



ImageCLEF evaluation activities

Theodora Tsikrika

University of Applied Sciences Western Switzerland (HES-SO)
Switzerland



Media Search cluster meeting
8th FP7 Networked Media concertation meeting
Brussels, December 13, 2011

Overview

- ImageCLEF
 - History (2003-2011)
 - Evaluation tasks
 - Resources: Collections, Topics, Ground Truth
- ImageCLEF impact
- ImageCLEF 2012 evaluation activities
- Future directions
- Conclusions

ImageCLEF : cross–language image retrieval

- Introduced in 2003 as part of the Cross-Language Evaluation Forum
 - February-September annual cycle
- Aim: advancement of visual media analysis, indexing, classification, retrieval
 - develop **infrastructure** for the evaluation of visual information retrieval systems operating in monolingual and cross–language contexts
 - provide reliable and reusable **resources** for such benchmarking purposes
 - promote the **exchange of ideas**
- **Multi-disciplinary**
 - (visual) information retrieval, cross–lingual information retrieval, computer vision and pattern recognition, medical informatics, etc.
- Tasks
 - (ad-hoc) image **retrieval**
 - image **annotation** (object recognition, image classification, concept detection)
 - interactive image retrieval

Organisation

Overall **coordination**: MedGIFT group, HES-SO, Switzerland

Task **organisation**: several researchers per task

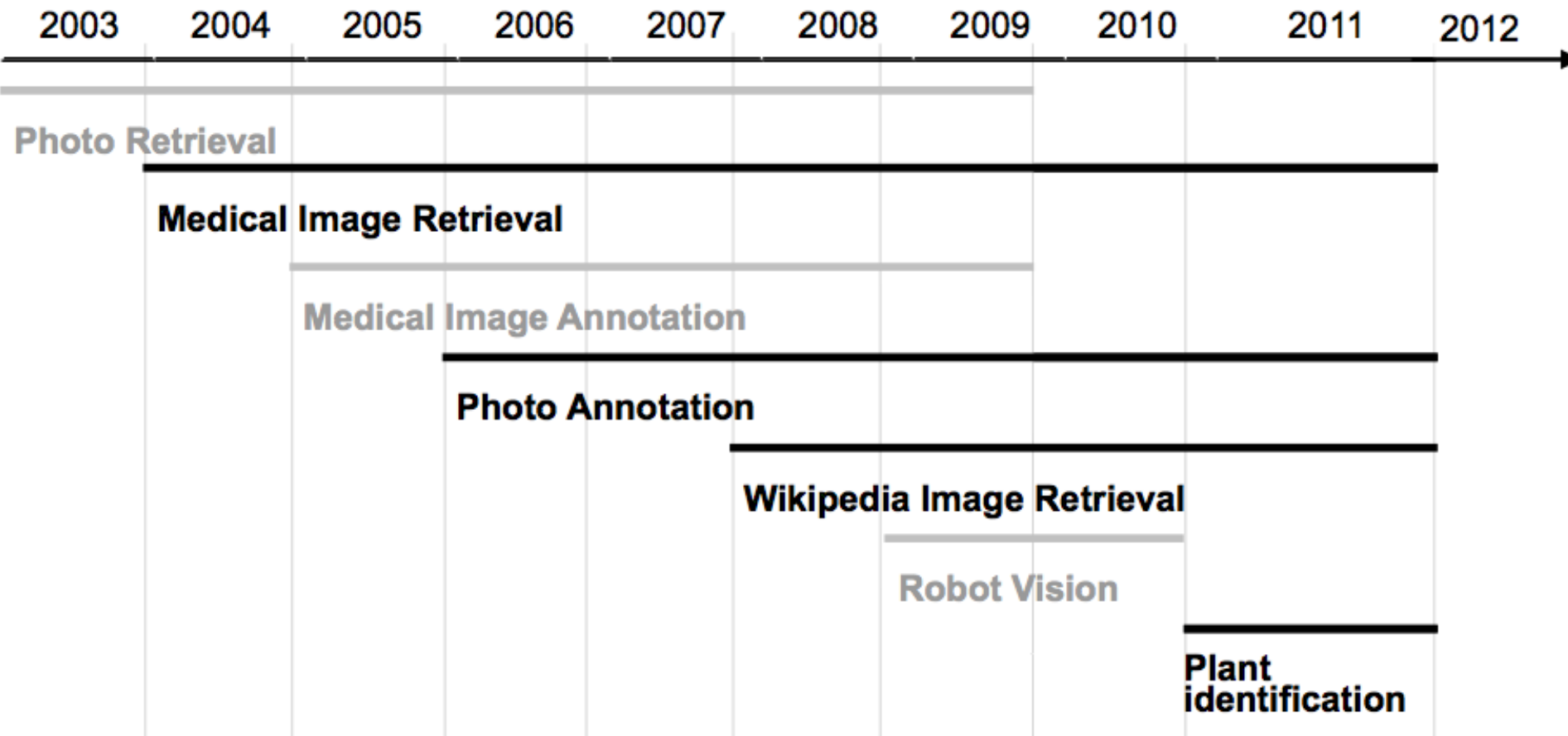
Support: international and national funding bodies, voluntary effort



ImageCLEF tasks

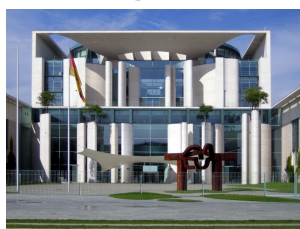
- **General images** (generic photos, historic archives, news, Web, Flickr, Wikipedia)
 - photo retrieval, photo annotation, Wikipedia image retrieval
- **Medical images**
 - medical image retrieval, medical image annotation
- **Scientific multimedia data**
 - plant identification
- **Robotics**
 - robot vision

ImageCLEF tasks timeline (2003-2011)



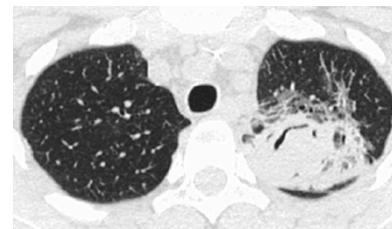
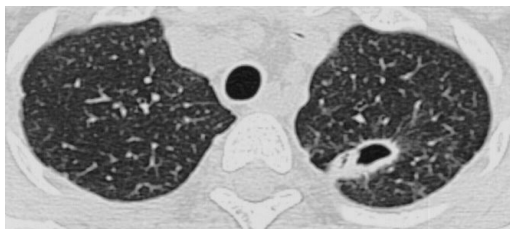
ImageCLEF tasks : general images

- **Photo retrieval** (2003-2009)
 - Given a multimedia information need, find relevant images (as many as possible)
 - Evaluation criteria: topical relevance, diversity
 - Datasets: historical, generic (travel), news
- **Photo annotation** (2006-)
 - object recognition, **concept** detection (objects, scene, abstract, quality, sentiment)
 - Datasets : PASCAL VOC, generic (travel), FlickrR (+ tags)
- **Wikipedia image retrieval** (2008-2011)
 - simulate image retrieval on the Web
 - larger **scale** image collections
 - highly **heterogeneous** textual descriptions and visual content
 - multimodal image retrieval approaches



ImageCLEF tasks : medical images

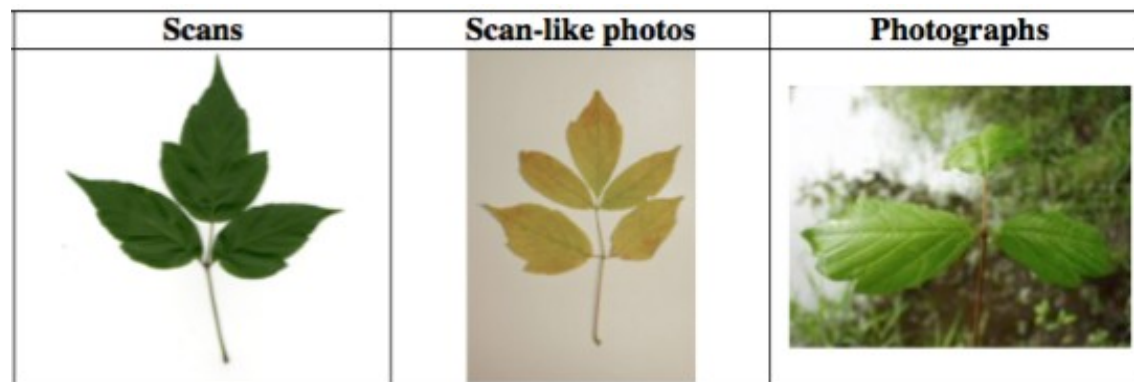
- **Medical image retrieval** (2004-)
 - (visual) clinical decision support for medical diagnosis
 - **Image-based** retrieval: clear information need for finding images
 - **Case-based** retrieval: find a case similar to the one under observation
 - Images + medical cases from clinical practice and biomedical literature
 - Topics + ground truth in collaboration with clinicians



- **Medical image annotation** (2005-2009)
 - Content-based image classification in medical applications
 - Dataset: medical radiographs collected from daily clinical routine
 - Classes: hierarchical Image Retrieval in Medical Applications code

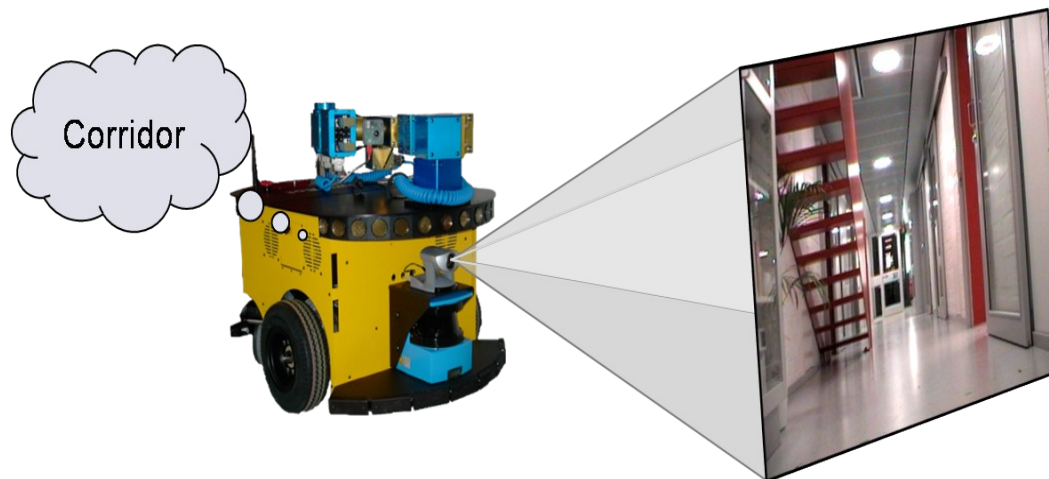
ImageCLEF tasks : scientific multimedia data

- **Plant identification** (2011-)
 - Automatic plant species identification based on images of leaves
 - Motivation: bridge **taxonomic gap** essential for
 - ecology management
 - biodiversity preservation
 - agriculture development
 - Organised in collaboration with botany scientists
 - **citizen sciences initiative**
 - Telabotanica: a French social network of amateur and expert botanists



ImageCLEF tasks : robotics

- **Robot vision** (2009-2010)
 - Semantic localisation of a mobile robot using **visual place recognition**
 - Determine the robot's **topological location** for each image in a sequence
 - classify correctly known rooms/functional areas
 - detect new rooms, not seen during training
 - Images acquired using a perspective camera or a stereo camera
 - Achieve **robustness** under varying imaging conditions



Collections (number of images x1000)

	2003	2004	2005	2006	2007	2008	2009	2010	2011
<i>General (non-medical) images</i>									
Photo retrieval	28.1	28.1	28.1	20	20	20	498.9		
Photo annotation (object and concept detection)				15.1	22.6	2.8	18	18	18
Wikipedia image retrieval						151.5	151.5	237.4	237.4
Robot vision							22.5	9.6	
Plant identification									5.4
<i>Medical images</i>									
Medical image retrieval		8.7	50	50	66.7	66	66	66	231
Medical image annotation			10	11	12	13	14.1		

Topics / Classes

	2003	2004	2005	2006	2007	2008	2009	2010	2011
<i>General (non-medical) images</i>									
Photo retrieval	50	25	28	60	60	39	50		
Photo annotation (object and concept detection)						17	53	93	99
Wikipedia image retrieval						75	45	70	50
Robot vision							5	14	
Plant identification									70
<i>Medical images</i>									
Medical image retrieval		26	25	30	30	30	30	30	30
Medical image annotation			57	116	116	196	57 116 196		

Ground truth

- Coverage
 - Whole collection for classification tasks
 - Pooling for retrieval tasks
- Assessors
 - **Volunteers**
 - Task organisers, participants, others
 - **Experts**
 - Clinicians, botanists
 - **Crowd**
 - Amazon Mechanical Turk workers, social networks
- Ground truth creation still a bottleneck in building benchmarks

ImageCLEF book

ImageCLEF: Experimental Evaluation in Visual Information Retrieval

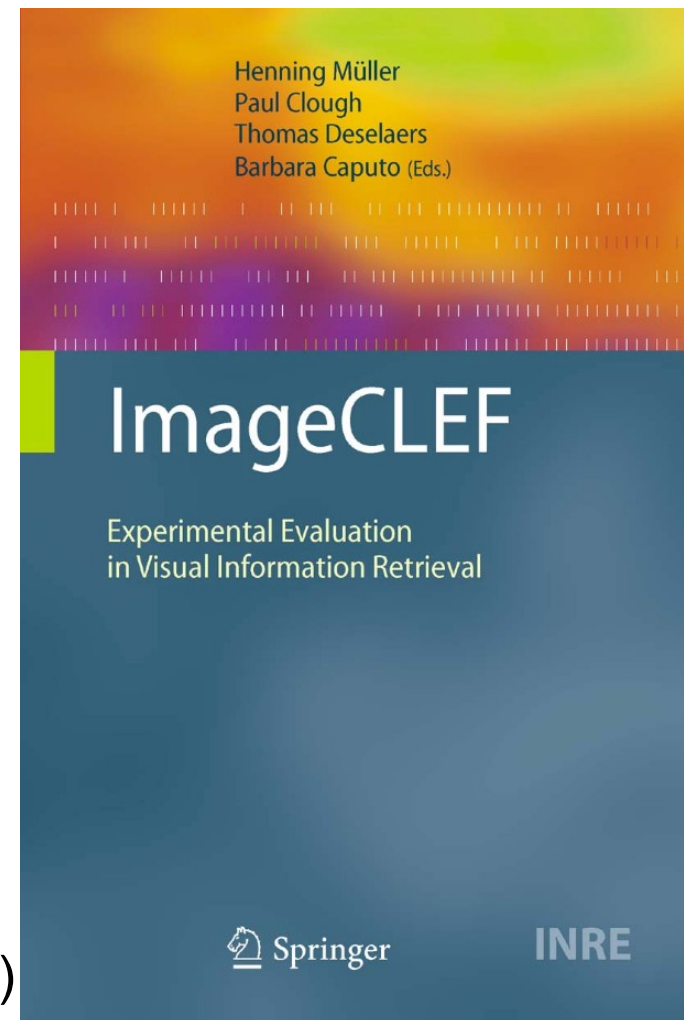
The Information Retrieval Series, Vol. 32

Müller, H.; Clough, P.; Deselaers, Th.; Caputo, B. (Eds.)

1st Edition., 2010, 495 pages

Contents

- Basic concepts (6 chapters)
 - history, datasets, topic development, relevance assessments, evaluation, fusion approaches
- Task reports (7 chapters)
- Participants' reports (11 chapters)
- External perspectives on ImageCLEF (3 chapters)

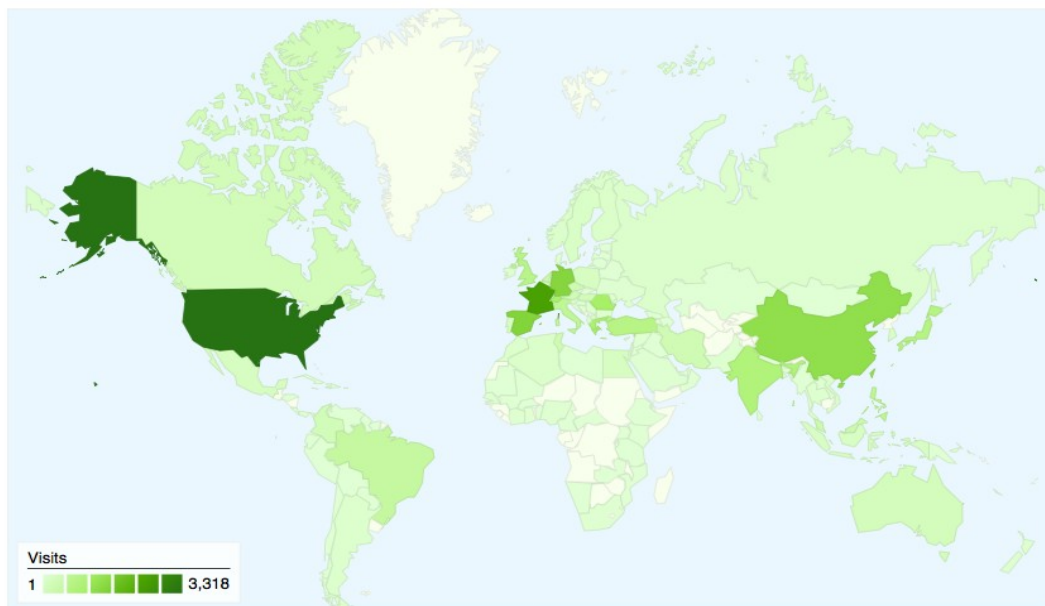


Participation

	2003	2004	2005	2006	2007	2008	2009	2010	2011
<i>General (non-medical) images</i>									
Photo retrieval	4	12	11	12	20	24	19		
Interactive image retrieval	1	2	2	3		6	6		
Photo annotation (object and concept detection)				4	7	11	19	17	18
Wikipedia image retrieval						12	8	13	11
Robot vision							7	7	
Plant identification									8
<i>Medical images</i>									
Medical image retrieval		12	13	12	13	15	17	16	17
Medical image annotation			12	12	10	6	7		
<i>Total</i>									
Participations	4	17	24	30	35	45	65	47	43
Registrations							84	112	141

ImageCLEF web site access (2011 campaign)

	Total Sept 2010 – Sept 2011	Per Month Average Sept 2010 – Sept 2011	Per Month Average Jan 2011 – Sept 2011
Unique visitors	~12,000	~1,100	~1,200
Visits	~25,000	~2,050	~2,250
Page views	~73,000	~6,100	~6,700



24,903 visits came from 138 countries/territories

Assessing the impact of evaluation campaigns

- Assess the impact of the research they foster
 - Research that otherwise would not have been possible
 - **Economic** impact: technology transfer, time/effort saved for researchers, ...
 - **Scientific** impact: scholarly impact, patents, ...
- **TREC**: B. R. Rowe, D. W. Wood, A. N. Link, and D. A. Simoni. Economic impact assessment of NIST's Text REtrieval Conference (TREC) Program. Technical Report, Project Number 0211875, RTI International, 2010.
- **TRECVID**: C. V. Thornley, A. C. Johnson, A. F. Smeaton, and H. Lee. The scholarly impact of TRECVID (2003–2009). JASIST, 62(4):613–627, 2011.
- **ImageCLEF**: T. Tsirikika, A. G. Seco de Herrera, and H. Müller. Assessing the Scholarly Impact of ImageCLEF. In Proceedings of CLEF 2011.

ImageCLEF scholarly impact: preliminary study

Preliminary study

Carried out by	HES-SO
Carried out in	April 2011
Published in	CLEF 2011 conference
ImageCLEF years	2003-2009

ImageCLEF publications

ImageCLEF papers in CLEF working notes	
ImageCLEF papers in CLEF proceedings	√
Papers describing ImageCLEF resources	√
Papers using ImageCLEF resources	
#papers	249

ImageCLEF publications and citations

	Year	CLEF proceedings			ImageCLEF resources			All		
		papers	citations	h-index	papers	citations	h-index	papers	citations	h-index
Scopus	2004	5	13	2	4	31	3	9	44	4
	2005	20	50	4	–	–	–	20	50	4
	2006	25	24	3	3	28	1	28	52	3
	2007	27	25	2	6	29	2	33	54	3
	2008	29	18	3	5	22	2	34	40	3
	2009	45	14	2	2	4	1	47	18	2
	2010	44	38	4	11	7	2	55	45	4
	Total	195	182	6	31	121	5	226	303	9
Google Scholar	2004	5	65	3	5	105	4	10	170	6
	2005	20	210	8	5	47	4	25	257	10
	2006	25	247	7	8	144	5	33	391	9
	2007	27	259	7	10	76	4	37	335	9
	2008	29	249	7	7	73	5	36	322	9
	2009	45	284	7	7	53	4	52	337	9
	2010	44	259	7	12	76	6	56	335	10
	Total	195	1573	18	54	574	13	249	2147	22

ImageCLEF scholarly impact: main findings

- 8.62 cites per paper on average
- Citations of **overview vs. participants'** papers
 - overviews : 15% of all publications, but attract around 50% of citations
 - 90% of the papers that have #citations \geq h-index are overviews
- Citations of **general vs. medical** images papers
 - publications in medical domain have had slightly higher impact
 - particularly during 2006-2008
- Citations **per task**
 - Peak in second or third year of operation
 - Followed by a decline – unless there is a major overhaul of the task
 - Tasks with greatest impact so far: photo retrieval, medical tasks
- ~70% of citations originating from papers not in CLEF proceedings

ImageCLEF scholarly impact: extended study

	Preliminary study	Extended study
Carried out by	HES-SO	HES-SO Royal School of LIS
Carried out in	April 2011	January 2012
Published in	CLEF 2011 conference	in preparation
ImageCLEF years	2003-2009	2003-2011

	ImageCLEF publications	
ImageCLEF papers in CLEF working notes		√
ImageCLEF papers in CLEF proceedings	√	√
Papers describing ImageCLEF resources	√	√
Papers using ImageCLEF resources		√
	#papers	249
		~ 1,000

Events in other forums

- **ImageCLEF @ ICPR 2010 contests**
 - 3 contests: photo annotation, robot vision, information fusion
 - ImageCLEF 2009 datasets
 - photo annotation, robot vision: more complex than in ImageCLEF
 - 90 registered groups, 30 submitting results
 - 4th contest: interactive image retrieval
 - not enough participants
- **MCBR-CDS workshops @ MICCAI 2009, 2011**
 - Medical Content-based Retrieval for **Clinical Decision Support** @ Intl. Conf. on Medical Image Computing & Computer Assisted Intervention
 - ImageCLEF 2009, 2011 medical image retrieval datasets
 - ~10 groups used the datasets
 - Need for standardised datasets & evaluation methodology

ImageCLEF 2012

- **Photo annotation and retrieval**
 - Training on manually annotated (reliable) samples
 - Training on automatically generated (noisy) samples
 - Concept-based retrieval
- **Medical image classification and retrieval**
 - Modality classification
 - Move closer to clinical routine: focus on case-based retrieval
 - Collaboration with National Library of Medicine (USA)
- **Robot vision**
 - Multimodal place recognition
 - Perspective camera and kinect mounted on mobile robot
- **Plant identification**
 - More plant species
 - Mobile phone pictures

Future directions

- Research **challenges**
 - Combination of media
 - Scalability
- **Scientific multimedia data**
 - Botanical, bio-imaging, earth observation, etc.
- **Continuous** evaluation
- **Component-based** evaluation
 - Groups can concentrate on the area of their expertise
- Interactive retrieval
- Automation of ground truth generation

Conclusions

- ImageCLEF evaluation campaign
 - Entering its 10th year with evidence of important impact
 - Bringing together researchers from multiple diverse communities
 - Various, continuously evolving tasks – always open for new suggestions!
 - Technology and benchmarks need to go hand in hand
- Interested in **participating** in ImageCLEF 2012 ?
 - Registration opens in February, closes in May
 - Datasets released in March-April
 - Runs submitted in late May-early June
 - Results released in mid July
- Interested in **proposing a task** for ImageCLEF 2013?
 - Please contact us (henning.mueller@hevs.ch, theodora.tsikrika@acm.org)
 - Task and datasets need to be fully defined by October 2012

Thank you!

**Image
CLEF**



<http://www.imageclef.org>