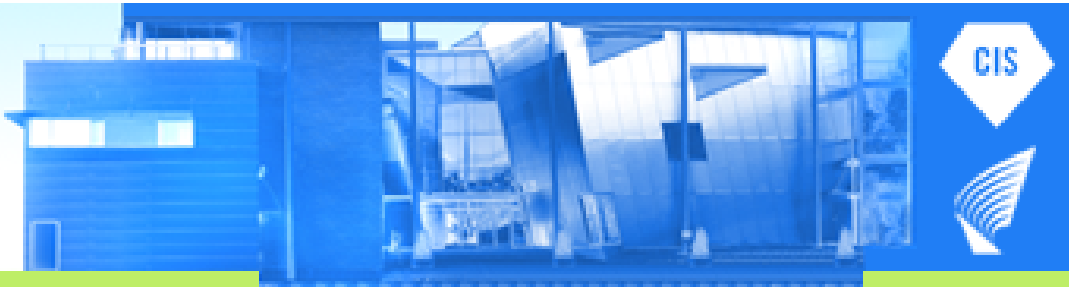
LABORATORY OF COMPUTER
AND INFORMATION SCIENCE
ADAPTIVE INFORMATICS
RESEARCH CENTRE



T-61.6030 Special Course in Computer and Information Science III: Multimedia Retrieval

Introduction

Markus Koskela

January 18, 2008

Overview of the course

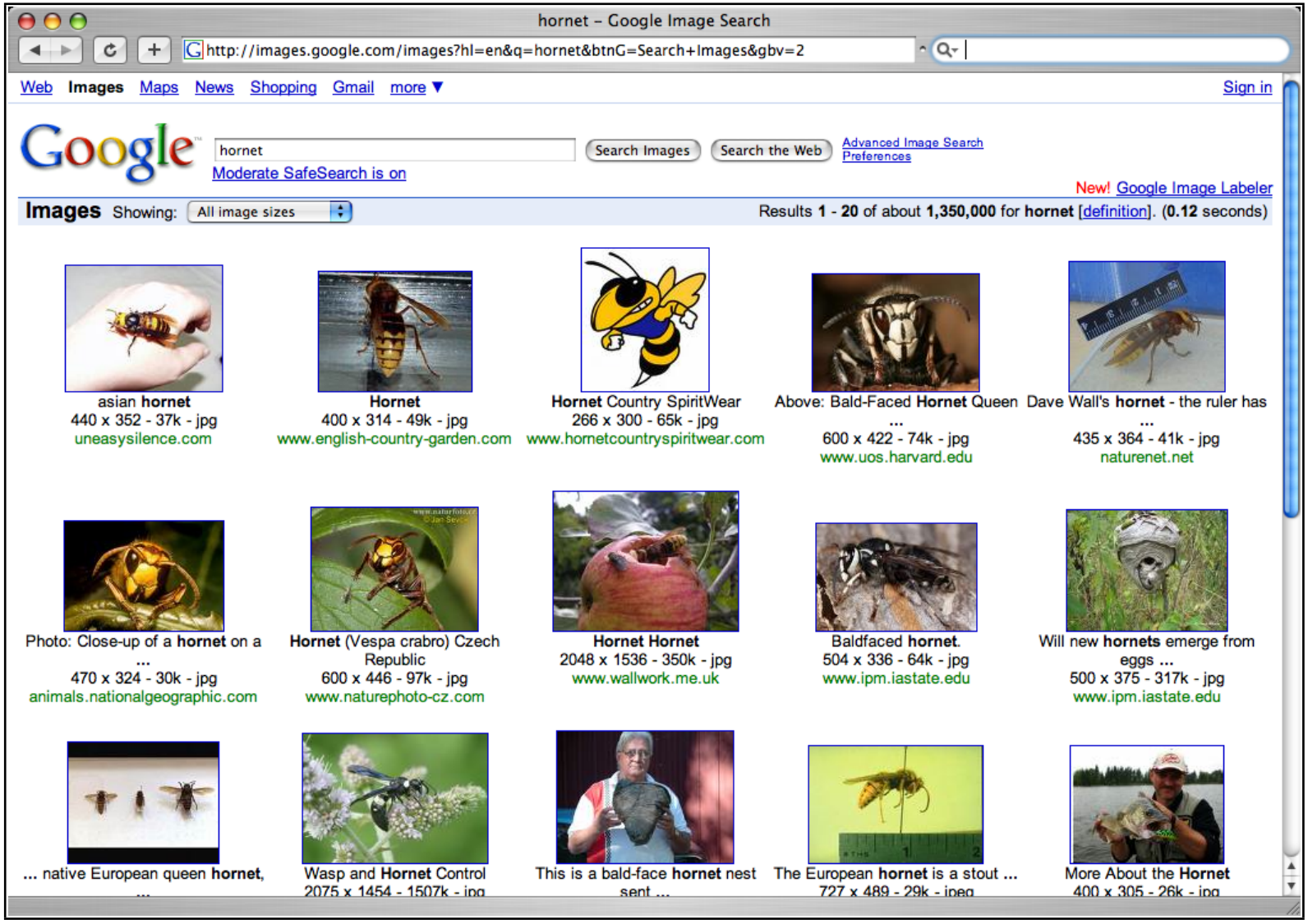
Lecturers	D.Sc. (Tech) Markus Koskela, docent Jorma Laaksonen
Assistant	M.Sc. (Tech) Mats Sjöberg
Credits (ECTS)	5
Semester	Spring 2008 (periods III-IV)
Sessions	On Fridays 12-14 in lecture hall T5
Language	English
Web	http://www.cis.hut.fi/Opinnot/T-61.6030/
Registration	WebTOPI or markus.koskela@tkk.fi

The topic

The **analysis and retrieval** of **multimedia data**.

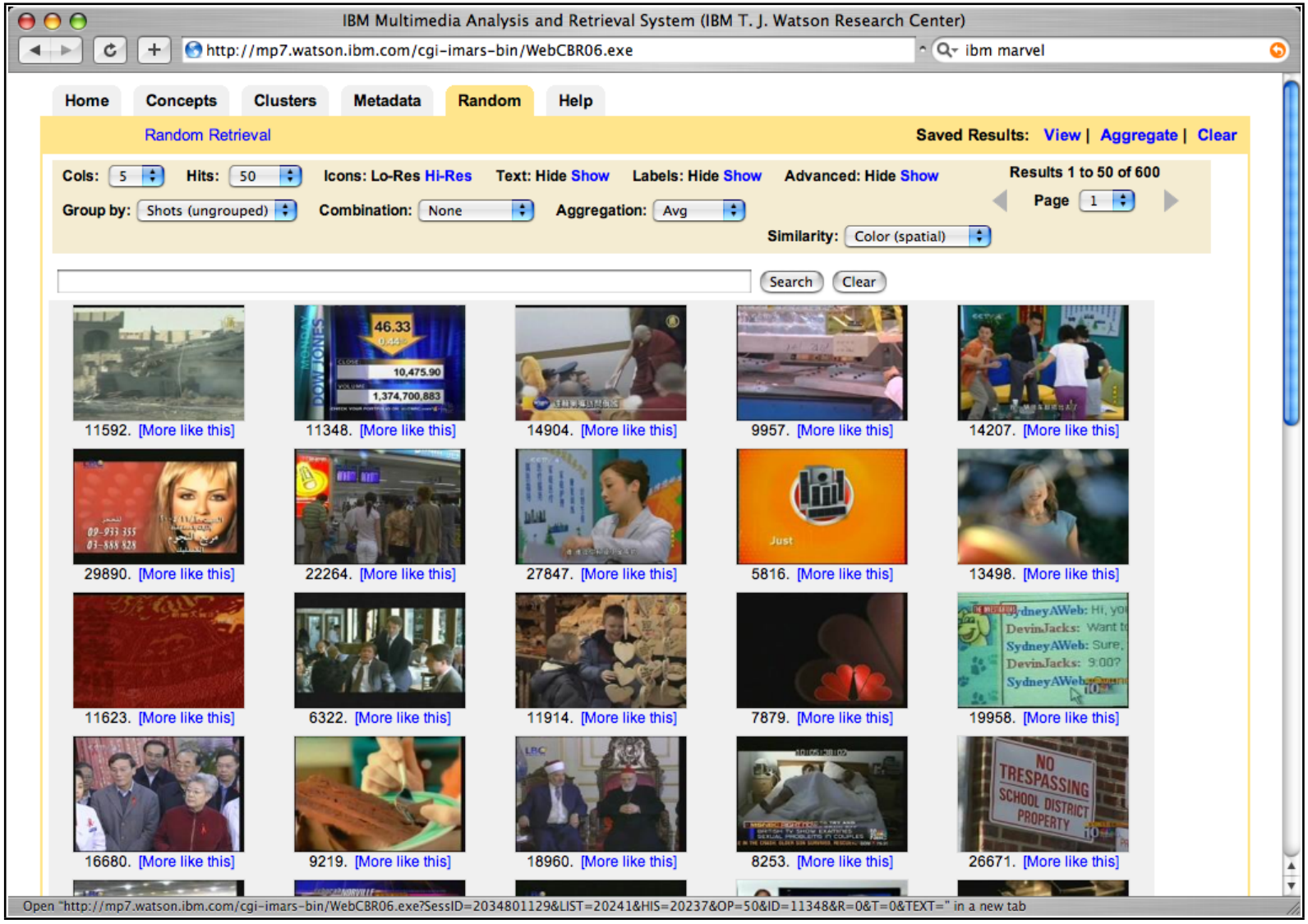
- combines multiple disciplines:
  - information retrieval
  - computer vision, signal processing
  - pattern recognition, machine learning
  - human-computer interaction
- a combination of image, video, audio, textual data

From



Google Image Search, © Google

To



IBM Marvel Search Engine, © IBM

Requirements

- seminar presentation
- project work (small-scale experiments & written report)
- active participation (one absence is allowed)

Material

Multimedia Retrieval by Blanken, H.M.,de Vries, A.P., Blok, H.E., and Feng, L. (Eds.). Springer 2007, 372 p. ISBN: 978-3-540-72894-8.

+ **Machine Learning for Multimedia Content Analysis** by Gong, Y. and Xu, W.. Springer 2007, 277 p. ISBN: 978-0-387-69938-7.

+ **A Unified Framework for Video Summarization, Browsing & Retrieval** by Xiong, Z., Radhakrishnan, R., Divakaran, A., Rui, Y., and Huang T. Elsevier 2006, 296 p. ISBN: 978-0-12-369387-7.

+ relevant journal articles, conference papers

Schedule

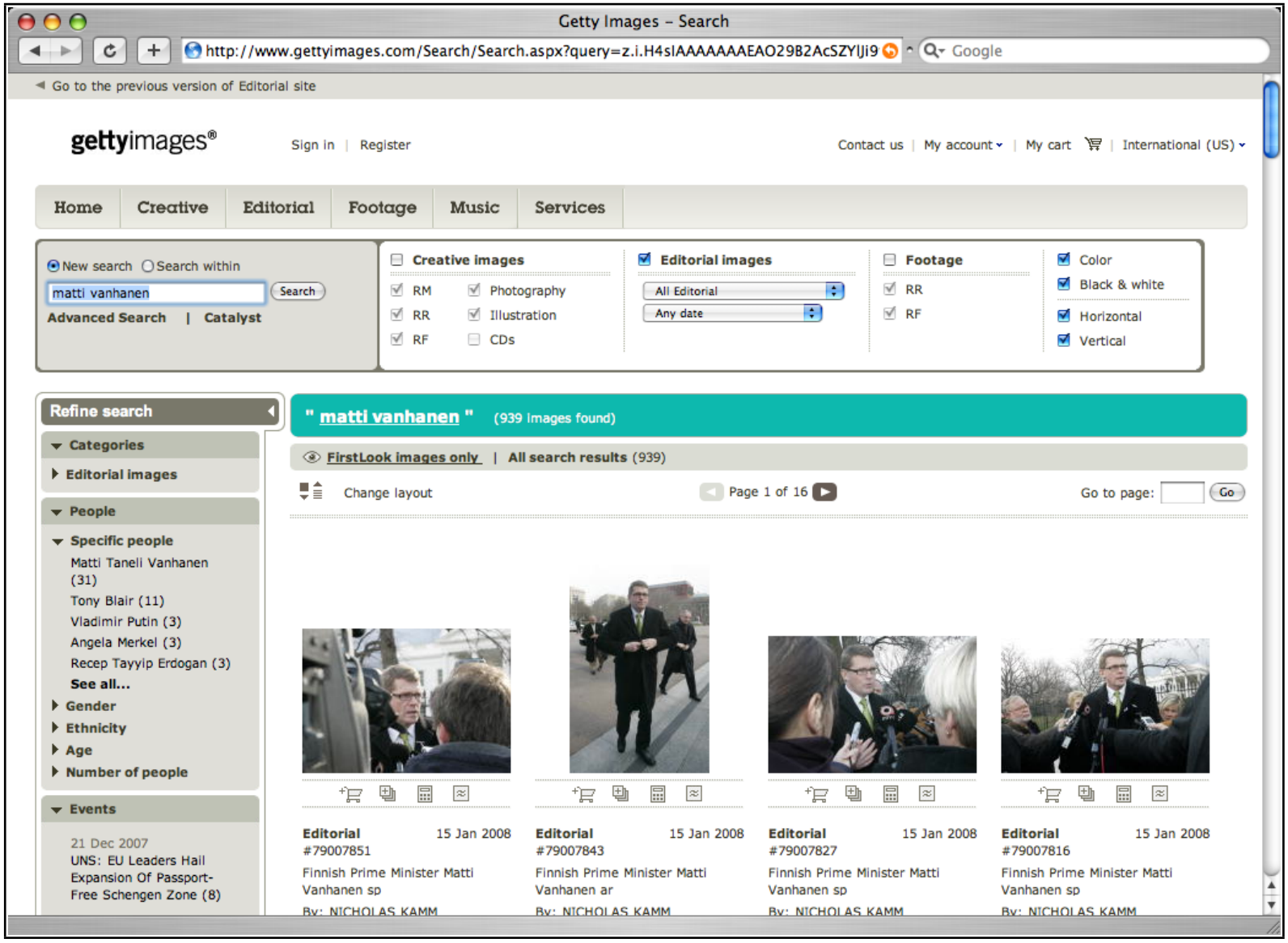
Date	Topic	Material
25.1.	Introduction	MR: Ch. 1
1.2.	Metadata, pattern recognition	MR: Ch. 2-3
8.2.	Text retrieval	MR: Ch. 4
15.2.	Image processing	MR: Ch. 5
22.2.	Generative probabilistic models	MR: Ch. 6
29.2.	Speech indexing	MR: Ch. 7
14.3.	Semantic video indexing	MR: Ch. 8
28.3.	Video highlight extraction	UF*: Ch. 3
4.4.	Video retrieval I	MR: Ch. 9
11.4.	Video retrieval II	MR: Ch. 10
18.4.	Interaction	MR: Ch. 11
25.4.	System evaluation	MR: Ch. 13

*: Xiong et al: A Unified Framework for Video Summarization, Browsing & Retrieval

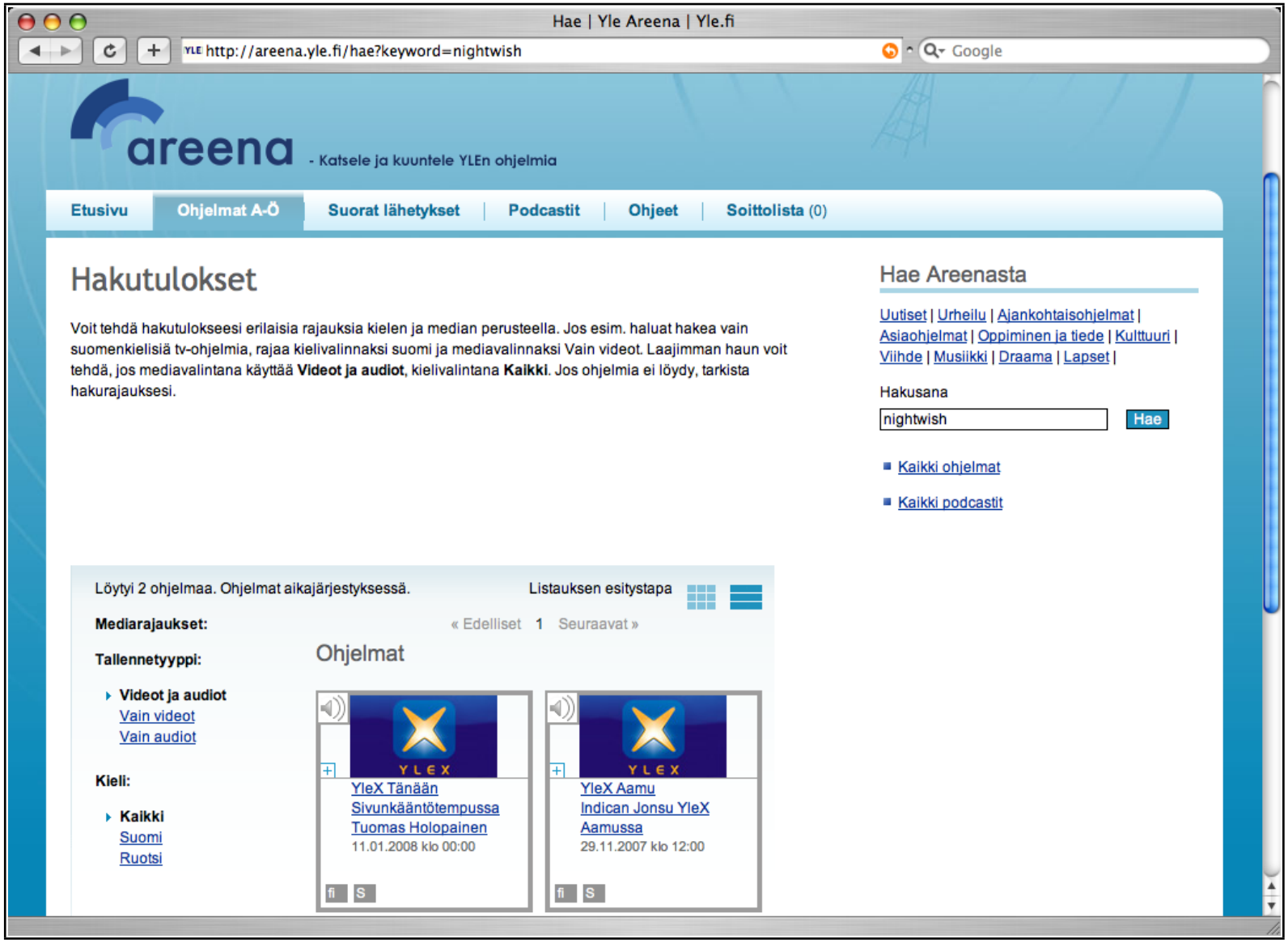
Chapter 1: Introduction

Example usage scenarios

- Journalism, advertising
  - stock photo & video archives
  - "Pretty girl doing something active, sporty in a summery setting, beach - not wearing lycra, exercise clothes - more relaxed in a tee-shirt. Feature is about deodorant so girl should look active - not sweaty but happy, healthy, carefree - nothing too posed or set up - nice and natural looking"
- Digital TV & movies
- Web search
- Medical images
- Surveillance



gettyimages.com, © Getty Images


Hae | Yle Areena | Yle.fi
http://areena.yle.fi/hae?keyword=nightwish
Google
areena
- Katsele ja kuuntele YLEn ohjelmia
Etusivu
Ohjelmat A-Ö
Suorat lähetykset
Podcastit
Ohjeet
Soittolista (0)
Hakutulokset
Voit tehdä hakutulokseesi erilaisia rajauksia kielen ja median perusteella. Jos esim. haluat hakea vain suomenkielisiä tv-ohjelmia, rajaa kielivalinnaksi suomi ja mediavalinnaksi Vain videot. Laajimman haun voit tehdä, jos mediavalintana käyttää Videot ja audiot, kielivalintana Kaikki. Jos ohjelmia ei löydy, tarkista hakurajauksesi.
Hae Areenasta
Uutiset | Urheilu | Ajankohtaisohjelmat | Asiaohjelmat | Oppiminen ja tiede | Kulttuuri | Viihde | Musiikki | Draama | Lapset |
Hakusana
nightwish
Hae
Kaikki ohjelmat
Kaikki podcastit
Löytyi 2 ohjelmaa. Ohjelmat aikajärjestyksessä.
Listauksen esitystapa
Mediarajaukset:
« Edelliset 1 Seuraavat »
Tallennetyyppi:
Videot ja audiot
Vain videot
Vain audiot
Kieli:
Kaikki
Suomi
Ruotsi
Ohjelmat
YLE X
YleX Tänään
Sivunkääntötempussa
Tuomas Holopainen
11.01.2008 klo 00:00
fi S
YLE X
YleX Aamu
Indican Jonsu YleX Aamussa
29.11.2007 klo 12:00
fi S

YLE Areena

Multimedia data

- Originally refers to a collection of media types, nowadays used also for (non-textual) single media types
- Often requires a lot of space for storage, time for processing, bandwidth for transfer
- Effective presentation not trivial either
  - indexing
  - summarization

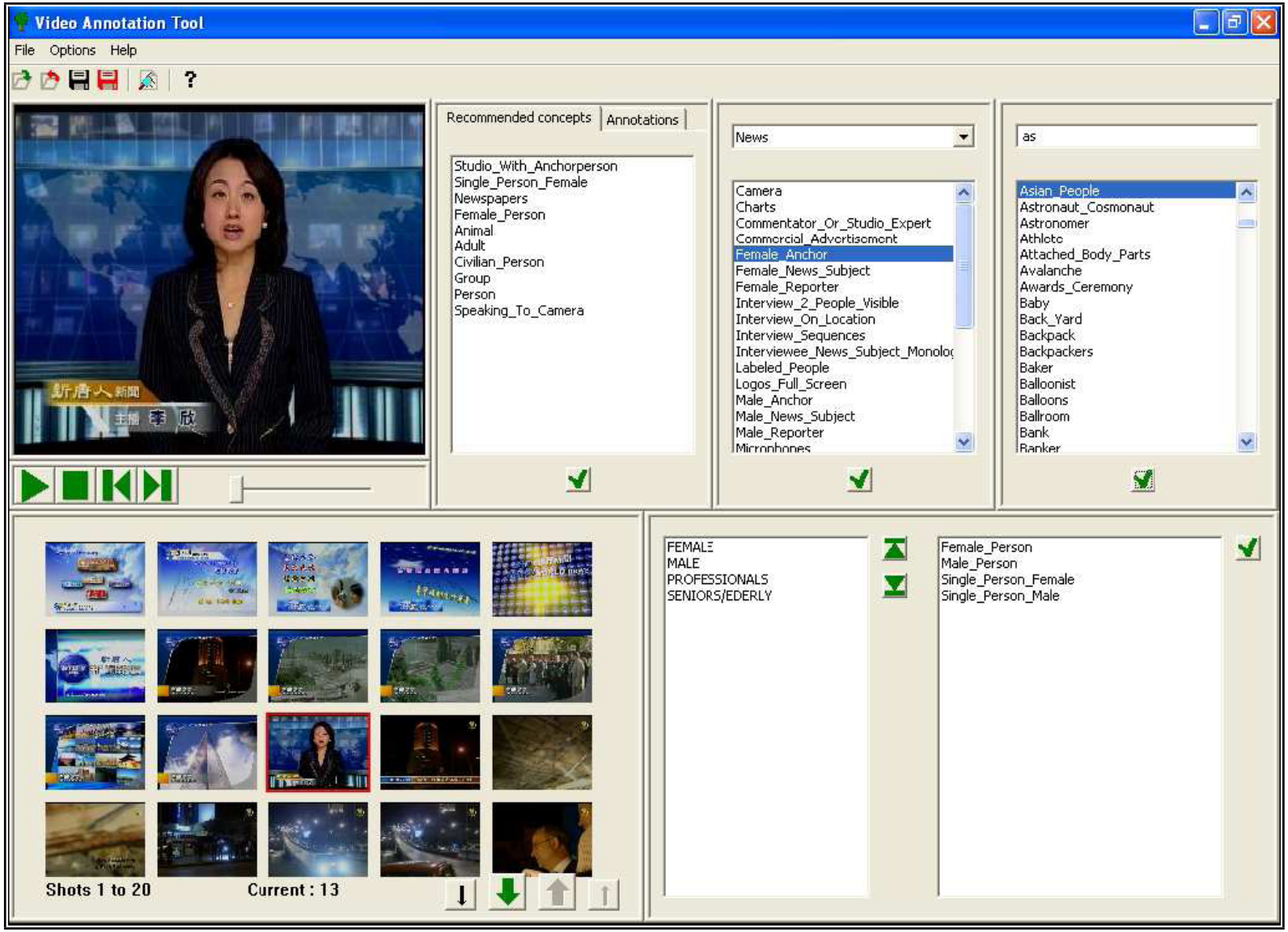
Metadata: descriptive

- Descriptive data in some specified format: *author*, *date*, *format*, *size*, ...
- Dublin Core, MPEG-7

The Internet Archive

Metadata: annotations

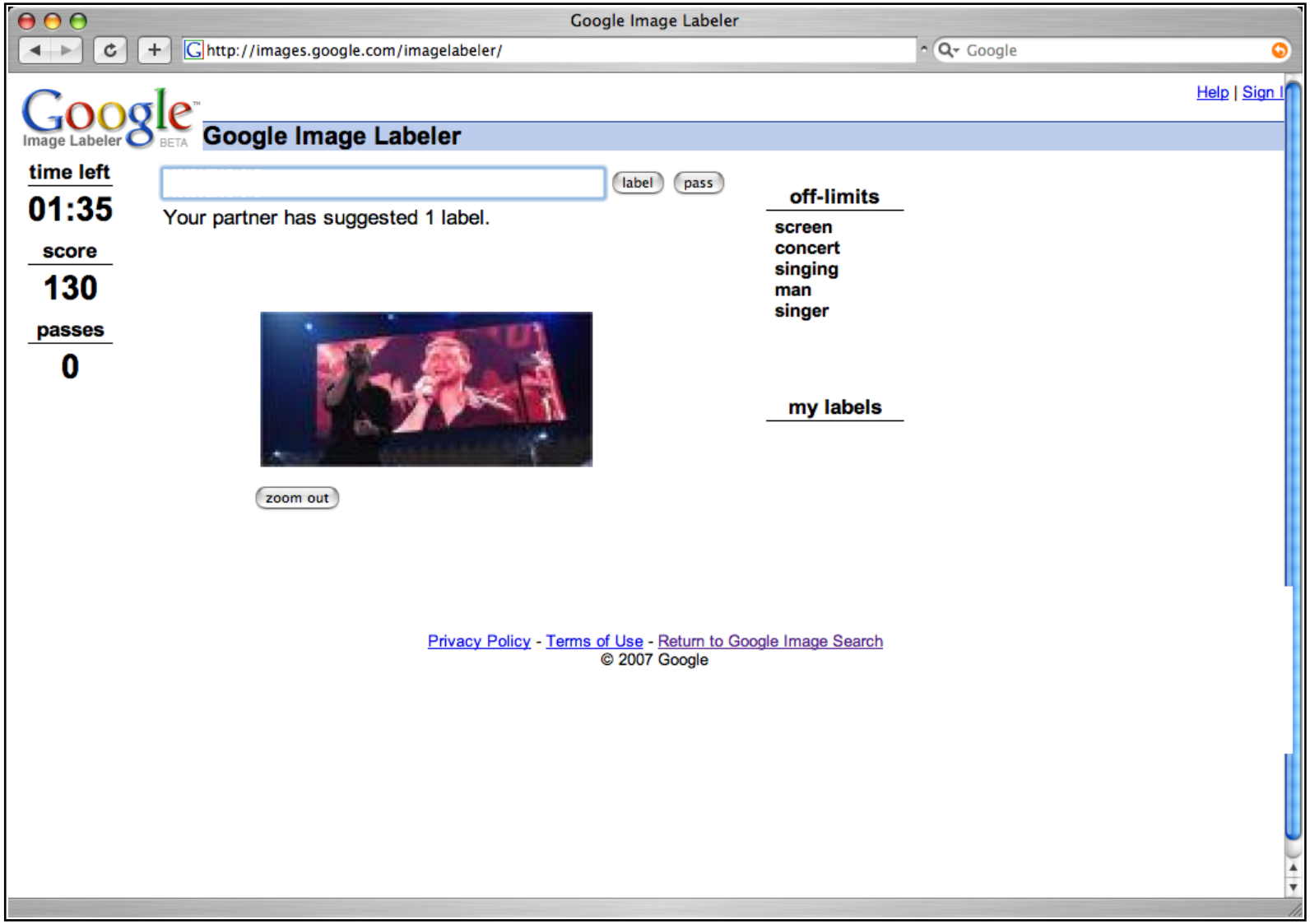
- Lists of keywords, free-format, ontology-based
- Manual annotation
  - time consuming, expensive, subjective
  - flickr, LabelMe, Google Image Labeler, ...
- Automatic annotation
  - a popular research topic



CDVP Video Annotation Tool, Dublin City University


LabelMe: The open annotation tool
http://labelme.csail.mit.edu/tool.html?collection=LabelMe&mode=i&folder=static_hotel_room_indoor&im
LabelMe
Please contact us if you find any bugs or have any suggestions.
Show me another image
Label as many objects and regions as you can in this image
Sign in (why?)
There are 207222 labelled objects
Instructions (Get more help)
Use your mouse to click around the boundary of some objects in this image. You will then be asked to enter the name of the object (examples: car, window).
Good
Bad
Labeling tools
Erase segment
Zoom
Fit Image
Polygons in this image (XML)
tv
ceiling
curtain
curtain
furniture
bag
shoolbag
chair
table occluded
Done

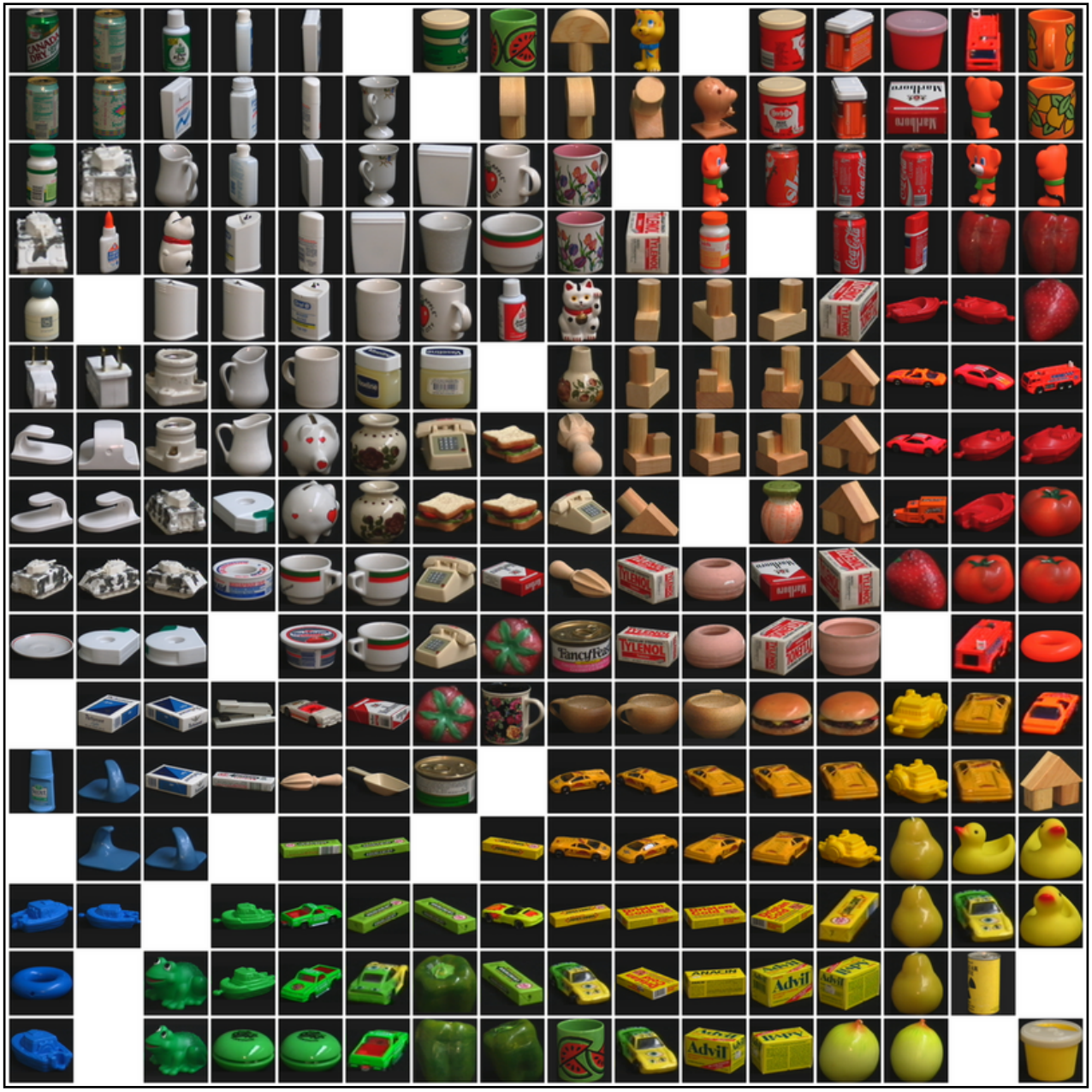
LabelMe, MIT

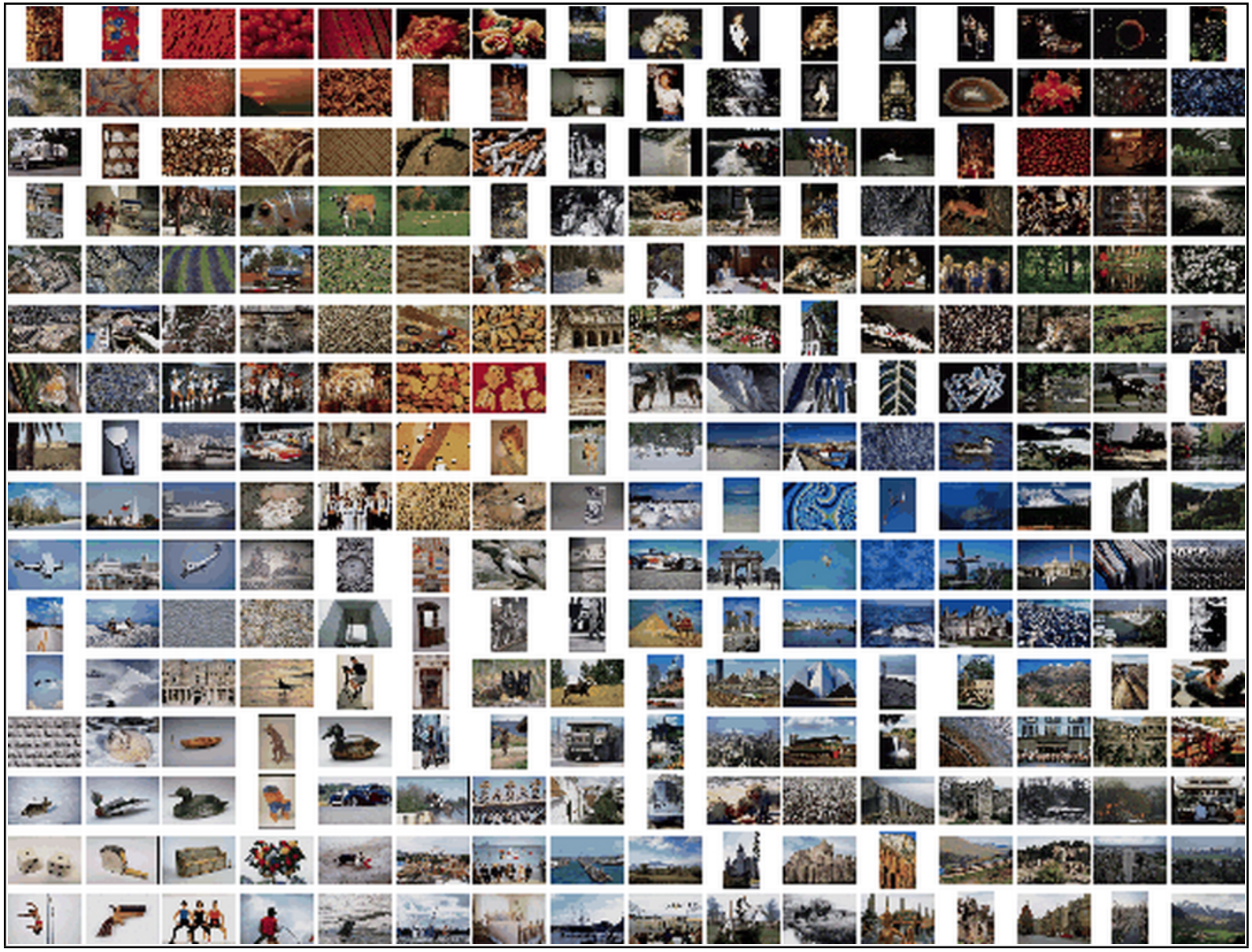


Google Image Labeler, © Google

Metadata: low-level features

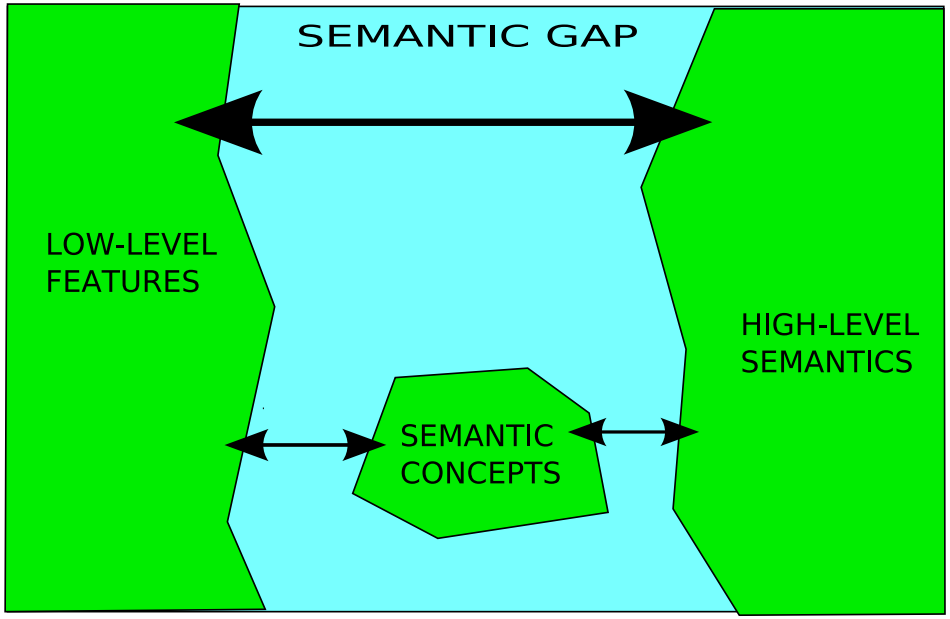
- Features that can be extracted automatically, usually represented as fixed-dimensional vectors
- **Text**: bag of words, n-grams, ...
- **Audio**: ZCR, DFT, silence ratio, ...
- **Images**: color histogram, color layout, edge histogram, Gabor texture, shape features, ...
- **Video**: all of above, motion, ...


TYLENOL
Advil

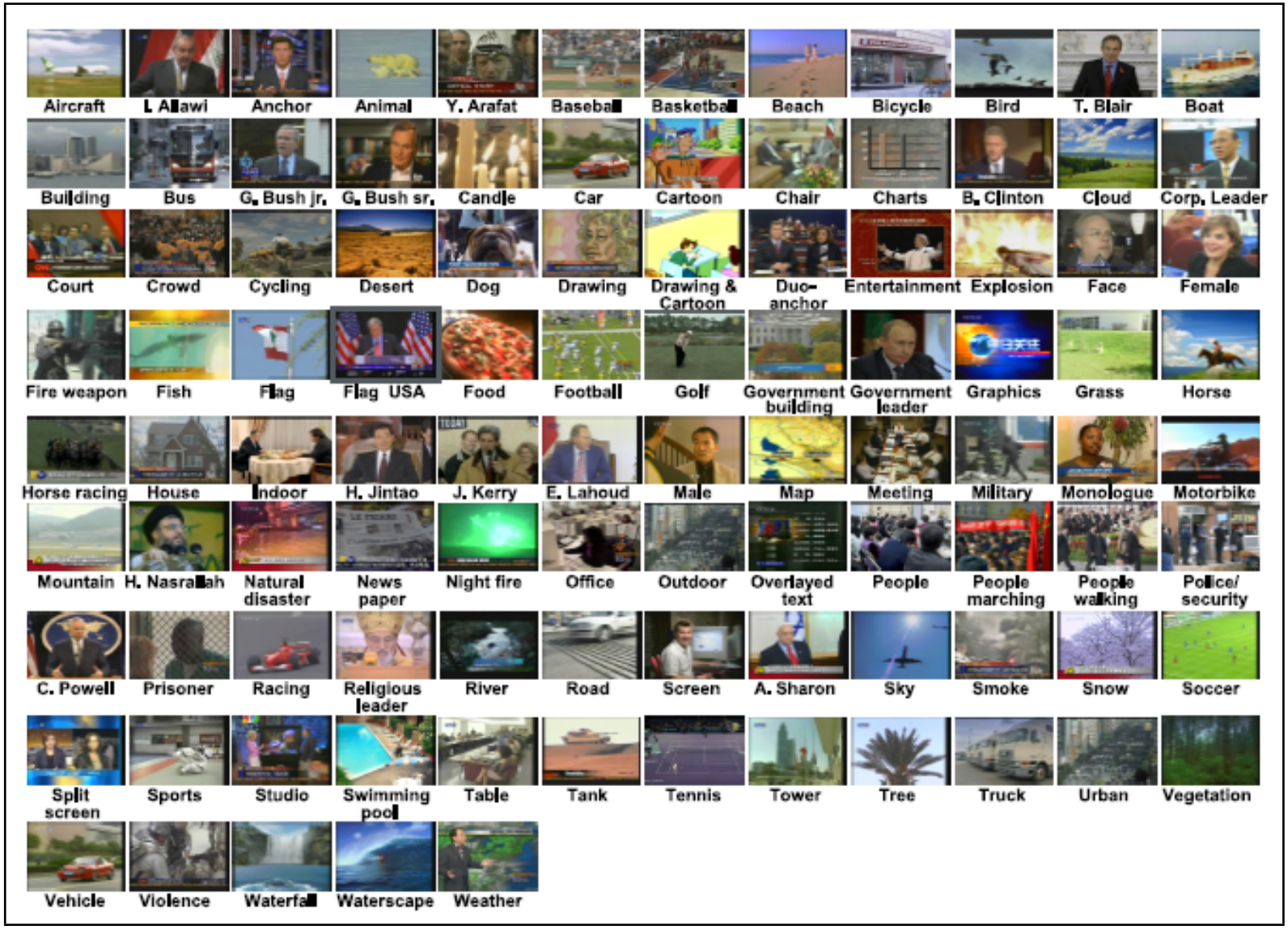


Metadata: high-level features

- End users are interested in semantically meaningful metadata, not low-level features: **the semantic gap**



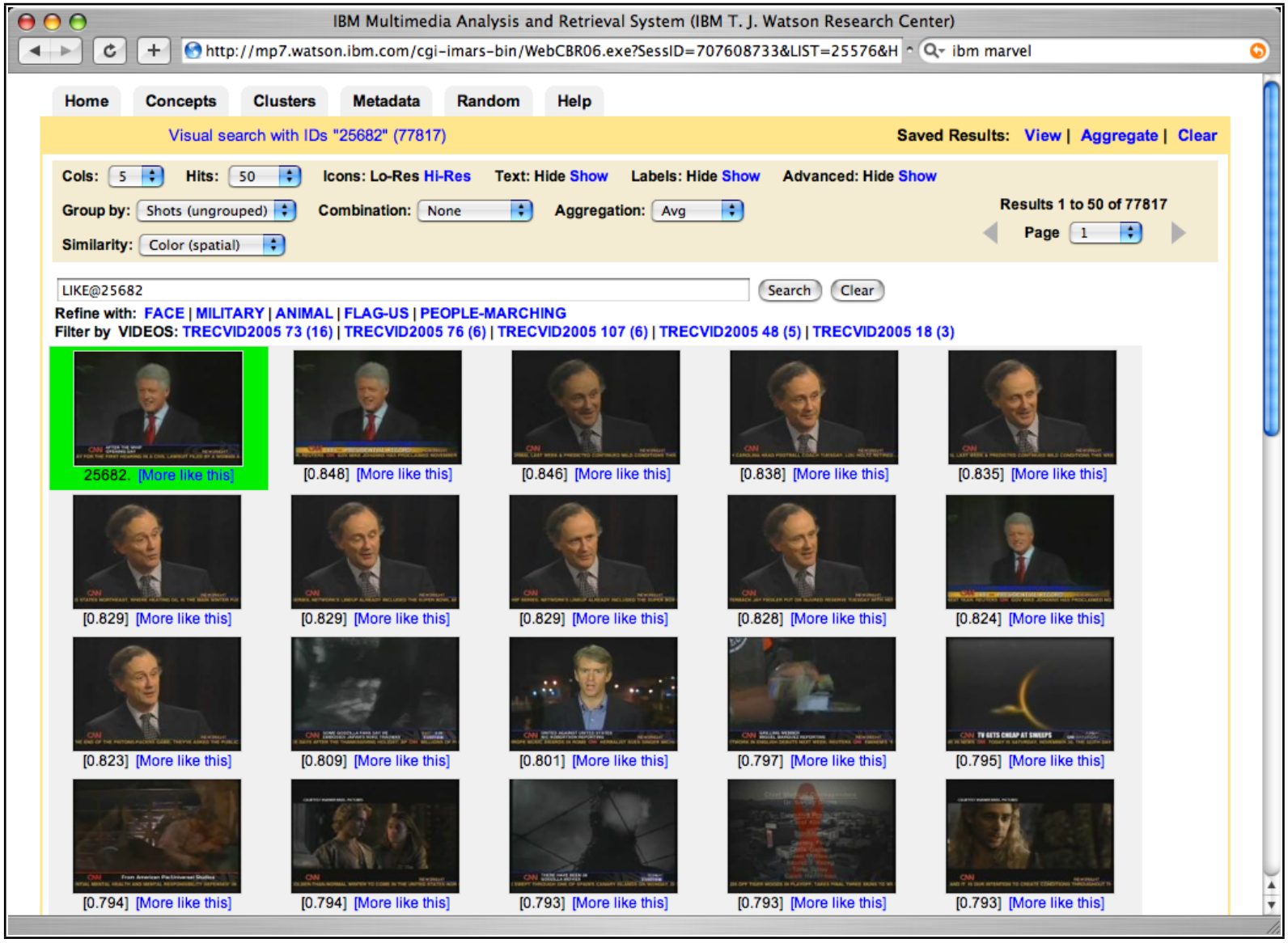
- An important question: **How to derive high-level features from the low-level ones?**



MediaMill, University of Amsterdam

Multimedia queries

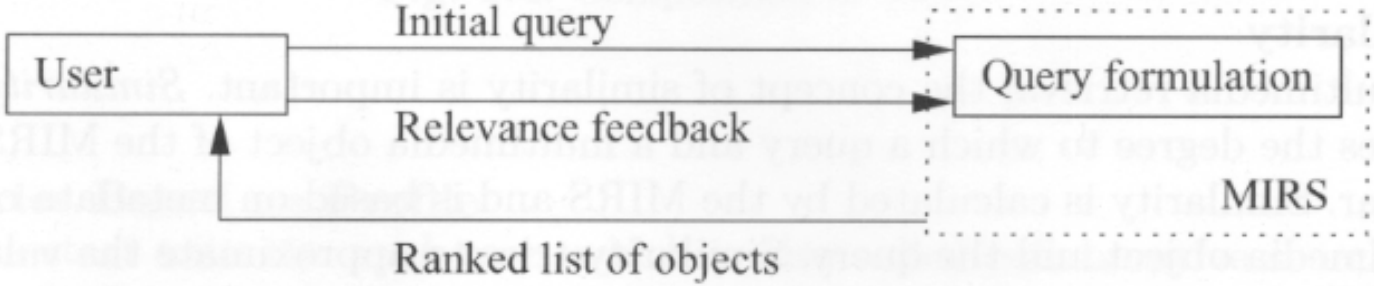
- More complex than text searches or SQL queries
- **Degree of similarity** instead of exact match
- Different types of queries:
  - query by descriptive metadata
  - query by annotations
  - query by low-level and high-level features
  - query by example
  - query by user profile

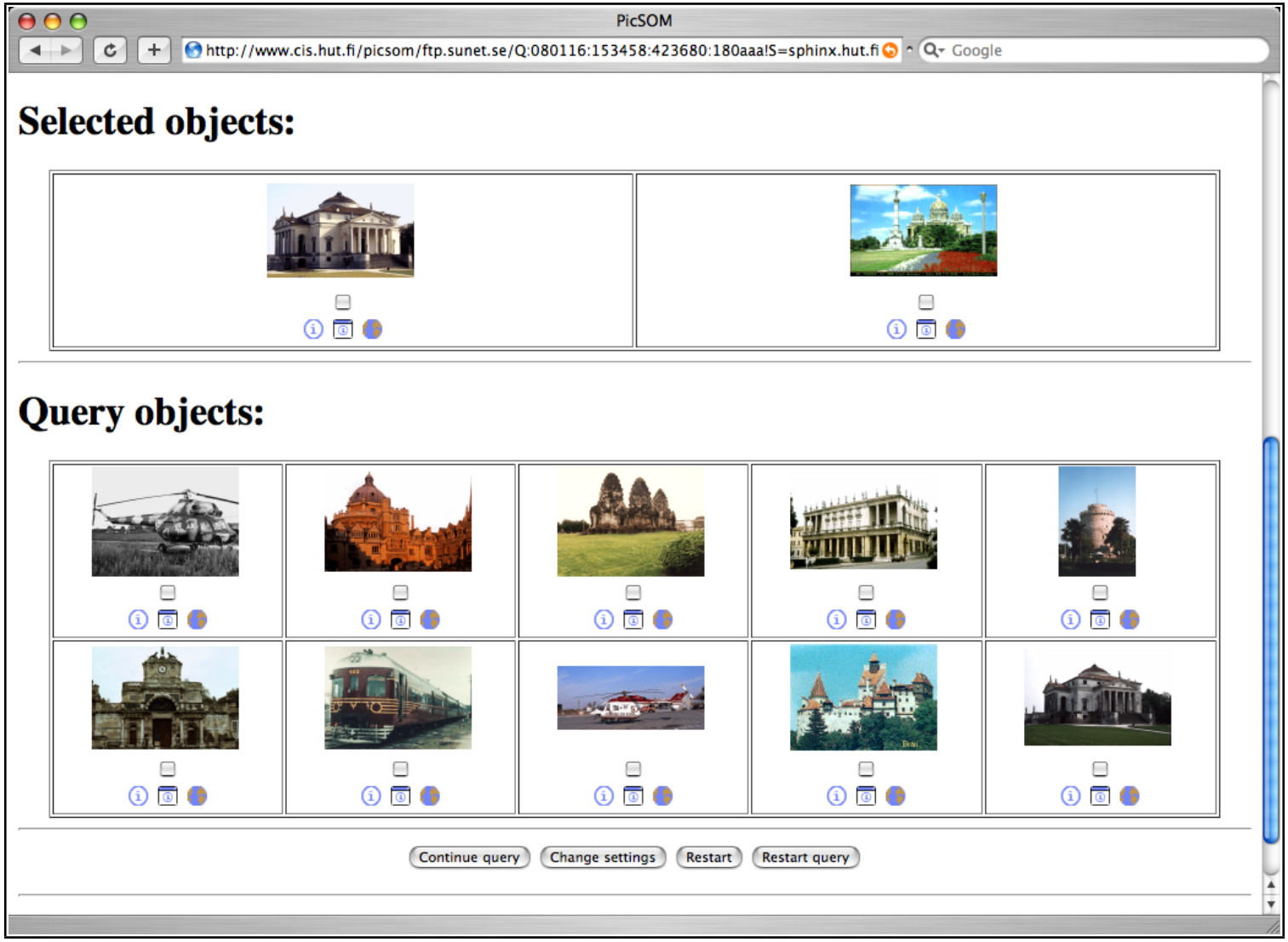


IBM Marvel Search Engine, © IBM

Relevance feedback

- Formulating a multimedia query can be difficult
- One solution: iterative and interactive refinement of the original query
- The user evaluates the relevance of retrieved items, which is then fed back to the system

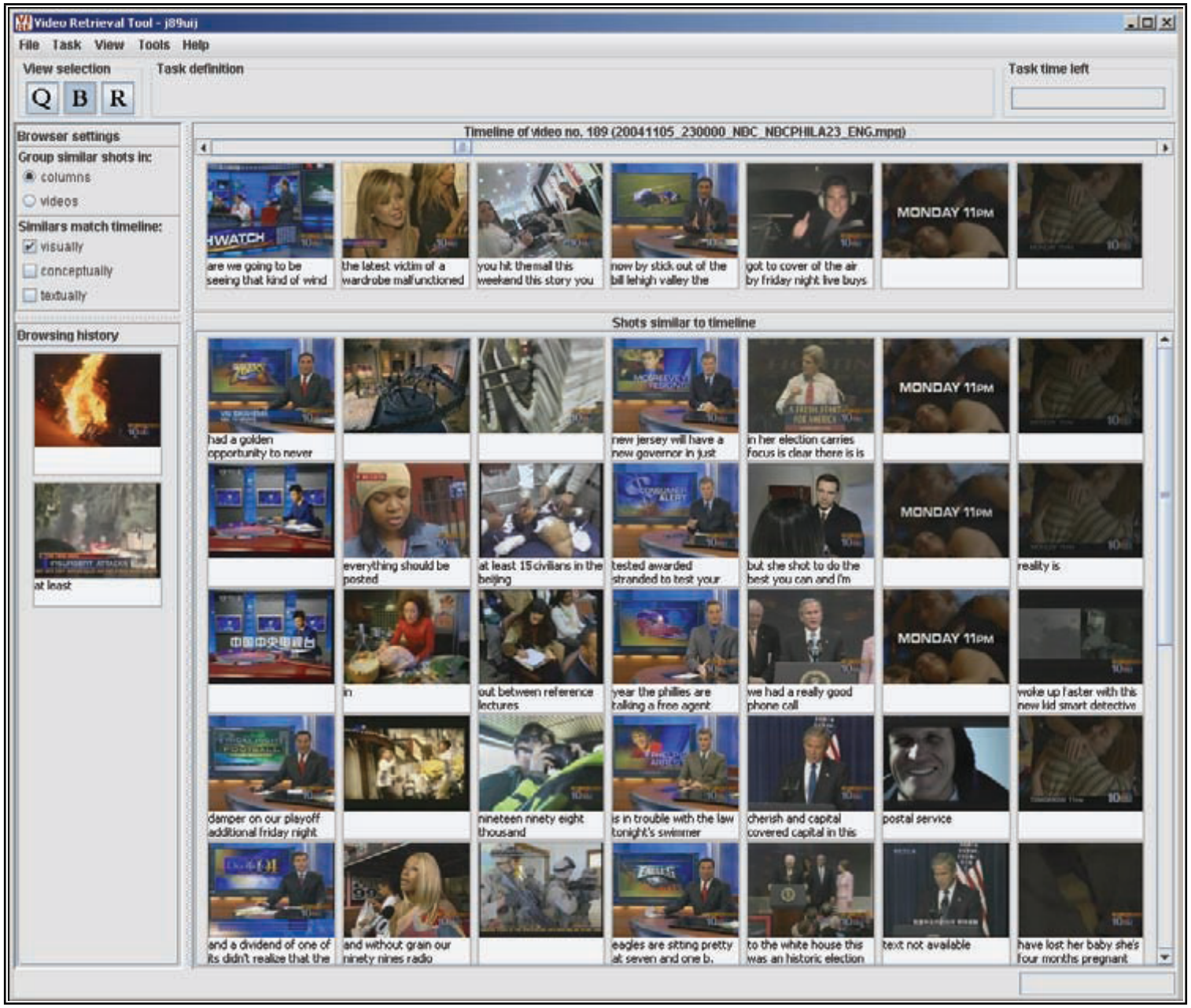




PicSOM, Adaptive Informatics Research Centre, TKK

Browsing

- Another solution: let the user to quickly scan through the (summarized) data
- Requires an effective interface
- Browsing and retrieval can be combined



Cluster-temporal browser, MediaTeam, University of Oulu

Sphere Browser, MediaMill, University of Amsterdam