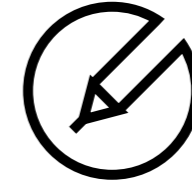


Evaluation metric:

NDCG (Normalized Discounted Cumulative Gain)

Different depths are used, i.e.,  $NDCG@N$ , where  $N$  will vary based on the type of the topics, for the recall oriented topics  $N$  will be larger ( $>20$ ), and for the precision oriented topics  $N$  will be smaller  $N$  (5 or 10 or 20).

Final score: Arithmetic mean of all topics.



Evaluation metrics:

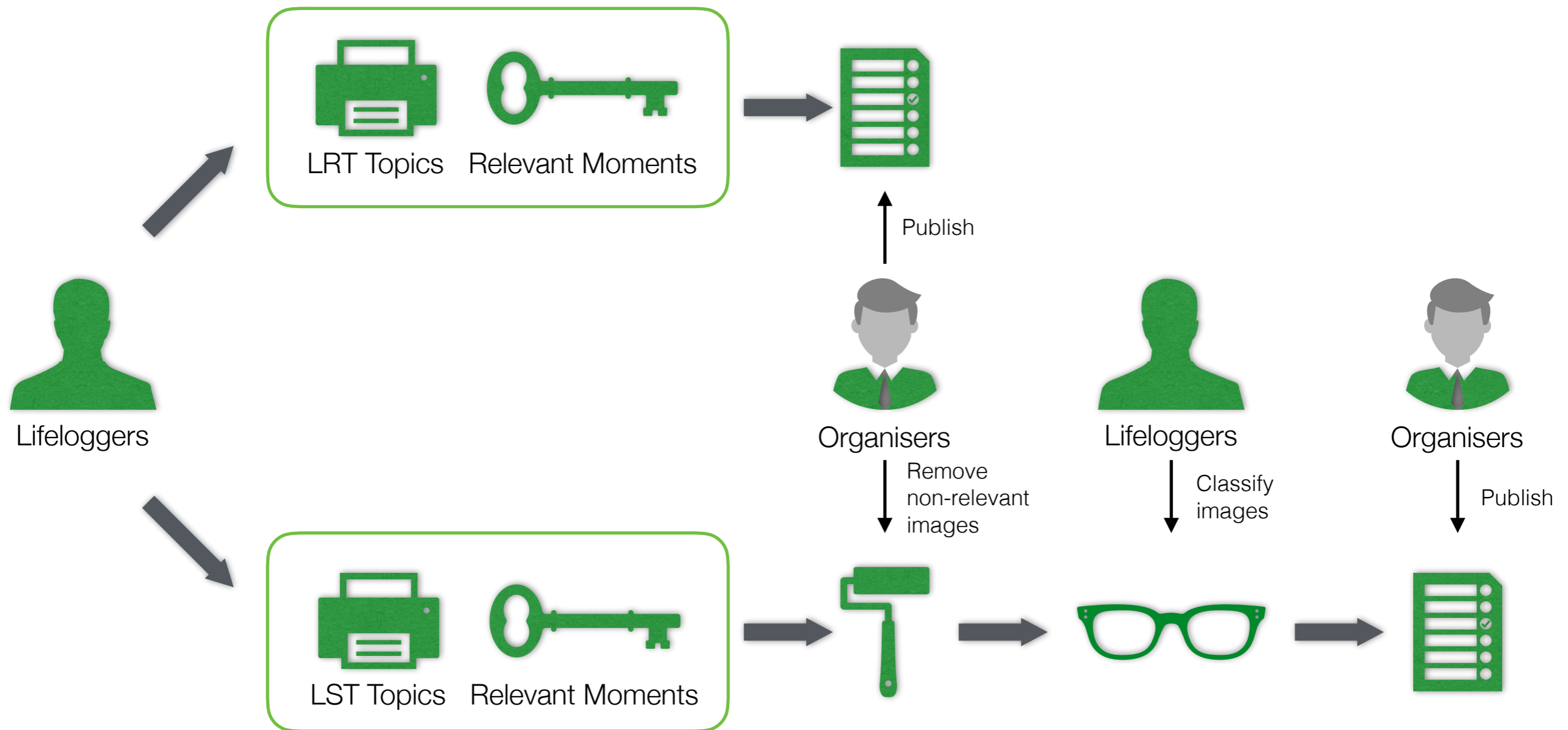
- Cluster Recall at  $X$  ( $CR@X$ ) - a metric that assesses how many different clusters from the ground truth are represented among the top  $X$  results;
- Precision at  $X$  ( $P@X$ ) - measures the number of relevant photos among the top  $X$  results;

Official metric: **F1-measure at 10** ( $F1@10$ ) - the harmonic mean of the previous two at cut off point of 10.

Final score: Arithmetic mean of all topics.



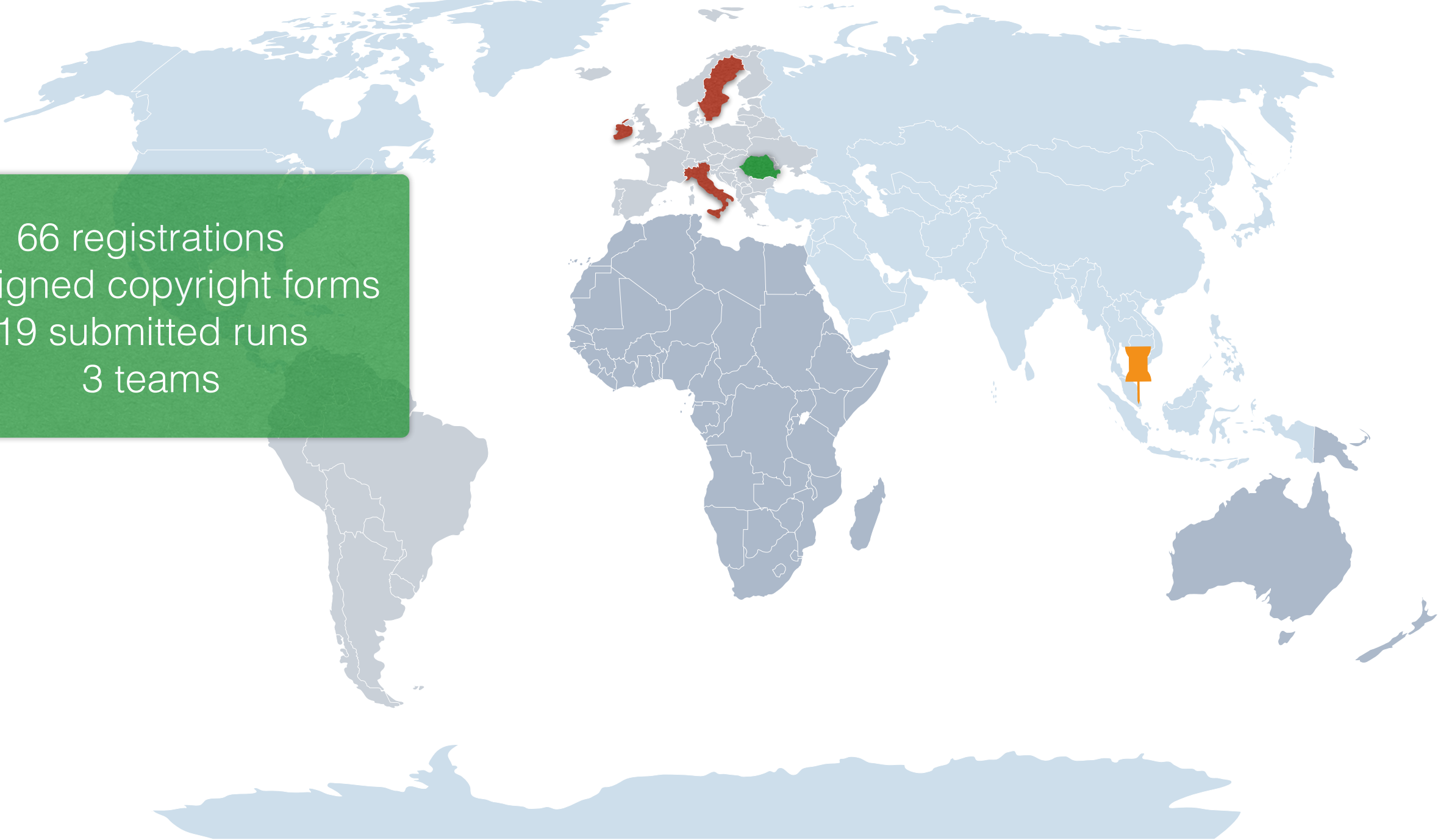
# Topics and Ground-truth



# Some statistics

	<b># Topics</b>	<b>Min</b>	<b>Max</b>	<b>Average</b>
<b>Devset - LRT</b>	16	14	1087	259
<b>Testset - LRT</b>	20	25	1039	214
<b>Devset - LST - Clusters</b>	5	5	34	16.6
<b>Testset - LST - Clusters</b>	10	5	47	15.6
<b>Devset - LST - Images per cluster</b>	5	5	343	30
<b>Testset - LST - Images per cluster</b>	10	3	746	52

# Teams



A world map with several regions highlighted in different colors. In Europe, the British Isles, Scandinavia, and Italy are highlighted in red. A small area in Eastern Europe is highlighted in green. In Southeast Asia, a region is highlighted in orange and marked with a pushpin. A green callout box on the left contains the following statistics:

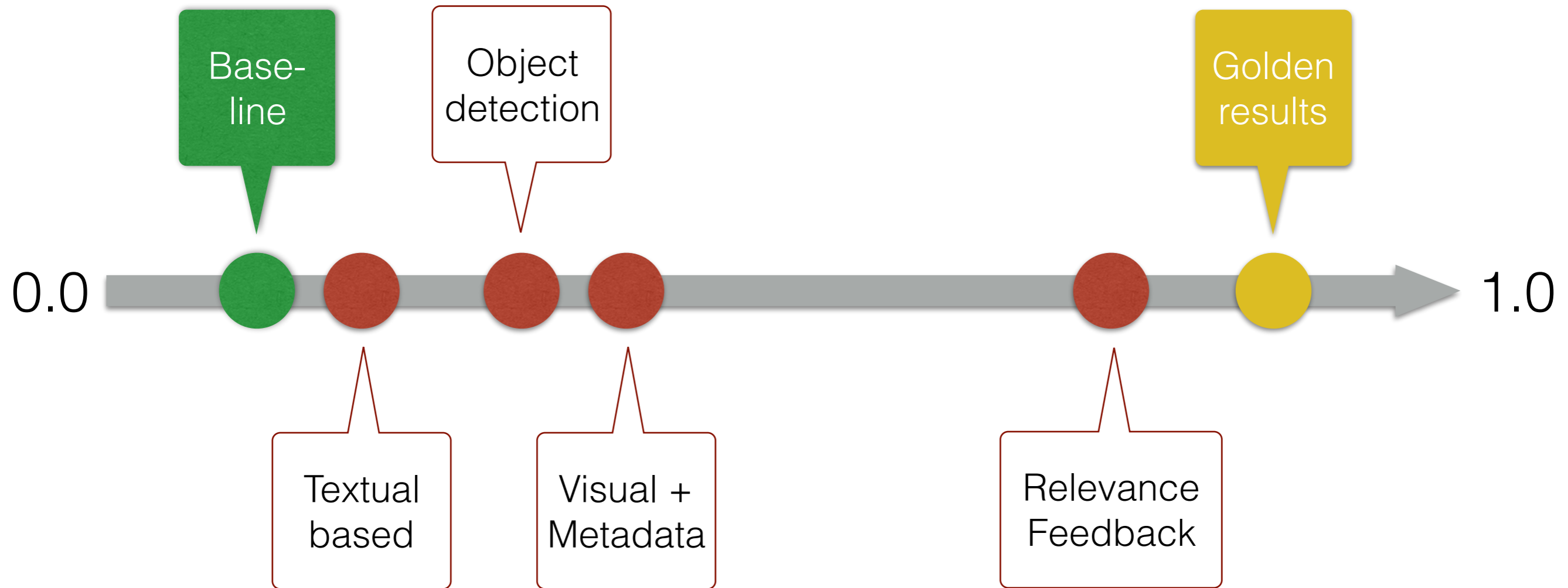
66 registrations  
21 signed copyright forms  
19 submitted runs  
3 teams

Team	Run	Description
Organisers	Baseline	Baseline method, fully automatic.
	Segmentation	Apply segmentation and automatic retrieval based on concepts.
	Fine-tuning	Apply segmentation and fine-tuning. Using all information.

Team	Run	Description
Organisers	Baseline	Baseline method, fully automatic.
	Segmentation	Apply segmentation and automatic retrieval and diversification based on concepts.
	Filtering	Apply segmentation, filtering, and automatic diversification. Using all information.
	Fine-tuning	Apply segmentation, fine-tuning, filtering, and automatic diversification. Using all information.
	RF	Relevance feedback. Using all information.
UPB	Run 1	Textual filtering and word similarity using WordNet and Retina.
I2R	Run 1	Parameters learned for maximum F1 score. Using only visual information.
	Run 2	Parameters learned for maximum F1 score. Using visual and metadata information.
	Run 3	Parameters learned for maximum F1 score. Using metadata.
	Run 4	Re-clustering in each iteration; 20% extra clusters. Using visual, metadata and interactive.
	Run 5	No re-clustering. 100% extra clusters. Using visual, metadata and interactive.
	Run 6	Parameters learned for maximum F1 score. Using visual, metadata, and object detection
	Run 7	Parameters learned for maximum F1 score, w/ and w/o object detection.
	Run 8	Parameters learned for maximum F1 score. Using visual information and object detection.
	Run 9	Parameters learned for maximum precision. Using visual and metadata information.
	Run 10	No re-clustering. 20% extra clusters. Using visual, metadata and interactive.



# Results

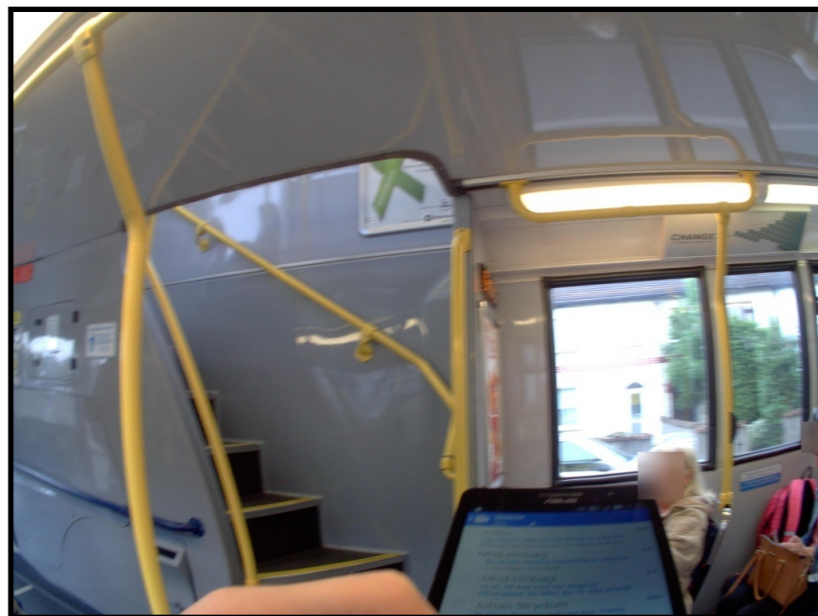


# Easy vs. Hard topics

## T8. Transporting

Query: Summarize the moments when user u2 using public transportation.  
Description: Photos taken inside a car or a taxi are not relevant. Blurred or out of focus images are not relevant. Images that are covered are not relevant.

0.81





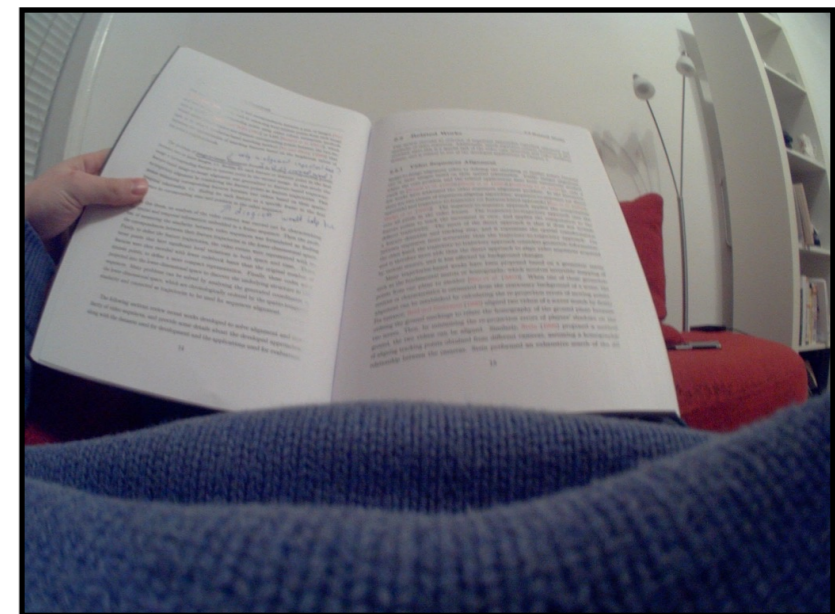
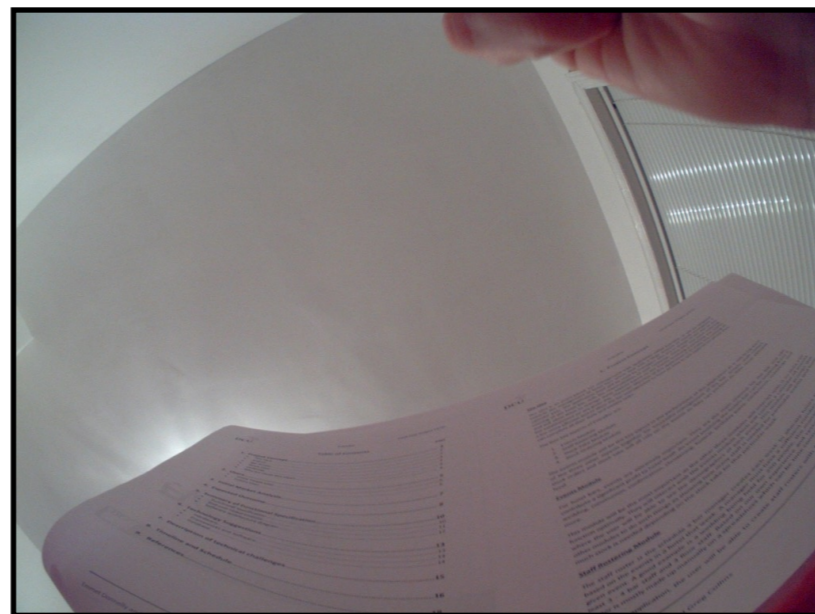
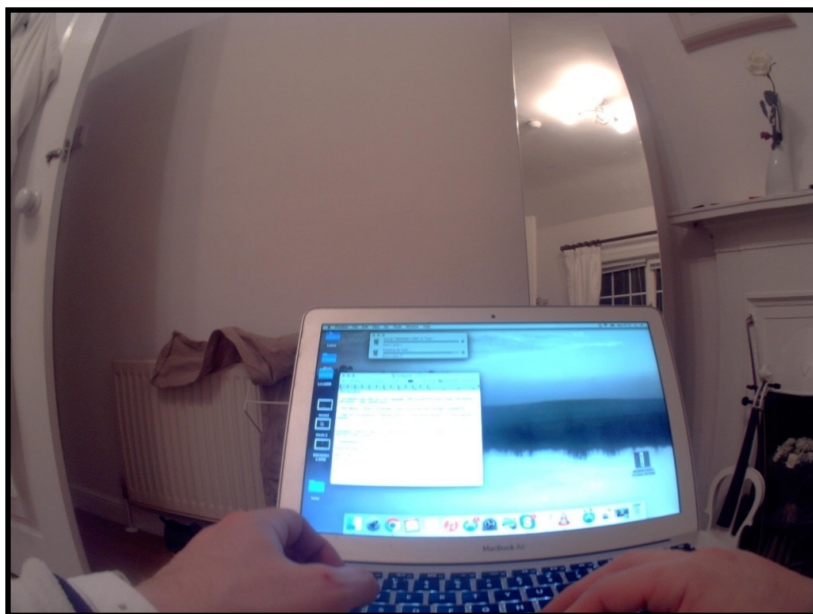
# Easy vs. Hard topics

## T4. Working at home

Query: Find the moment(s) in which user u1 was working at home.

Description: To be consider to relevant, the user should be using computer for work, reviewing an article or taking some notes at home. Using computer for entertainment is not relevant.

0.08



# Official Results

## Lifelog Retrieval

Team	Best run	Run type	Avg. NDCG
Organisers*	Fine Tuning	Mixed	0.386

## Lifelog Summarisation

Team	Best run	Run type	Avg. F1@10
I2R	Run 2	Mixed	<b>0.497</b>
UPB	Run 1	Mixed	0.132
Organisers*	Relevance Feedback	Mixed	0.769

\* **NOTE:** These runs are not ranked since they are from the organisers team.



Raise more awareness for lifelogging in general



Rich of data, more are coming 01

International teams from different continents 02

Different multimodal methods from different ideologies, exploiting different sources of information 03

01 Tasks seemed to be very complex.

02 The scope of the subtasks should be better defined.

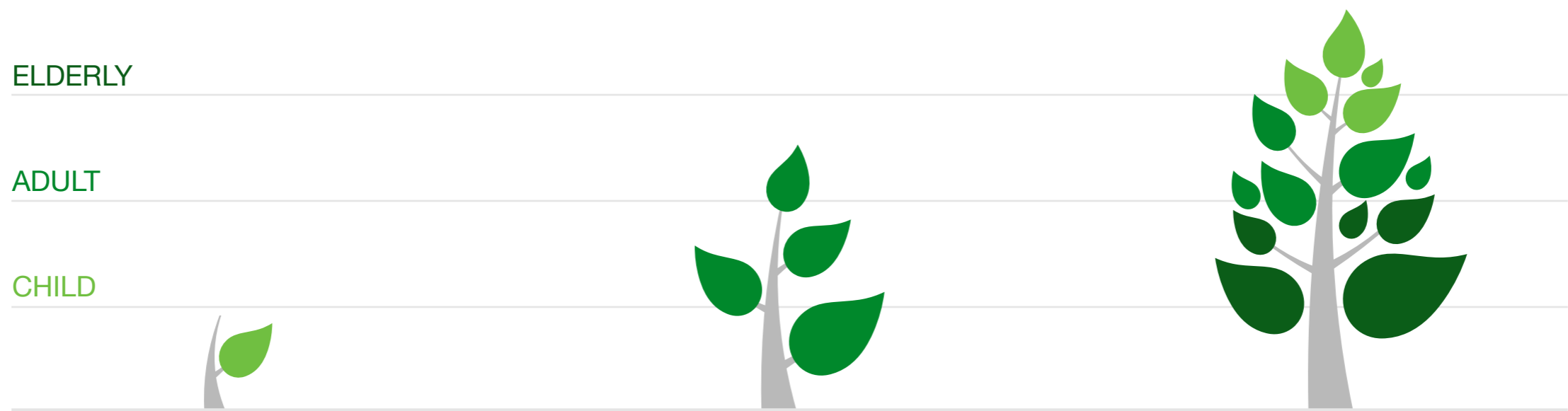
03 Lacking of pre-extracted features.



Large gap between signed-up teams and the submitted ones

# Vision

In reality, we can't yet imagine the use-cases of lifelogs, but it could become a **permanent companion** assisting you throughout life. Constantly growing in size.



## 1. Knowledge Support

Focus on supporting knowledge acquisition and learning in the early years.

## 2. Productivity

From education to the workplace, providing information and insights to assist productivity and fitness.

## 3. Health

Into old age, providing support for cognition and health to maintain independence and activity.

# Future tasks

ImageCLEFlifelog 2018 and NTCIR-14 (proposals submitted)





January 16-18, 2018 FUNCHAL, MADEIRA–PORTUGAL ([www.icpram.org/INDEED.aspx](http://www.icpram.org/INDEED.aspx))

CALL FOR PAPERS: 7<sup>th</sup> International Conference on Pattern Recognition Applications and Methods

SPECIAL SESSION ON  
**INSIGHTS DISCOVERY FROM LIFELOG DATA (INDEED)**

SUMMARY AND SCOPE

The aim of this special session is to draw together researchers involved in this new and challenging area. The sought contributions can be either methodological on learning and data analysis (e.g., active and adaptive learning, hypothesis generation, data managed techniques, combining classifier techniques, etc.) or focused on solving real world challenges (e.g., human behavior analysis, content-based multimedia retrieval, and many others).

TOPICS

- Multimedia data analytics for lifelog data
- Search and retrieval from lifelogs
- User experiences of lifelog gathering and application
- Real-world applications of lifelogging
- Captology and behaviour change from personal data
- Egocentric vision and first-person camera vision analytics

IMPORTANT DATES

Submission Deadline:  
**November 7<sup>th</sup>, 2017**

Authors Notification:  
**November 21<sup>st</sup>, 2017**

Camera Ready and  
Registration:  
**November 29<sup>th</sup>, 2017**

ICPRAM Conference:  
**February 16<sup>th</sup>-18<sup>th</sup>, 2018**

# Thank you!

Question?



[dangnguyen@insight-centre.org](mailto:dangnguyen@insight-centre.org)

