



# FP6 project VITALAS

Arjen P. de Vries

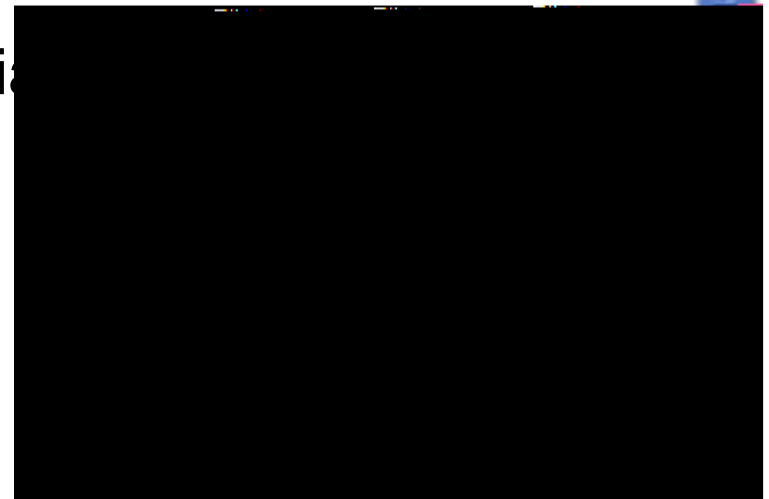
arjen@acm.org





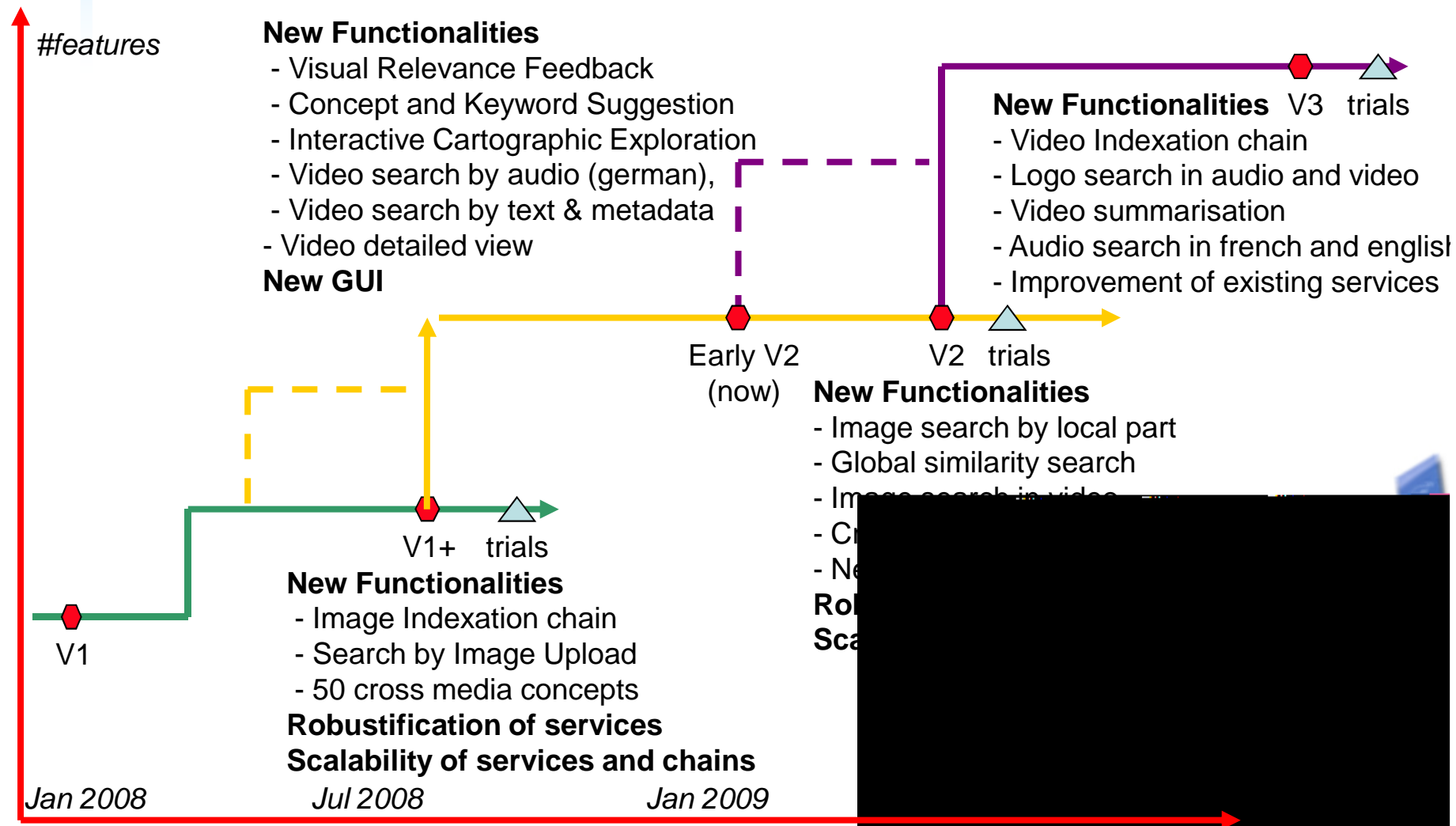
Information Society  
Technologies

- Three characteristics:
  - Use-case driven
    - Professional users
  - Larg-*ish* collections
    - Target: 10M images, 10k hours of video
  - Cross-media
    - Target: 1K – 3K ‘cross-media



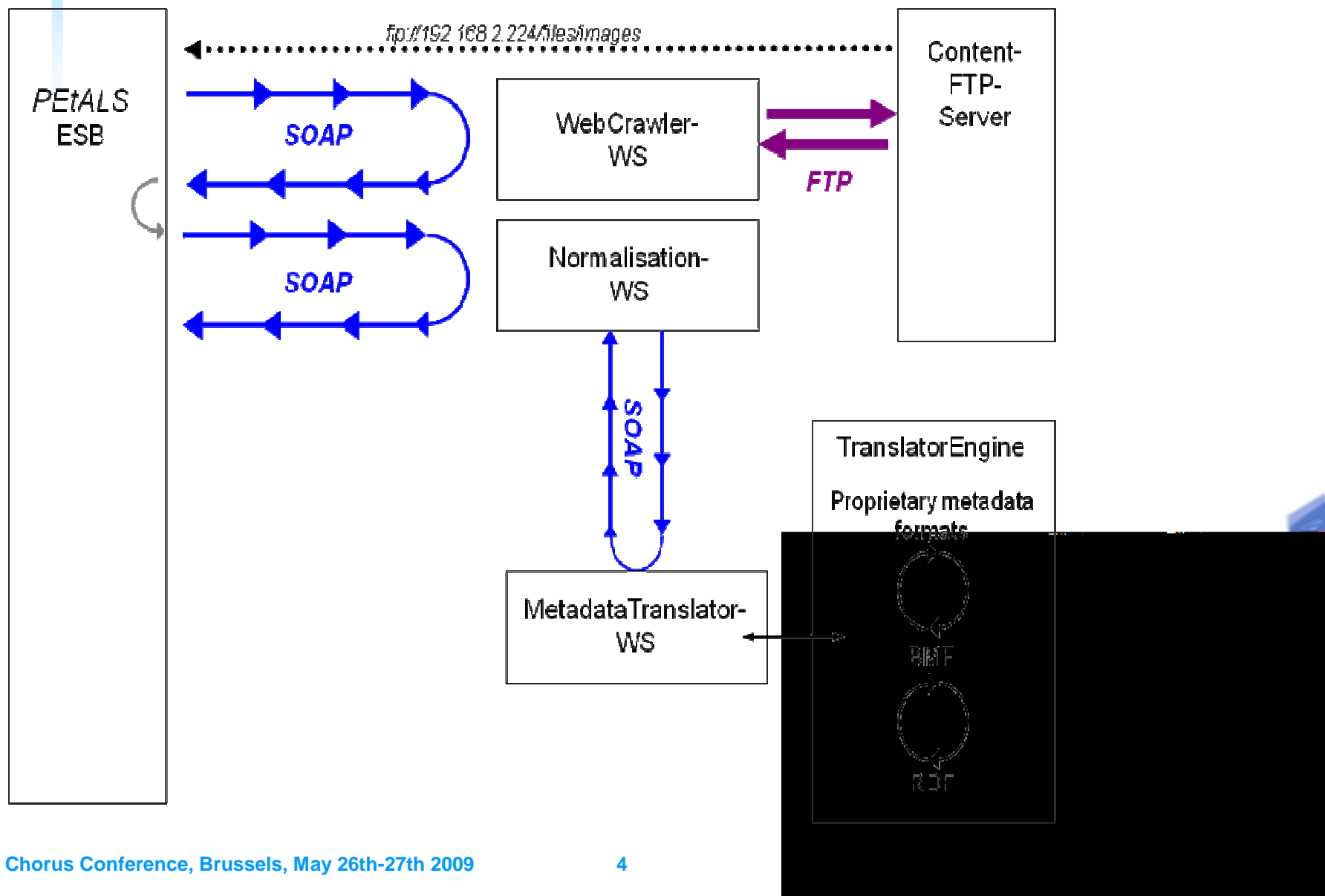


# Integration Project





# Metadata Translator





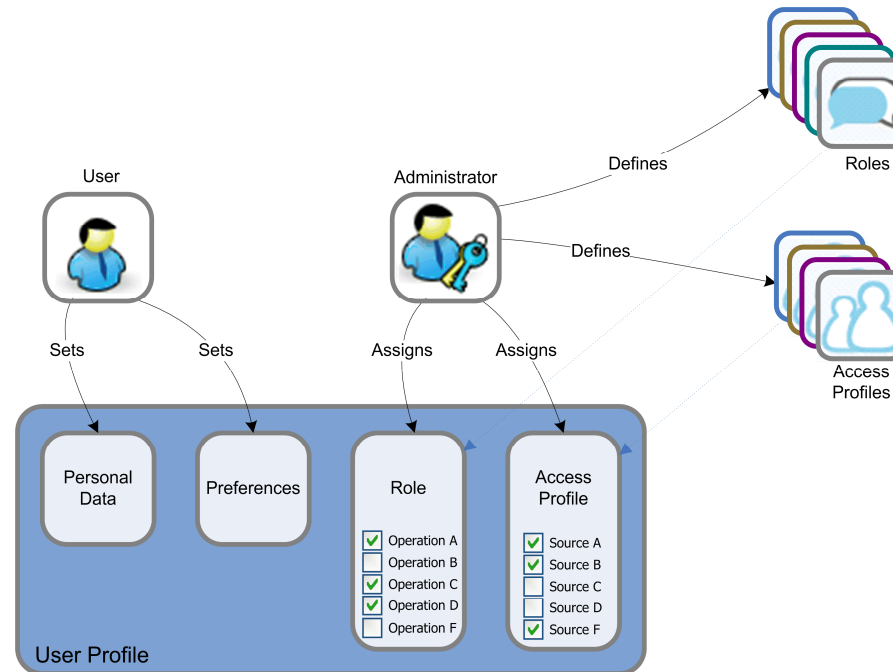
# Personalization, access rights, ...

## User sets:

- Personal Data
- Preferences

## Administrator establishes:

- Role (rights to perform system operations)
- Access Profile (rights to access data sources)



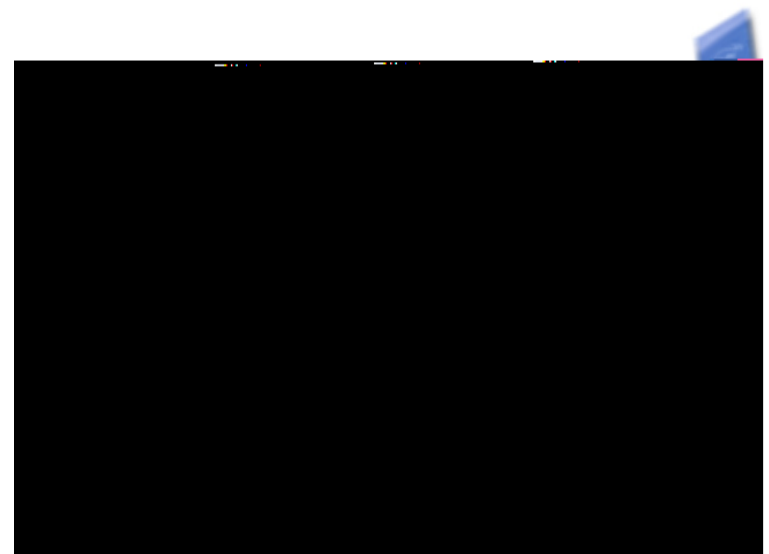


Information Society  
Technologies

# Disclaimer



$(15\text{min}-4\text{min})/11 \text{ partners} \approx 1 \text{ min} /$   
partner





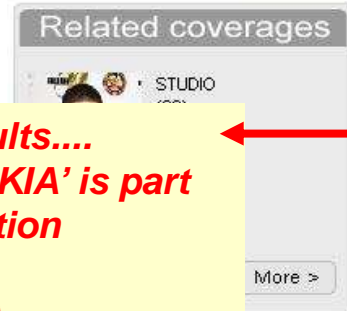
# Use-case driven?!

**QUERY:**  
**'KIA'**

belga picture EDITORIAL • ALL >> news sports entertainment | creative

7 items in my order basket Monday, Feb 18 2008

home search galleries coverages selected coverages download history welcome Tom Wuytack logout change preferences



## Picture search results

Save as default search Page 1 of 16 - 377 picture(s) found | Go to page 1



Kia Shine  
Common performs at Studio 54's  
After-Fight Party hosted by Winky Wright  
Las Vegas, Nevada - 16.02.08  
Credit: (Mandatory): Judy Eddy / WENN

**Unwanted results....  
Just because 'KIA' is part  
of the title/caption**

**e.g. SLOVAKIA**



8287257 / ARZKQM  
14/01/2008 15:45  
BELGA/AFP  
US-DETROIT AUTO  
SHOW-FORD-LINCOLN

8287250 / ARZKQF  
14/01/2008 15:44  
BELGA/AFP  
US-DETROIT AUTO  
SHOW-FORD-LINCOLN MKT

8287170 / ARZKND  
14/01/2008 15:37  
BELGA/AFP  
US-DETROIT AUTO SHOW-FORD

8287155 / ARZKMO  
14/01/2008 15:36  
BELGA/AFP  
US-DETROIT AUTO SHOW-FORD





Information Society Technologies



QUERY: 'KIA' AND 'LOGO'

Picture search results

Save as default search

Page 1 of 1 - 10 picture(s) found | Go to page 1











Related coverages



BELGIUM SOCCER  
JL DAY11  
BRUSSELS  
CHARLEROI  
(2)

More >

Results seem correct but.. Only 10 pictures found

 8309375 / ASARJE 17/01/2008 19:30 BELGA/AFP US-DETROIT AUTO SHOW-KIA	 7786539 / AQWXYC 02/11/2007 12:56 BELGA BELGIUM SOCCER JL DAY11 BRUSSELS CHARLEROI	 6982279 / APDEFA 29/06/2007 02:15 MAXPPP KIA AUTO	 6630146 / AOJDEJ 04/05/2007 08:54 MAXPPP 3294159
 6630141 / AOJDEE 04/05/2007 08:54 MAXPPP 3294158	 6630136 / AOJDDZ 04/05/2007 08:53 MAXPPP 3294155	 6630133 / AOJDDV 04/05/2007 08:53 MAXPPP 3294153	 5266521 / ALJNRH 16/10/2006 22:27 BELGA/AFP US-AUTO SHOW-KIA-LOGO
			





Information Society  
Technologies

# Missing results...

http://app2.belga.be:8086 - Picture details - Mozilla Firefox

Picture details Picture 554 of 872

**Name:** AUSTRALIA TENNIS AUSTRALIAN OPEN

**Caption:** epa00911824 Kim Clijsters of Belgium jubilates after winning against Martina Hingis of Switzerland



Done

http://picture.belga.be - Picture details - Mozilla Firefox

Picture details Picture 41 of 218

**Name:** SOCCER-D1-FC BRUSSELS-PRESENTATION

**Caption:** 20060717 - BRUSSELS, BELGIUM : Michael Jonckheere poses for a portrait picture, Monday 17 July 2006, during the official presentation of FC Brussels soccer team for the 2006-2007 season. BELGA PHOTO

**Keywords:** FOOTBALL

**Date:** 17/07/2006 22:19

**Credit:** BELGA

**Source:** BELGA

**Dimensions:** 2048 x 3072

**Photo:** -

**Category:** SPO

**Sub Category:** SOCCER

**City:** BRUSSELS

**Country:** BELGIUM

**ID:** 4627646

**Theme:** sports

**Exported to:**

**Coverages:**

**Galleries:**



Done



# Scalable Sub-image Search

- Improved access structure PMH V2
  - Random projections to reduce dimensionality, and compressed signatures to further reduce memory resource consumption
  - Real-time Global similarity search on 20 M images and more ...
  - Real-time Local search in 1 (150M local feature descrip



Information Society  
Technologies



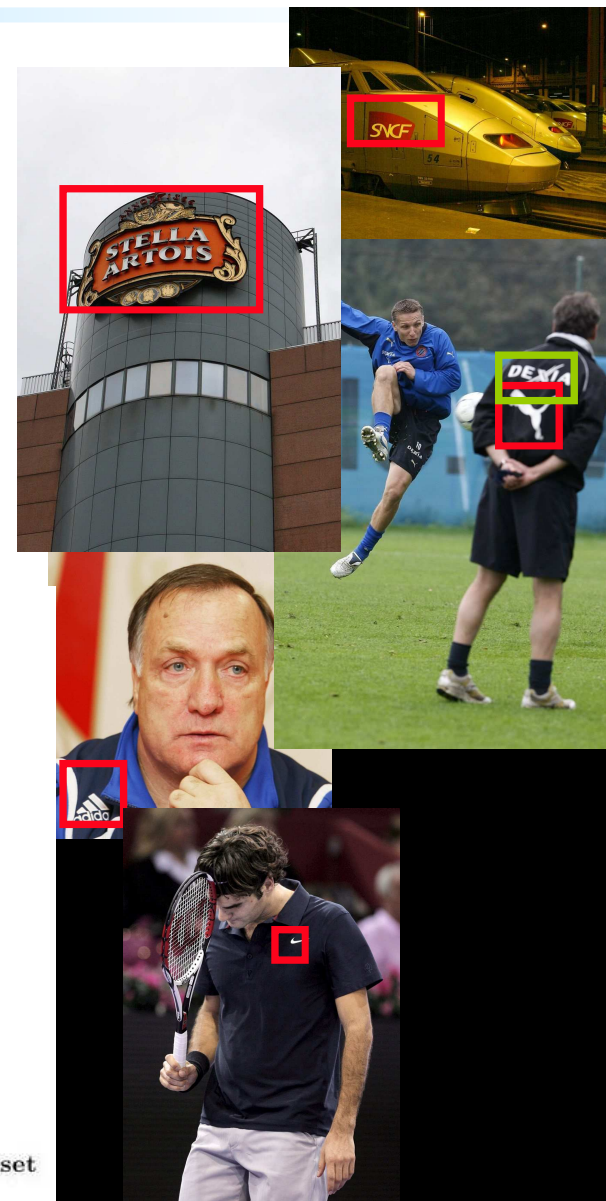
# Scalable Sub-image Search

- Logos ground truth generation
  - Creation of a new challenging dataset (INRIA)
  - 10,000 BELGA images
  - Annotated with 26 logos (> 10 pixels)
  - 55 internal queries and 24 external queries
  - Made available to third parties for research usage

Logo name	Illustration	nb of images
Addidas		114
Addidas-text	<b>adidas</b>	47
Base		408
Bouigues		11
Citroen		46
Citroen-text	<b>CITROËN</b>	97
CocaCola		32
Cofidis		17
Dexia		494
Standard Liege		372
Eleclerc		8
Ferrari		76
Gucci		2
Kia		82

Mercedes		73
Nike		116
Peugeot		5
US President		14
Puma		54
Puma-text	<b>PUMA</b>	22
Quick		56
Roche		2
SNCF		6
StellaArtois		15
TNT		66
VRT		9

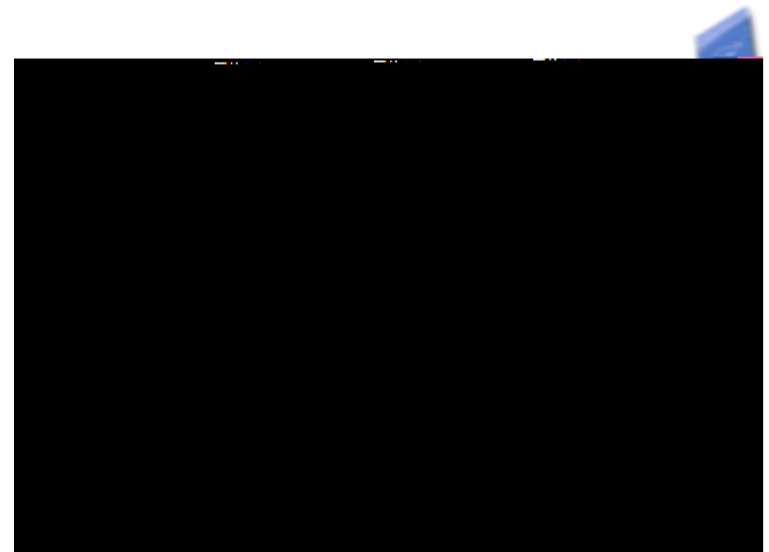
Table 1: Annotated logos in BelgaLogos dataset





# Cross-media Concepts

- Scalability issues for 1K – 3K concepts:
  - How to define the lexicon?
  - How to use the lexicon?
  - How to obtain training data for supervised ML?

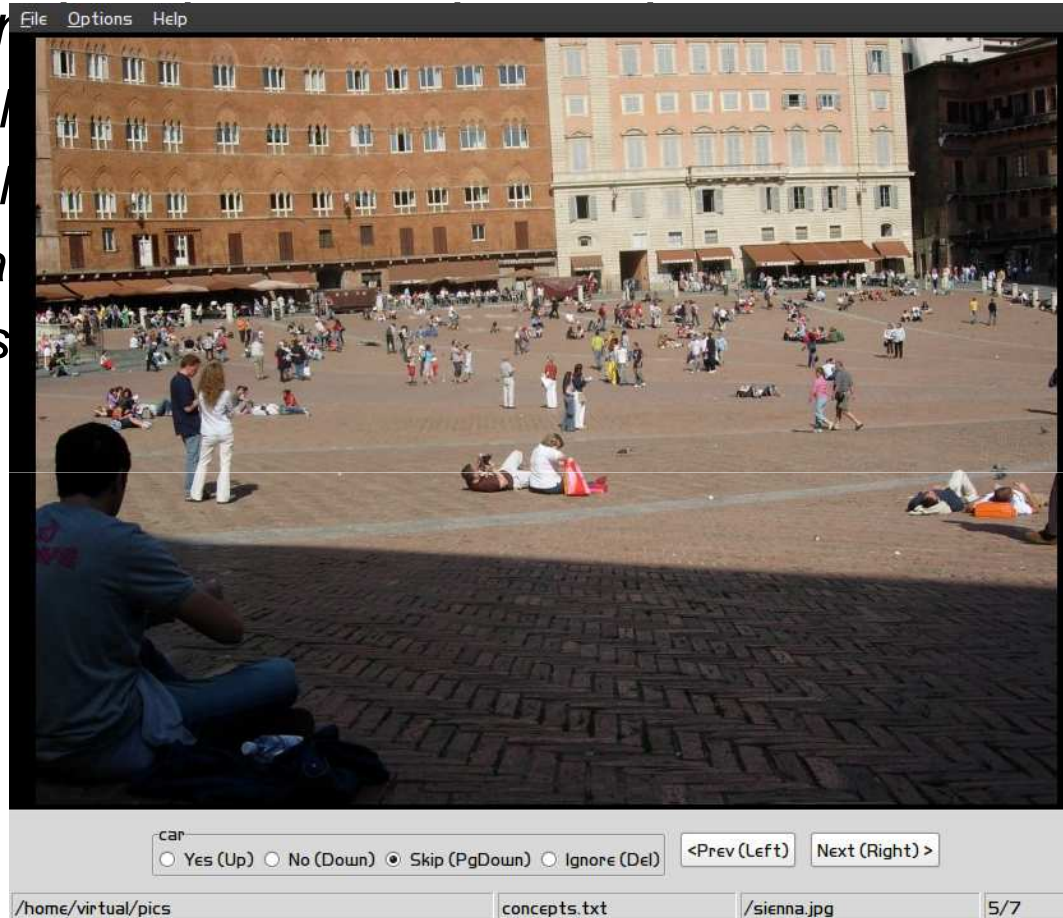






# Cross-media Concepts

- *Group*
- *Se*
- *Se*
- *Ma*
- *Dis*



*e/concept*



Information Society  
Technologies



# How to define the lexicon?

- Candidate concepts are selected semi-automatically
  - Derived from Belga captions and/or logs
  - Statistical test with Europarl as reference corpus





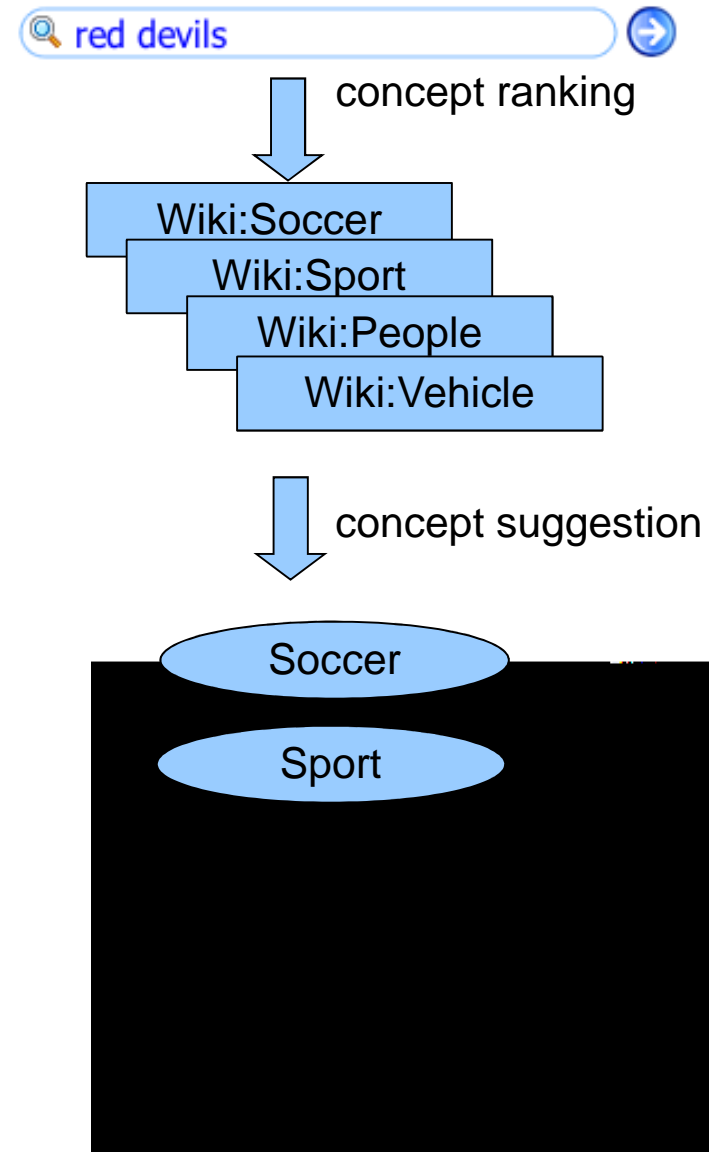
# Automatically-selected concept vocabulary

- soccer (soccer, soccers, soccere): 27278.65
- pictur (picture, pictures, pictured, picturing, pictur, picturs, pictureds): 22565.06
- minist (minister, ministers, ministe, minist, ministeer, ministes, ministered): 21035.98
- team (team, teams, teamed, teaming): 18273.36
- cup (cup, cups, cupping, cupped): 18187.04
- citi (city, cities, citi, citys, citis, citie): 17369.72
- leagu (league, leagues, leagu, leagus): 16693.42
- celebr (celebrates, celebrate, celebrations, celebration, celebrated, celebrating, celebrities, celebrity, celebrants, celebrer, celebre, celebrators, celebrateing, celebreated, celebreates, celebres, celebrational, celebral, celebratings, celebrant): 16545.52
- won (won, wons): 16436.22
- championship (championships, championship): 16256.48



# How to use the lexicon?

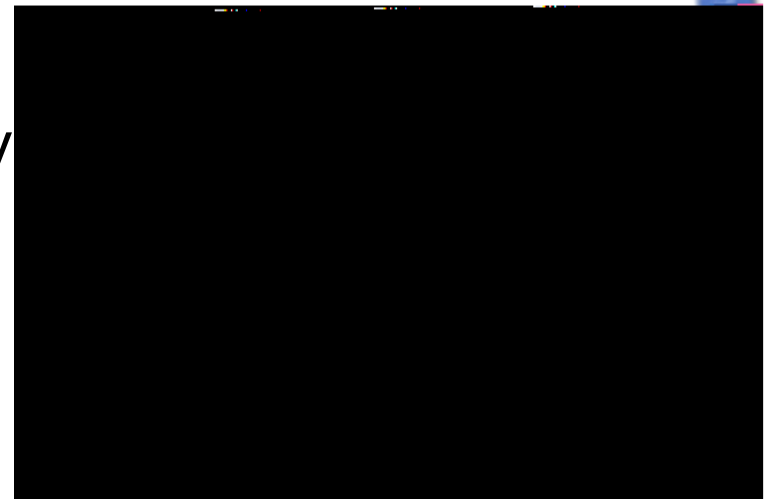
- Wikipedia-based Concept Suggestion:
  - Each Vitalas concept is associated with a Wikipedia article
  - Rank concept-related articles by keyword query
  - Suggest top-k concepts





# How to obtain training data?

- Can we use click-through data instead of manually labelled samples?
  - Advantages:
    - Large quantities, no user intervention, collective assessments
  - Disadvantages:
    - Noisy & sparse
    - Queries not based on strict v



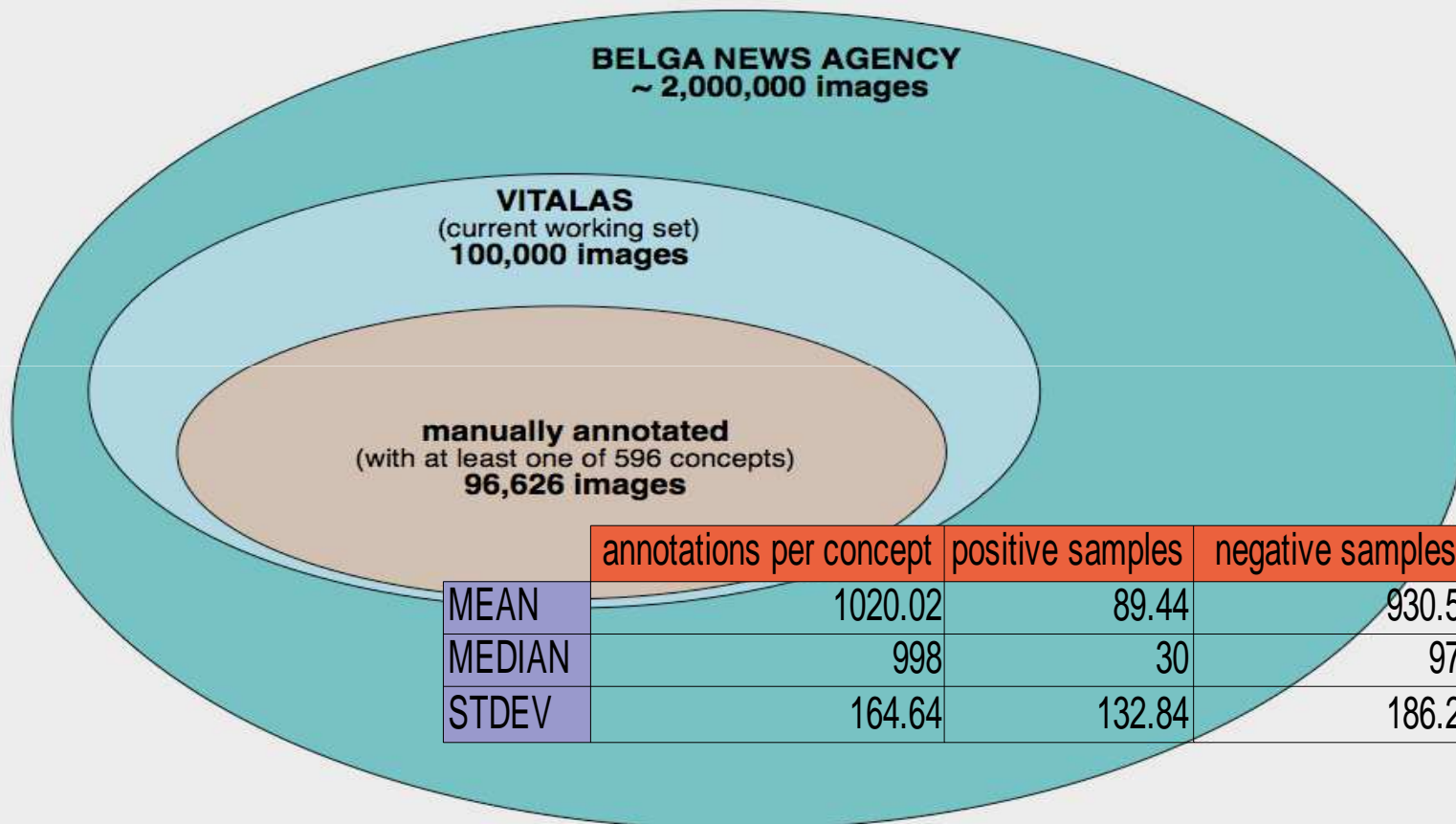


Information Society  
Technologies

# VITALAS manual annotations



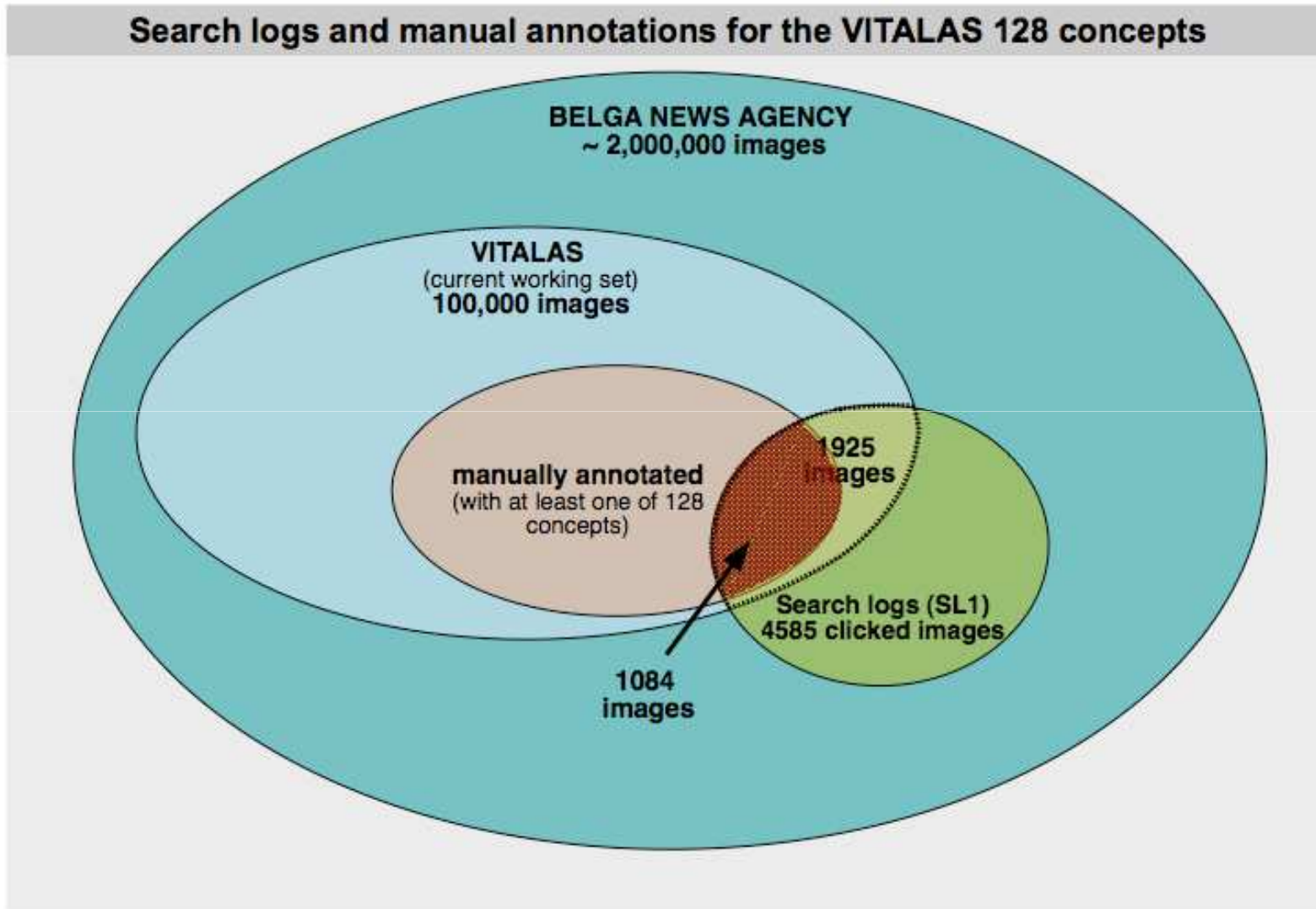
## Manual annotations for the VITALAS 596 concepts



	annotations per concept	positive samples	negative samples
MEAN	1020.02	89.44	930.57
MEDIAN	998	30	970
STDEV	164.64	132.84	186.21



# VITALAS manual annotations vs. search logs





# Automatic Image Annotation



- Research questions:
  - How to annotate images with concepts using click-through data?
  - How reliable are click-through data based annotations?
  - What is the effectiveness of these annotations as training samples for con





# How to annotate?

- Use queries for which images were clicked
  - Inherent noise: gap between queries/captions and concepts
    - queries describe the **content+context** of images to be retrieved
    - clicked images retrieved using their captions: **content+context**
    - concept-based annotations: based on visual **content**-only criteria
  - Sparsity: **only cover part of** previously accessed
  - Mismatch between terms in descriptions and queries



Information Society  
Technologies



# Effectiveness

## •Experiment 1 (visual features):

- training: search-log based annotations
- test set for each concept: manual annotations (~1000 images)
- feasibility study: in most cases, AP considerably higher than the prior

Concept $c$	$T_{c,exact}$	$T_{c,LM}$	$T_{c,LMS}$	$T_{c,LMS_{key}}$	$T_{c,LM_{stem}}$	$T_{c,LMS_{stem}}$	$T_{c,LMS_{stem}_{key}}$	$T_{c,clickgraph}$	prior
airplane_flying				0.0589			0.0589		0.0262
airport	0.2269	0.3736	0.3736	0.3637	0.3736	0.3736	0.3637	0.3032	0.2181
anderlecht	0.5920	0.6003	0.6003	0.5496	0.6003	0.6003	0.5501	0.5168	0.3223
athlete				0.3419	0.4547	0.4547	0.4859		0.3968
basketball	0.5172	0.5499	0.5499	0.5473	0.5499	0.5499	0.5521	0.5172	0.3855
building		0.2166	0.2166	0.2166	0.2166	0.2166	0.2166	0.0779	0.1034
club.brugge	0.5353	0.5786	0.5224	0.6082	0.5786	0.5224	0.6056	0.5030	0.4080
crowd				0.3854			0.3854		0.1494
farms				0.0623			0.0677		0.0090
fashion_model			0.7227	0.7227		0.7116	0.7116		0.4333
fire	0.3868	0.4523	0.4523	0.3371	0.4395	0.4395	0.3424	0.3620	0.0972
flood	0.5159	0.5794	0.5794	0.5333	0.5948	0.5948	0.4511	0.4261	0.3627
formula_one	0.4322	0.5478	0.5890	0.7242	0.5813	0.5813	0.7242	0.7257	0.4208
highway				0.3062			0.3223		0.1623
logo								0.5453	0.4322
meadow				0.0162			0.0162		0.0010
rally_motorsport		0.6899	0.6899	0.7197	0.6899	0.6899	0.7018	0.5556	0.2763
red_devils	0.7540	0.7837	0.8046	0.7191	0.7977	0.8046	0.7191	0.6584	0.4624
sky				0.2662			0.2448		0.1454
soccer	0.5694	0.6435	0.6435	0.6475	0.6435	0.6435	0.6475	0.5600	0.4297
stadium		0.3954	0.3954	0.2169	0.3954	0.3954	0.1913		0.1091
team		0.2153	0.2153	0.1905	0.2153	0.2153	0.2153	0.0675	0.0371
tennis	0.4471	0.5044	0.5044	0.5016	0.5044	0.5044	0.5016	0.4588	0.3717
volleyball	0.5561	0.5678	0.5678	0.5119	0.5678	0.5678	0.5119	0.4541	0.3403
war		0.1737	0.1737	0.1737	0.1737	0.1737	0.1737	0.3120	0.2076
<b>MAP</b>	<b>0.5030</b>	<b>0.4920</b>	<b>0.5059</b>	<b>0.4050</b>	<b>0.4928</b>	<b>0.5022</b>	<b>0.4067</b>	<b>0.4402</b>	<b>0.2523</b>



# Example: soccer

manually annotated positive samples

search log based annotated positive samples



test set results



More? See paper published at CIVR 2009!  
Final Chorus Conference, Brussels, May 26th-27th 2009



# Scalable Speech Search

## ■ *Speech Search*

### **Near-Realtime ASR using Multipass Decoding**

- SoA models for German
- Baseline setups for French and English

### **Hybrid Speech Search Approach**

- Parallel word and syllable ASR decoding
- Hybrid retrieval from word and subword lattices
- Advantage: no fixed vocabulary

### **Focus on both Accuracy and Efficiency in Retrieval**

- High precision/recall without need of fixed lexicon
- Enable vocabulary independent search on 10.000 words



# Scalable Speech Search

- *Speech Search Results - Hybrid Spoken Term Detection*
  - Data set: Broadcast News + Conversation Shows (German)
  - Exact hybrid on a simulated 10.000h syllable index: 0.06 seconds
  - Enables vocabulary independent speech search on large archive
  - Several retrieval algorithms available
    - ⇒ Choice depends on precision / recall / efficiency requirements

System Setup	Precision	Recall
Exact syllable transcript search	0,96	0,6
Syllable lattice retrieval	0,91	0,7
Syllable lattice retrieval ( + fuzzy search)	0,80	0,7
Exact syllable-word-hybrid	0,97	0,8





Information Society  
Technologies

# Scalable Speech Search

## Treffer

### Treffer 1

Start: 14:44  
Match: 100%



... die weltweite **finanzkrise** hat die schwächen des kapitalistischen wirtschaftssystems ans nicht gebracht ...

### Treffer 2

Start: 26:23  
Match: 100%



... doch noch retten zu können unter anderem eine firma heidelberg zement war durch die **finanzkrise** erheblichen bedrängnis geraten außerdem hatte merkel bei fehlt spekulationen ...

### Treffer 3

Start: 27:41  
Match: 100%



... die den tod des ihren siebzig jährige bestätigte in dem schreiben heißt es die durch die **finanzkrise** verursachte wirtschaftlichen notlage seiner firmen zu wie die ohnmacht nicht behandeln ...

**Alle Informationen anzeigen**

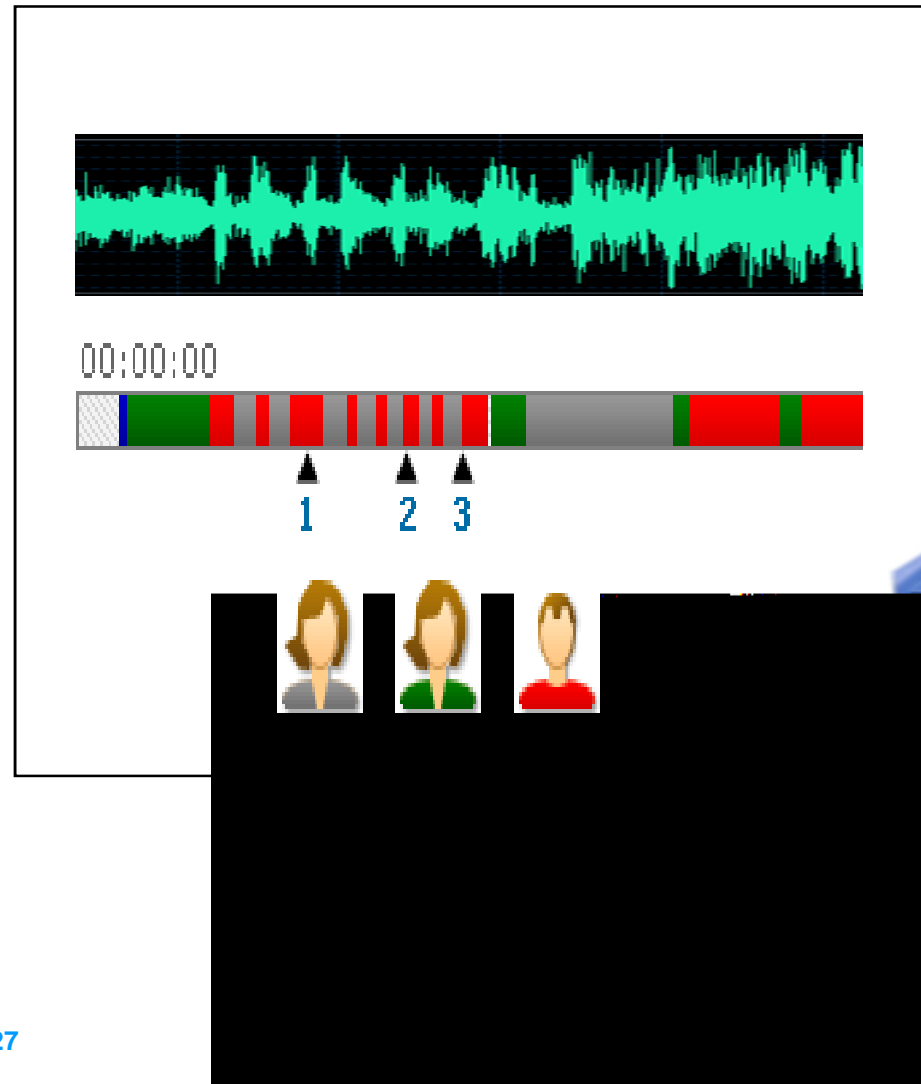






# Audio Structuring

- Separation into homogeneous audio segments
- Detection of pre-defined acoustic scenarios
  - ⇒ *Speech parts, speaker gender*
  - ⇒ *Re-appearing speaker clusters, acoustic concepts*





Information Society  
Technologies

# Scalable Audio Search

Suchen | Browsen

Suche

merkel sarkozy

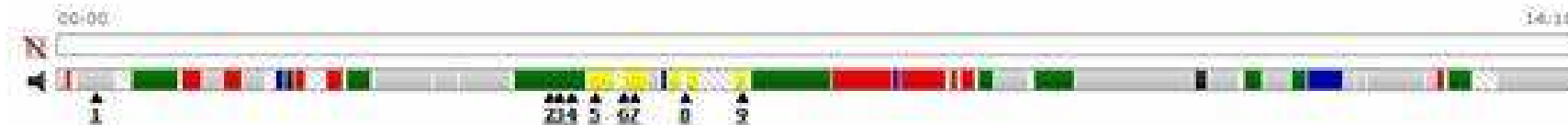
Neue Suche

[Zurück zur Übersicht](#)

Asset 1 von 42: 000501\_h19\_p

[vorheriges](#) [nächstes](#)

Fraunhofer  
Institut  
Intelligente Analyse- und  
Informationssysteme



## Treffer

### Treffer 1

Start: 00:22  
Match: 100%



... zum kampf gegen dumpinglöhne auf bundeskanzlerin **merkel** wird mit dem karlpreis der stadt aachen geehrt ...

### Treffer 2

Start: 04:38  
Match: 100%



... es gab zahlreiche verletzte frankreichs präsident **sarkozy** ließ seinen ganzen scharonis ihnen heute in aachen weiterhin bundeskanzlerin merkel ...

### Treffer 3

Start: 04:43  
Match: 100%



... sarkozy ließ seinen ganzen scharonis ihnen heute in aachen weiterhin bundeskanzlerin **merkel** den internationalen karlpreis für ihren einsetz zur lösung der europäischen ...

### Treffer 4

Start: 04:51  
Match: 100%



... ihren einsetz zur lösung der europäischen verfassungskrise in seiner weder lobte **sarkozy** die kanzlerin in den nächsten film so als wolle er alle bericht da über ein kübler ...

### Treffer 5

Start: 05:04  
Match: 100%



... es war so voll wie selten vor dem aachener rathaus kein wunder bundeskanzlerin **merkel** ist die diesjährige karlpreis kägerin und der französische staatspräsident sagte sie der laudator ...

### Treffer 6

Start: 05:20  
Match: 100%



... ich liebe angela **merkel** mehr als manche sagen ich habe viel von ihr die wähler hat die angela mit der ...

### Treffer 7

Start: 05:26  
Match: 100%



... manche sagen ich habe viel von ihr die wähler hat die angela mit der angela **merkel** sei es zu verdanken dass die europäische union doch noch einen weg aus der verfassungskrise gefunden ...

### Treffer 8

Start: 05:35  
Match: 100%



... **merkel** dankte sagte sie dafür dass er frankreich zurück nach europa geführt habe was versprach weiter ...

### Treffer 9

Start: 06:27  
Match: 100%



... ein friedliches miteinander der kulturen überall auf der welt das sei ihre vision so **merkel** der heutige tag der nach könnte ein guter anfang ...





# Automatic Detection of Repeated Objects

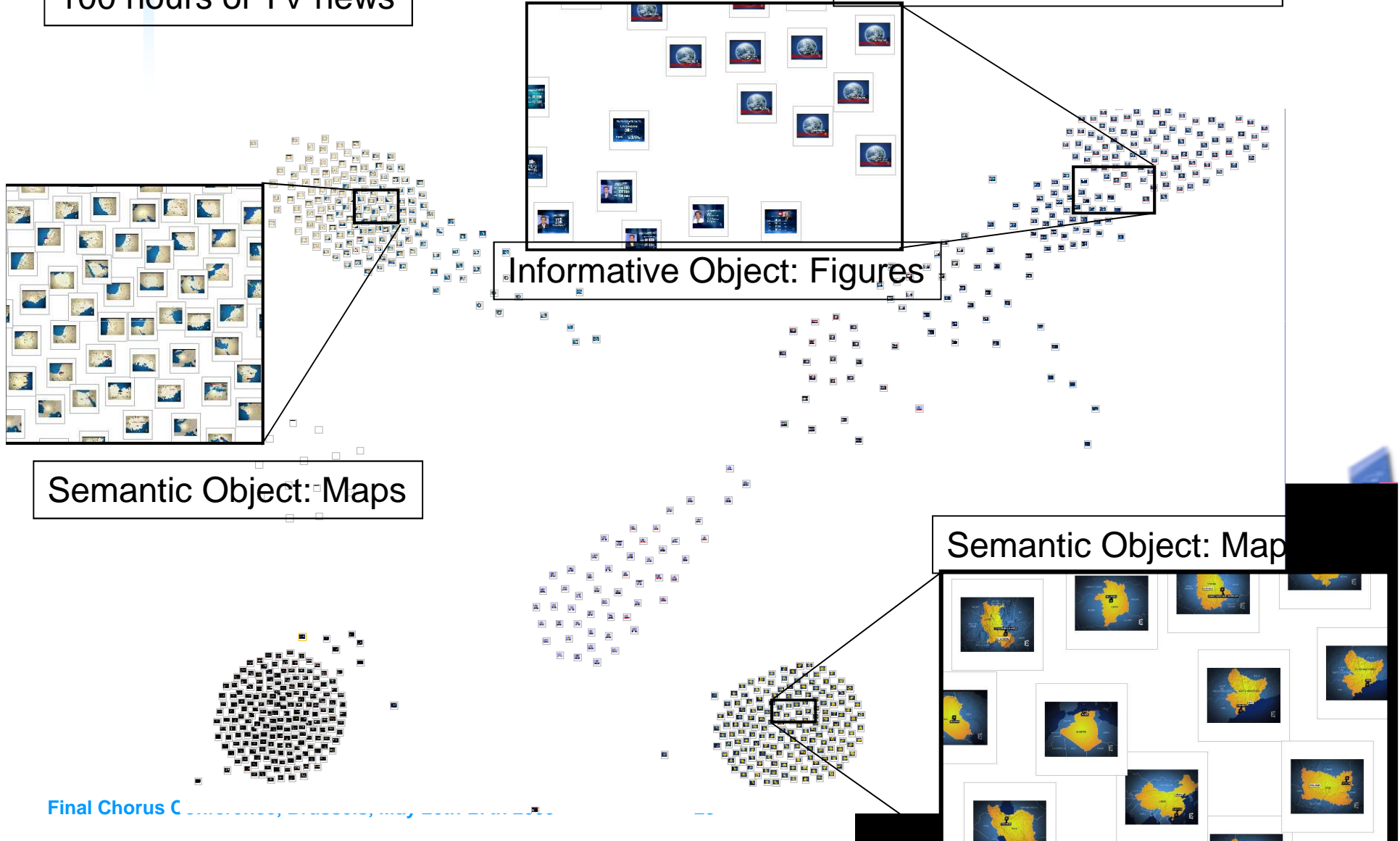
100 hours of TV news

Structuring Object: Generic

Informative Object: Figures

Semantic Object: Maps

Semantic Object: Map

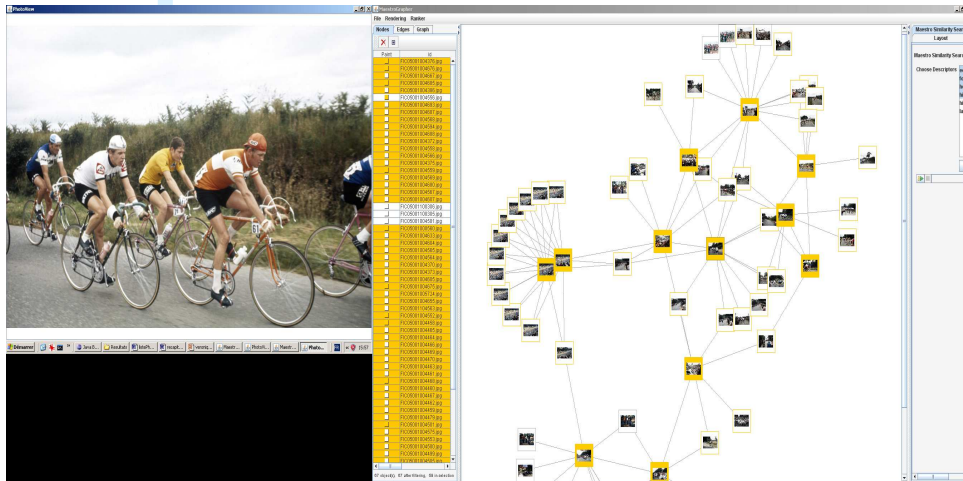




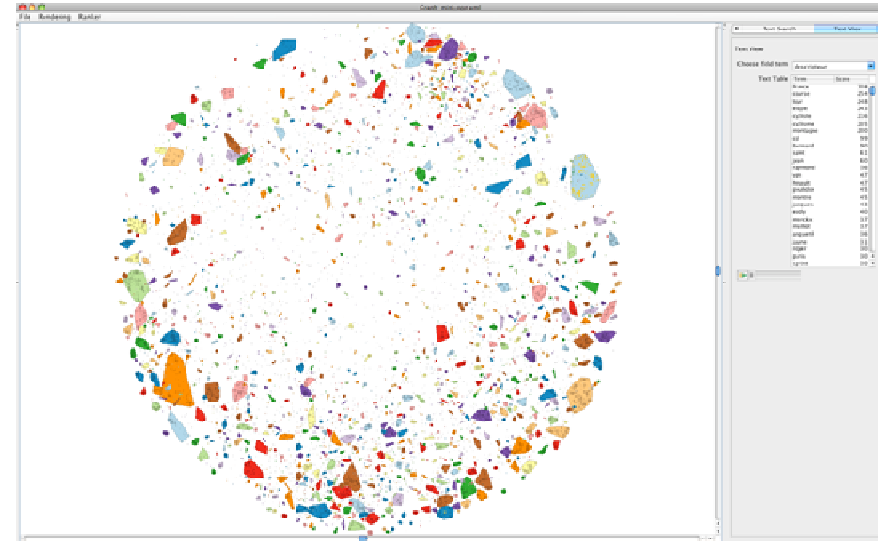
Information Society  
Technologies



# Interactive Exploration



Local Interactive Cartographic Explorer



Global Interactive Cartographic Explorer

*Presented at CIVR'09*



Information Society  
Technologies

# Browse 100K Belga documents

Global map: the colored clusters are built from textual similarities  
Documents selected by search functionalities are highlighted

Global map  
and zoom area

Text or image  
search

Labels of the  
Cluster  
pointed  
by the mouse

The screenshot shows the Vitalas V2 interface. On the left, there is a sidebar with an 'Overview' section showing a circular global map and an 'Image' section showing a photo of a road with a crack. Below the image is a 'text' section with the text 'Cracks in the ground in Bengkulu'. On the right, there is a search panel with a 'Text search' field containing 'earthquake', a 'Search' button, and a 'Term frequency' table. The main area is a large map of colored clusters representing document similarities.

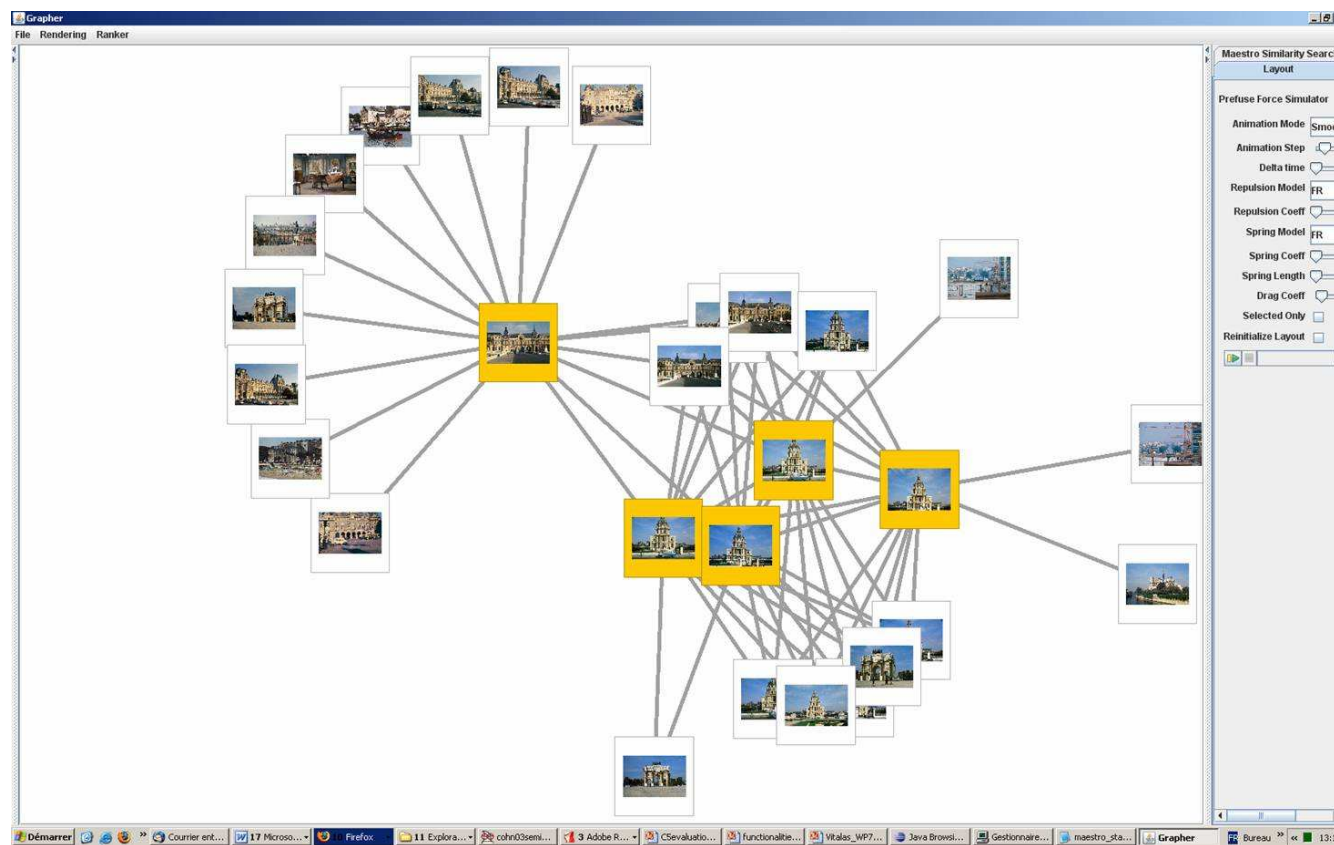
Term	Score
earthqua...	58
indonesian	22
indonesia	20
pakistan	16
house	13
powerful	12
ruins	11
earthqua...	11
aid	10
field	10
hospital	9
treated	8
24	7
truck	7
jolted	7
unload	7
hours	7
goods	7
relief	6
main	6

Image and text of the document pointed by the mouse



# Navigate by textual or visual proximity

Local map: local expansion based on image, textual or hybride proximity  
Documents close from each others are connected



Modality  
setting





# (First) User Evaluations

- *Search with VITALAS integrated system - Corpus: 100K images (BELGA)*
  - *Users were very enthusiastic about the possibilities offered by the advanced functionalities*
  - *More practice needed on new functionalities to perform better*
  - *Response time is not sufficient*
- *Logo search, evaluation - Corpus 10000 images (BELGA)*
  - *Very positive end user feedback*
  - *Our expert would like to extend the use beyond logos*
- *Audio search, evaluation - Corpus: 1 year from*
  - *Results: very positive end user feedback*
  - *Timeline representation pertinent*



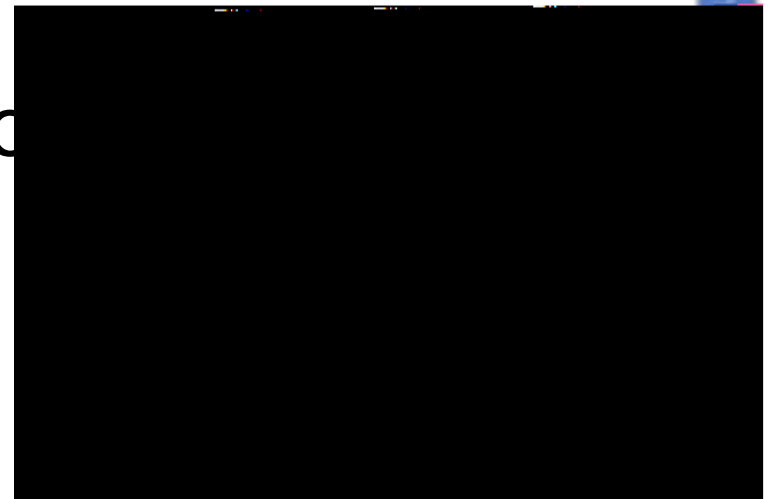
# International Benchmarks

- Organize
  - Wikipedia MM task (from INEX → ImageCLEF)
    - *More Vitalas spin-off... ImageCLEF 2009 uses a 500k Belga collection!*
- Participate
  - TRECVID High level features task (2008, 2009), Search task (2008, 2009)
    - Ambition for 2009: Interactive professional archivists
  - INEX Efficiency, Entity Rank



# Conclusion

- Scalability:
  - Local and global visual search: 20M global features, 100K locals features
  - Vocabulary independent speech search on 10k hours of video
- Competitive cross-media concept detectors
- Overcome training data bottleneck  
log file analysis





# Final year's work

- Main scientific objectives:
  - Scalability
  - Closing the loop between cross-media retrieval and indexing
- Main emphasis in terms of effort:
  - Improve intuitiveness and user control
- Main deliverables:
  - Operational system at user



## Final word...

PF/Tijah

- Many search problems in Vitalas have been approached as XML Retrieval:
  - Concept / term suggestion, concept / meta-data search, query log analysis
  - Features: XQuery integration, incremental indexing, working with arbitrary