Multimedia Knowledge Laboratory

Research Agenda on Social Web and Media

Yiannis Kompatsiaris, MKLab Head

Symeon Papadopoulos, MKLab Researcher

Yahoo! Research Labs

Barcelona, 12 Dec 2008

Overview

- MKLab: Short Intro
- Our Vision: Harvest Intelligence from the Masses
  - WeKnowIt
  - Topic community detection in tagging systems
  - Co-clustering of tags & resources
  - Generating ground truth from Social Annotations
  - Case studies / Applications

MKLab (1)

- Located outside Thessaloniki, Greece.
- MKLab ∈ Informatics & Telematics Institute (ITI) ∈ Centre for Research and Technology Hellas (CERTH)
- Main research interests
  - Image & Video analysis (feature extraction, content classification, clustering)
  - Semantics (ontology design, reasoning)
  - Image & Video Retrieval
  - Social Web

MKLab (2)

- 30 people
  (researchers, developers, administration)
- Participation in many European and national research projects.
  - FP7: WeKnowIt (coordination), JUMAS
  - FP6: AceMedia, X-Media, MESH, BOEMIE, VIDI-Video, K-Space, PATExpert, ELU, etc.
- Possibilities for:
  - Bilateral collaborations for R&D and EU projects
  - Joint Publications / Patents

MKLab / traditional lines of research

Computer Vision

- Feature extraction, image filtering
- Segmentation (images, videos)
- Event detection (videos)

Machine Learning / Classification

- SVM
- Hidden Markov Models

Logic

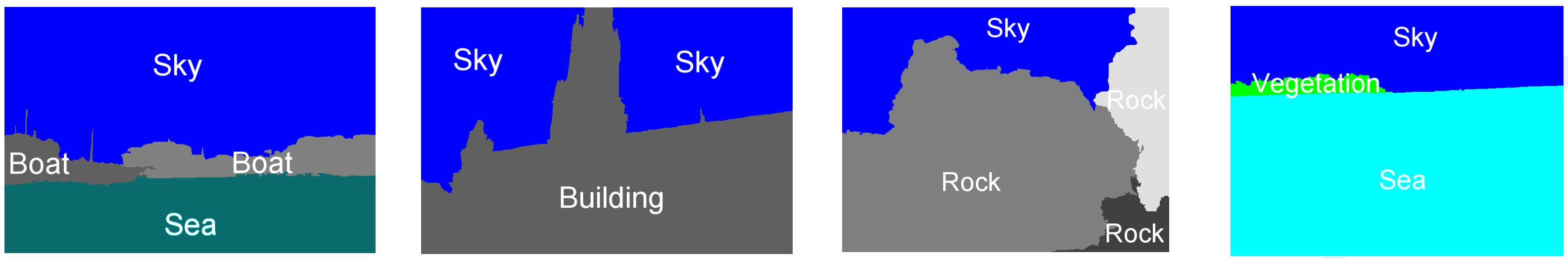
- Ontology design
- Reasoning, Rules

- Context estimation and usage
- Hybrid Techniques

Image Segmentation and Region Classification

Input images

Region-concept associations



GA PSO SOM SVM

Evaluation of Image Region Classification

Concepts

c1: Building, c2: Foliage, c3: Mountain, c4: Person, c5: Road, c6: Sailing-boat, c7: Sand, c8: Sea, c9: Sky, c10: Snow

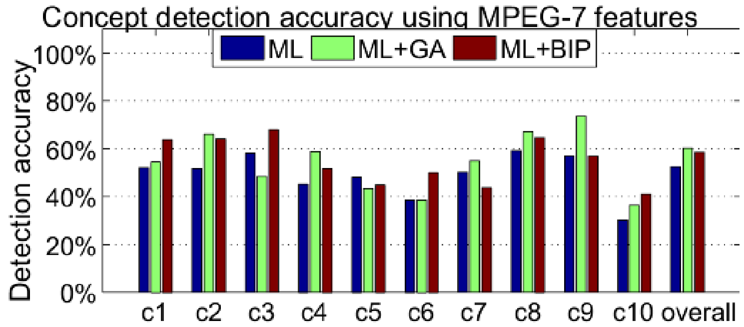
Classification Techniques

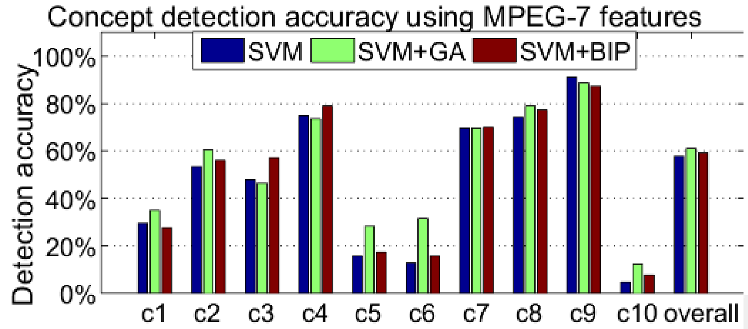
ML	Maximum Likelihood
SVM	Support Vector Machine
GA	Genetic Algorithm
BIP	Binary Integer Programming

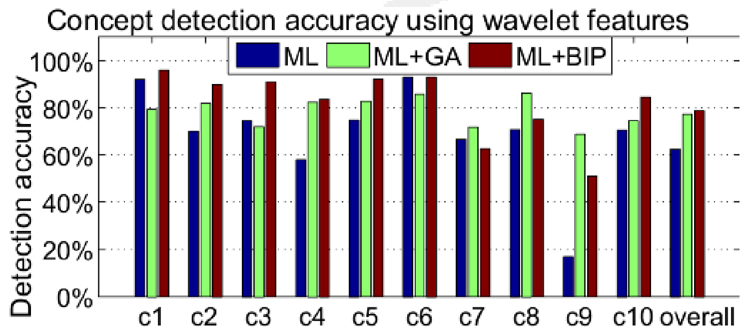
Data set

~450 images train / ~ 3000 regions

~450 images test



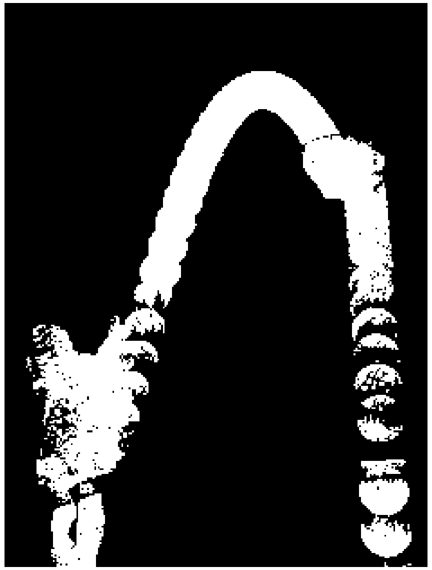


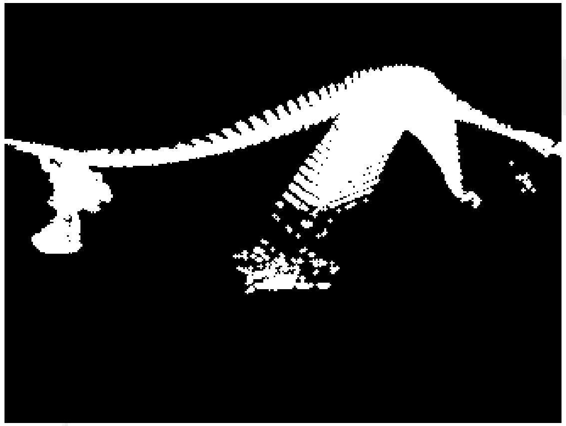


Event Detection in Videos

Activity area detection by means of motion features.

Event detection based on activity area characteristics.





Initial Assertions

(image:Indoor) ≥ 0.67
(image:$\exists$contains.Sea) ≥ 0.73
(image:$\exists$contains.Sand) ≥ 0.58
(image:$\exists$contains.Mountain) ≥ 0.85

Disjointness axioms removed

(image:Indoor) ≥ 0.67
(image:$\exists$contains.Sea) ≥ 0.73
(image:$\exists$contains.Sand) ≥ 0.58
(image:Coastal) ≥ 0.73
(image:Beach) ≥ 0.58
(image:Natural) ≥ 0.73
(image:Outdoor) ≥ 0.73
(image:$\exists$contains.Mountain) ≥ 0.85
(image:Mountainous) ≥ 0.85
(image:Natural) ≥ 0.85
(image:Outdoor) ≥ 0.85

Scene level hierarchy

Outdoor (0.85) Indoor (0.67)
Natural (0.85) ManMade
Coastal (0.58) Mountainous (0.85)
Beach (0.58)

T1 step

Domain TBox

Natural $\equiv$ Outdoors $\sqcup \neg$ ManMade
Mountainous $\equiv$ Natural $\sqcup \neg$ Coastal
Beach $\equiv$ Coastal $\sqcap \exists$contains.Sand
$\exists$contains.Mountain $\sqsubseteq$ Mountainous
$\exists$contains.Sea $\sqsubseteq$ Coastal
$\exists$contains.Sand $\sqcap$ Mountainous $\sqsubseteq \perp$
Outdoor $\sqcap$ Indoor $\sqsubseteq \perp$

directly disjoint

inferred disjoint

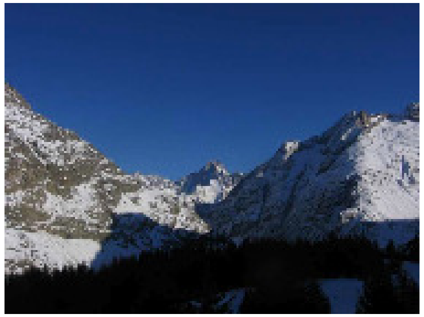
(image:Indoor) ≥ 0.67
(image:$\exists$contains.Sea) ≥ 0.73
(image:$\exists$contains.Sand) ≥ 0.58
(image:Coastal) ≥ 0.73
(image:Beach) ≥ 0.58
(image:$\exists$contains.Mountain) ≥ 0.85
(image:Mountainous) ≥ 0.85
(image:Natural) ≥ 0.85
(image:Outdoor) ≥ 0.85

Disjoint axioms restored

Inconsistency handling

(image:$\exists$contains.Mountain)≥ 0.85

T2 step



(image:$\exists$contains.Mountain) ≥ 0.85
(image:Mountainous) ≥ 0.85
(image:Natural) ≥ 0.85
(image:Outdoor) ≥ 0.85

Final Assertions

Web 2.0 as a Database

Information Retrieval
- keyword search
- tag-based browsing

Statistical Analysis
- tag frequency
- time series

Web 2.0: Multi-facet association

Social setting

Appearance

Location

Time

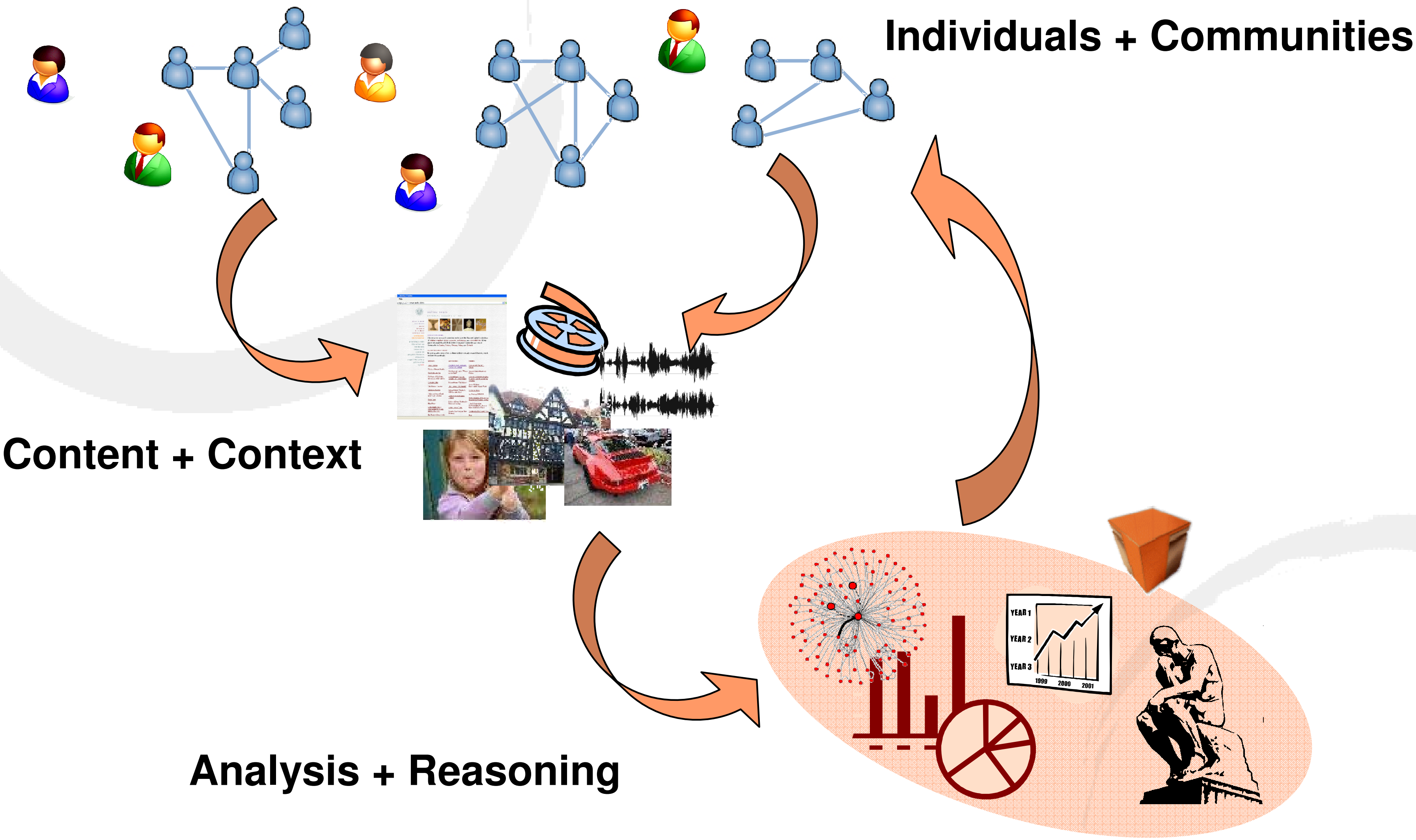
Tagging

camera barcelona SagradaFamilia
tourist pics spain catalunya monument
vacation **visitor** architecture **gaudi**

Machine Learning
- Context
- Rules

Facts
Relations
Trends

MKLab positioning in the ecosystem



Technologies involved

- Media processing
  Text / Image / Video → features, concepts
- Data mining & knowledge discovery
  Uncover structure (Clusters / Patterns)
- Reasoning with context
  Make inferences, check consistency of information

WeKnowIt

Emerging, Collective Intelligence for personal, organisational and social use

Mass user-generated content Web 2.0
Little understanding

Organizations – Processes
No benefits from community and mass content

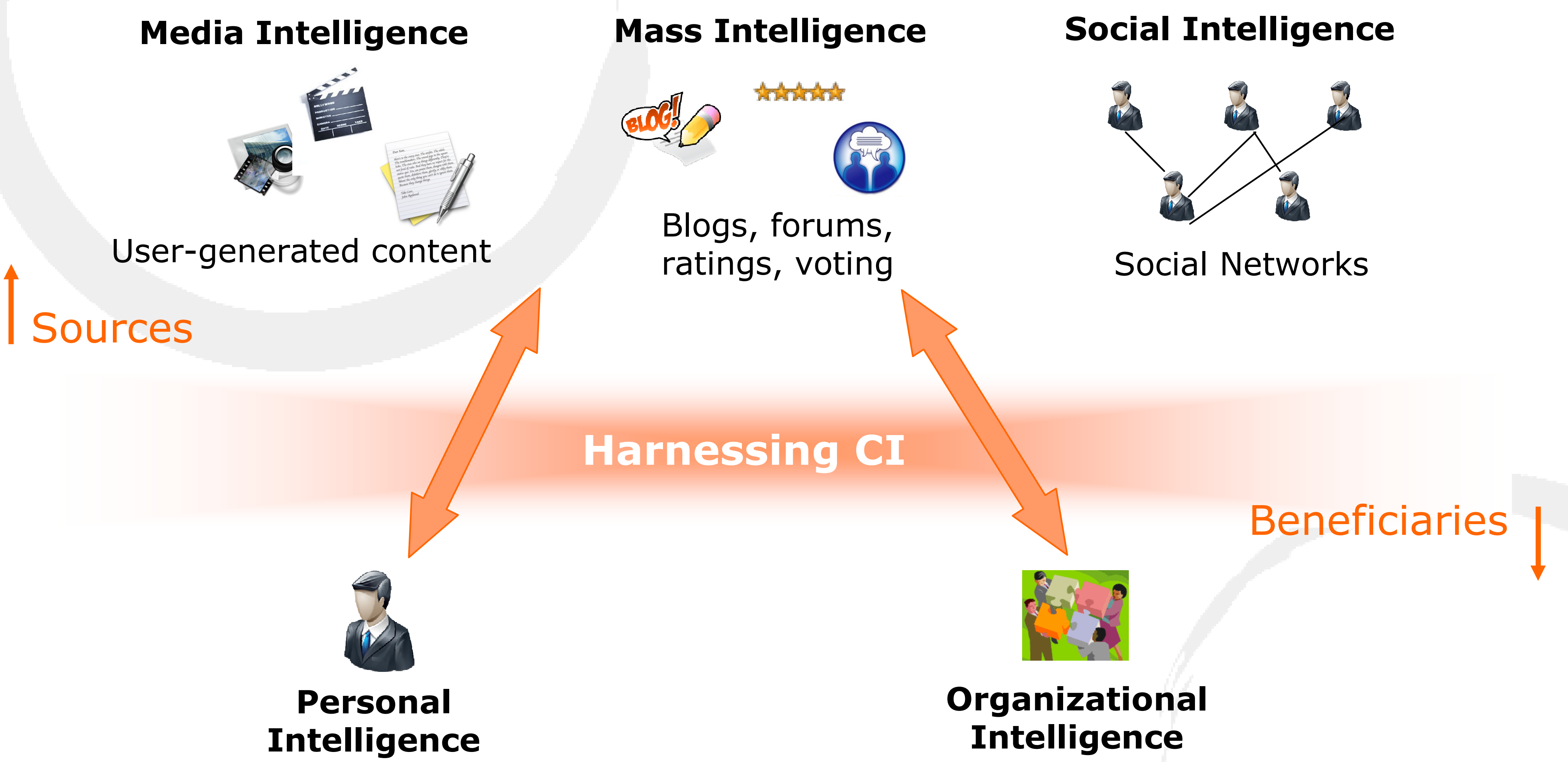
weknowit

Analysis techniques: Content, Social, Mass
Loose interaction

Users & Devices
Limited sharing and access

Collective Intelligence

Elements of Collective Intelligence



Personal Intelligence

Profile of contributor

>> What to send where, e.g. location, age, picture

Media Intelligence

Picture arrives at emergency response

>> Automatic detection of a fire event

Organizational Intelligence

The right knowledge to the right people at the right time

>> Whom (fire-fighters, ambulances,…) to inform about what

Mass Intelligence

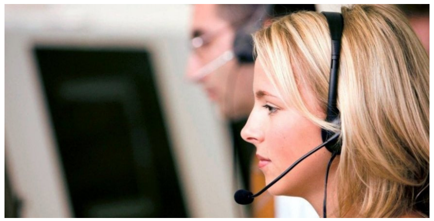
Many contributors

>> Extraction of trends about the scale of the incident

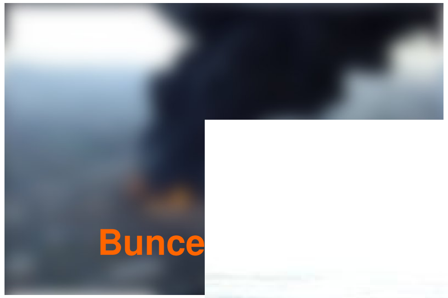
Social Intelligence

Trust and feedback

>> Determine trustworthiness and hub-structures by SNA



Personal Intelligence



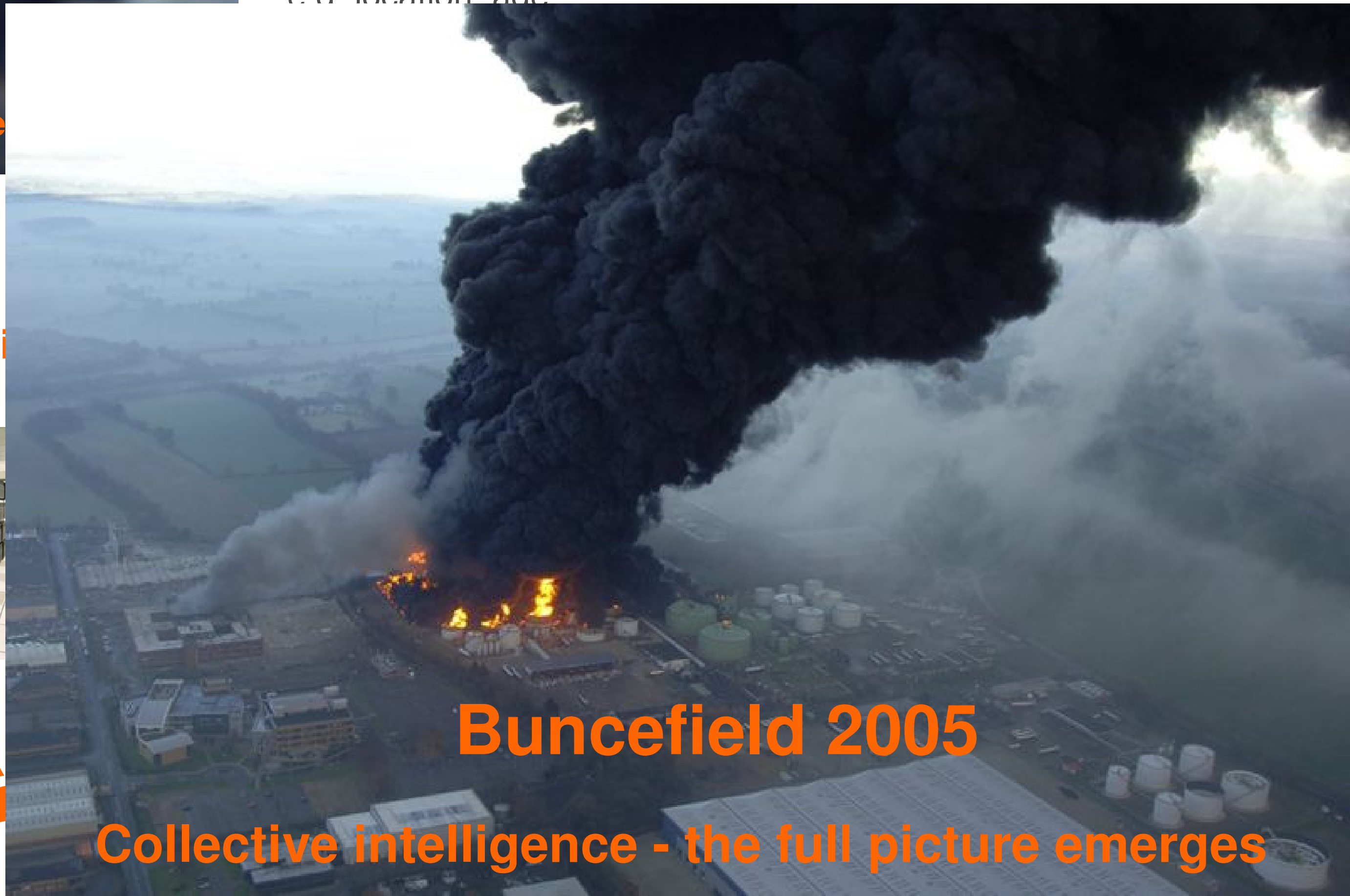
Profile of contributor

\>> What to send where,
e.g. location, age

Media Intelligence

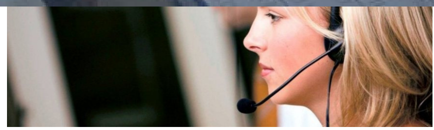
se
on

Organizati



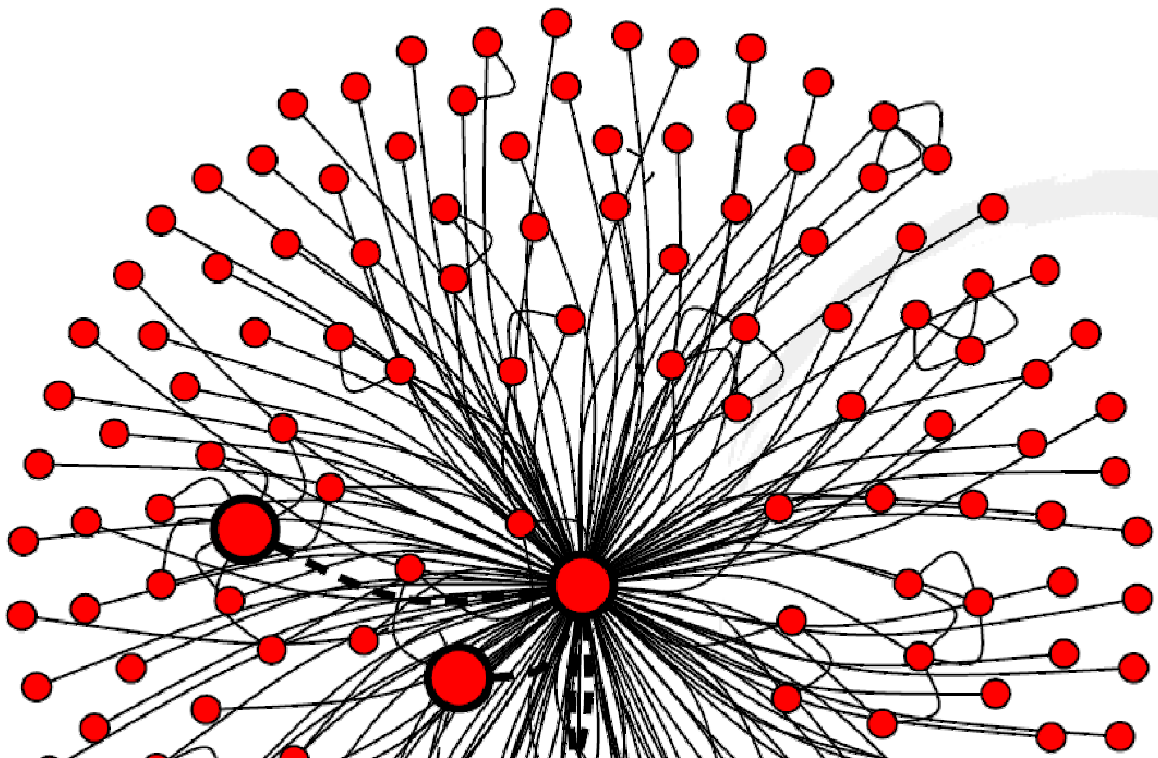
Trust and feedback

\>> Determine trustworthiness
and hub-structures by SNA



Community Detection in complex networks

A local approach

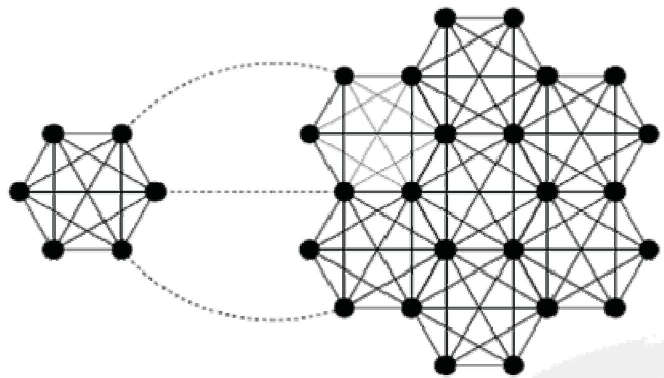


Symeon Papadopoulos
prof. Athena Vakali

Community Detection

Problem

Identify groups of vertices in a network that are closely intertwined → communities.



Applications

Social systems

Metabolic networks

Ecological webs

Web graphs

Web 2.0: blogs, tagging systems, content

Global vs. Local

Global measures for community:

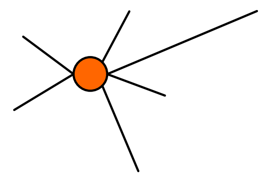
- Intra-edges $$\sum_{v \in C} w_{uv} \geq \sum_{v \in V-C} w_{uv} \ \textit{for all } u \in C.$$
- Modularity $$Q = \sum_{i} (e_{ii} - a_i^2) = \mathrm{Tr}\,\mathbf{e} - \| \mathbf{e}^2 \|$$
- Conductance $$\phi(S) = \frac{\sum_{i \in S, j \notin S} A_{ij}}{\min\{A(S), A(\overline{S})\}}$$

Although many good methods available for global detection, the problem is still hard in practical settings.

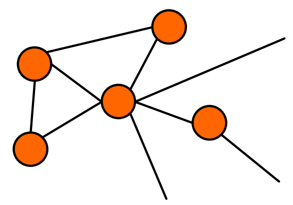
→ **Local community detection**

Local community detection

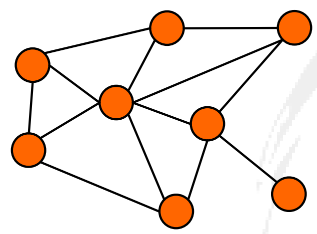
Start from a seed node.



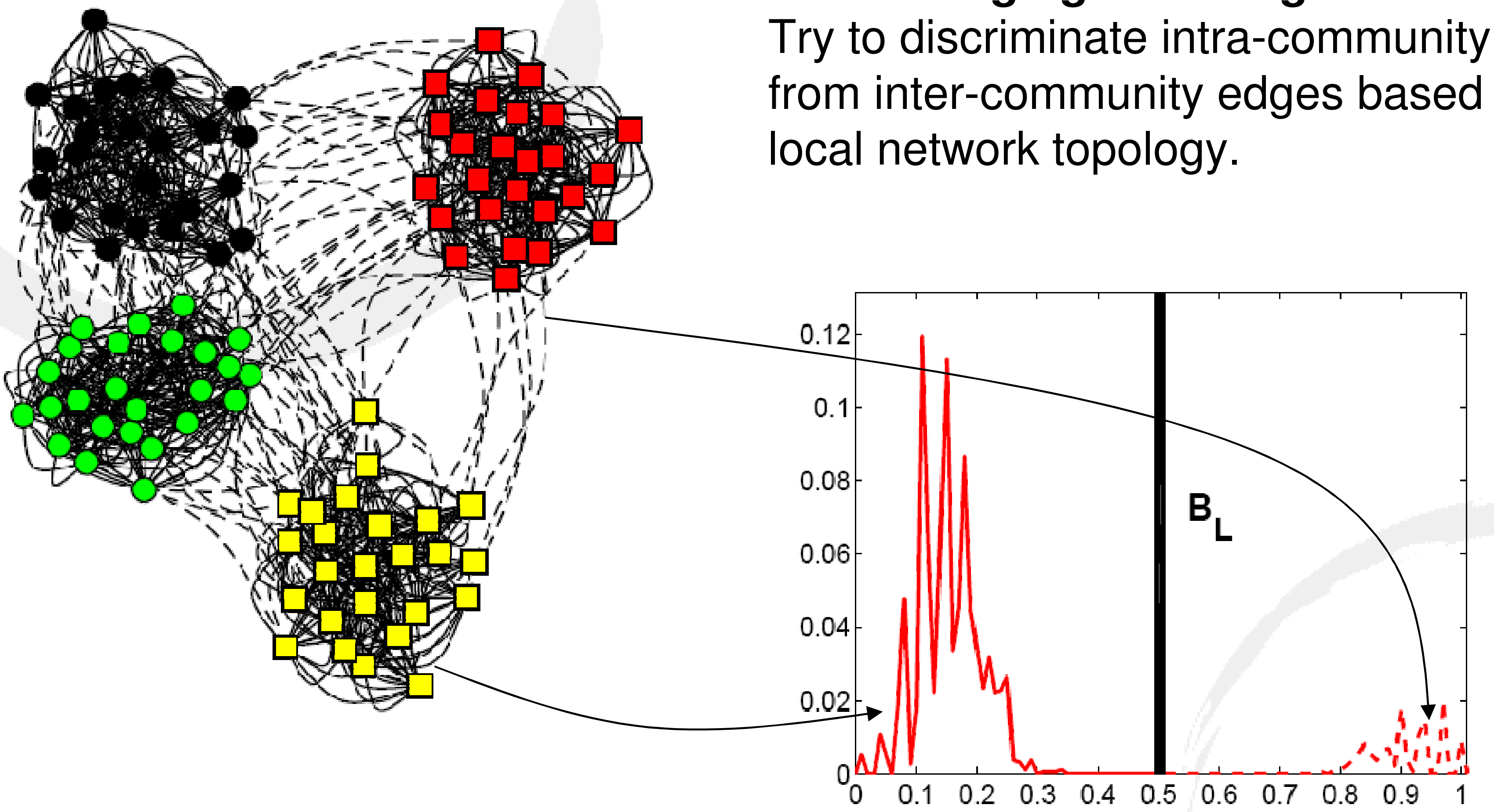
Attach neighboring nodes using some criterion.



Stop when a certain condition is satisfied.



Bridge Bounding

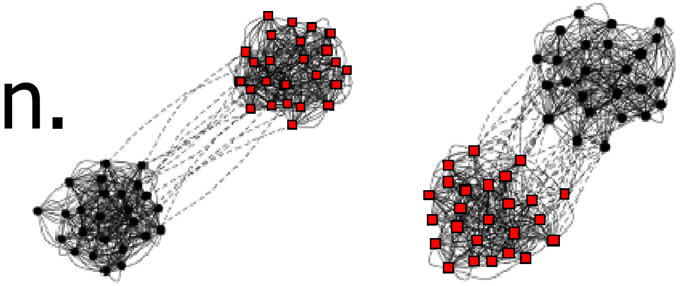


Local bridging of an edge
Try to discriminate intra-community from inter-community edges based on local network topology.

Evaluation (synthetic networks)

Synthetic networks according to method of Newman and Girvan.

$$S_{PAR} = \{N, K, z_{tot}, p_{out}, s_{var}\}$$



(a) $p_{out} = 0.01$ (b) $p_{out} = 0.08$

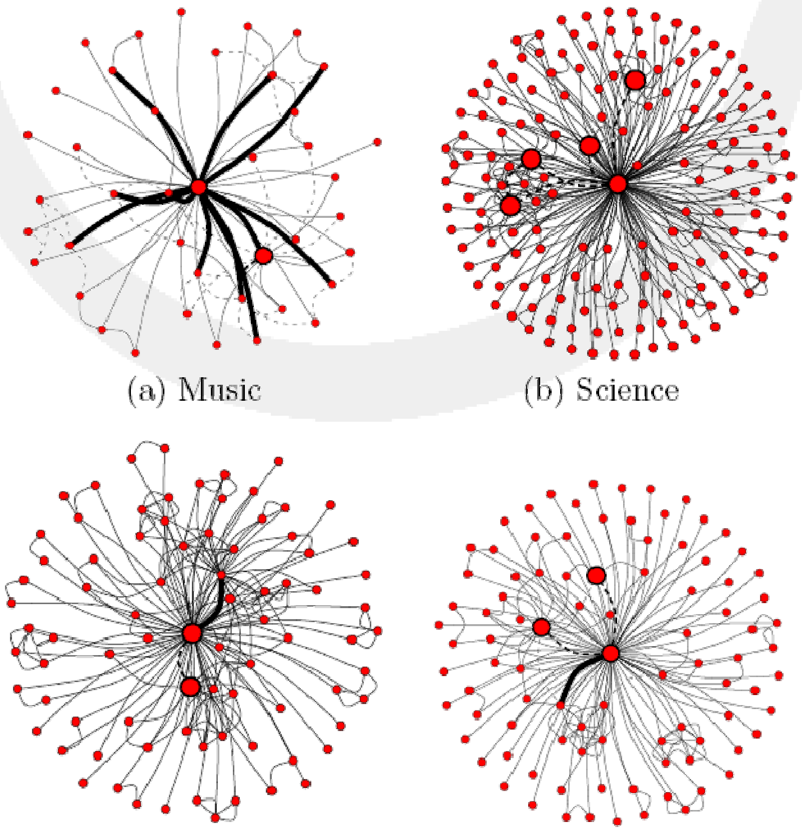
Change conspicuity of underlying communities (p_{out}).

	F_C			NMI		
p_{out}	BB	BB'	GN	BB	BB'	GN
0.01	100	100	100	1.0	1.0	1.0
0.05	100	100	100	1.0	1.0	1.0
0.1	100	100	50	1.0	1.0	0.86
0.15	100	99	50	1.0	.98	0.86
0.20	99	74	50	0.98	0.84	0.86
0.25	24	24	0	0.54	0.56	0.02

Change relative sizes of underlying communities (s_{var}).

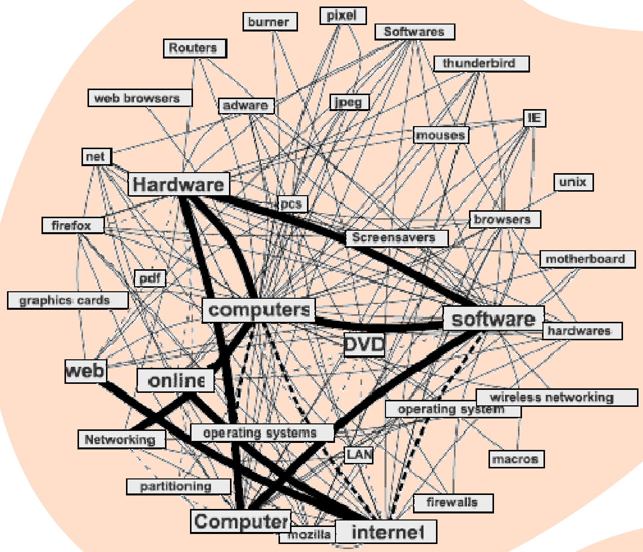
	F_C			NMI		
s_{var}	BB	BB'	GN	BB	BB'	GN
1.1	100	100	100	1.0	1.0	1.0
1.5	100	100	100	1.0	1.0	1.0
1.6	99.5	100	100	0.99	1.0	1.0
1.7	88	98	100	0.82	0.96	1.0
1.8	85.5	97	100	0.79	0.95	1.0
1.9	58.5	87	90	0.68	0.82	0.88
2.0	12.5	80	82	0.45	0.73	0.81
2.5	0	62	75	0.45	0.63	0.72

Case Study: LYCOS iQ

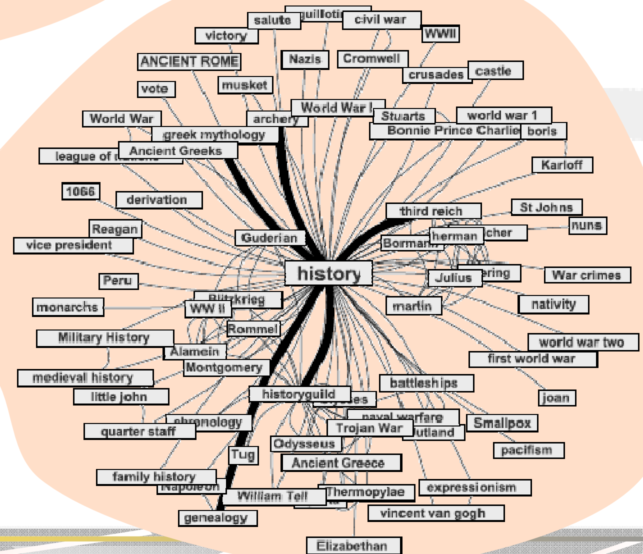


(a) Music (b) Science

(c) Film (d) Animals



Computers:
A densely interconnected community



History:
A star-shaped community

Clustering in Multimedia Social Tagging Systems

A case study in Flickr

Eirini Giannakidou
prof. Athena Vakali

Social Tagging & Multimedia

Setting

Abundance of

- multimedia content in social media sharing sites
- user-generated metadata

Convergence in the patterns of tagging behaviour

Motivation

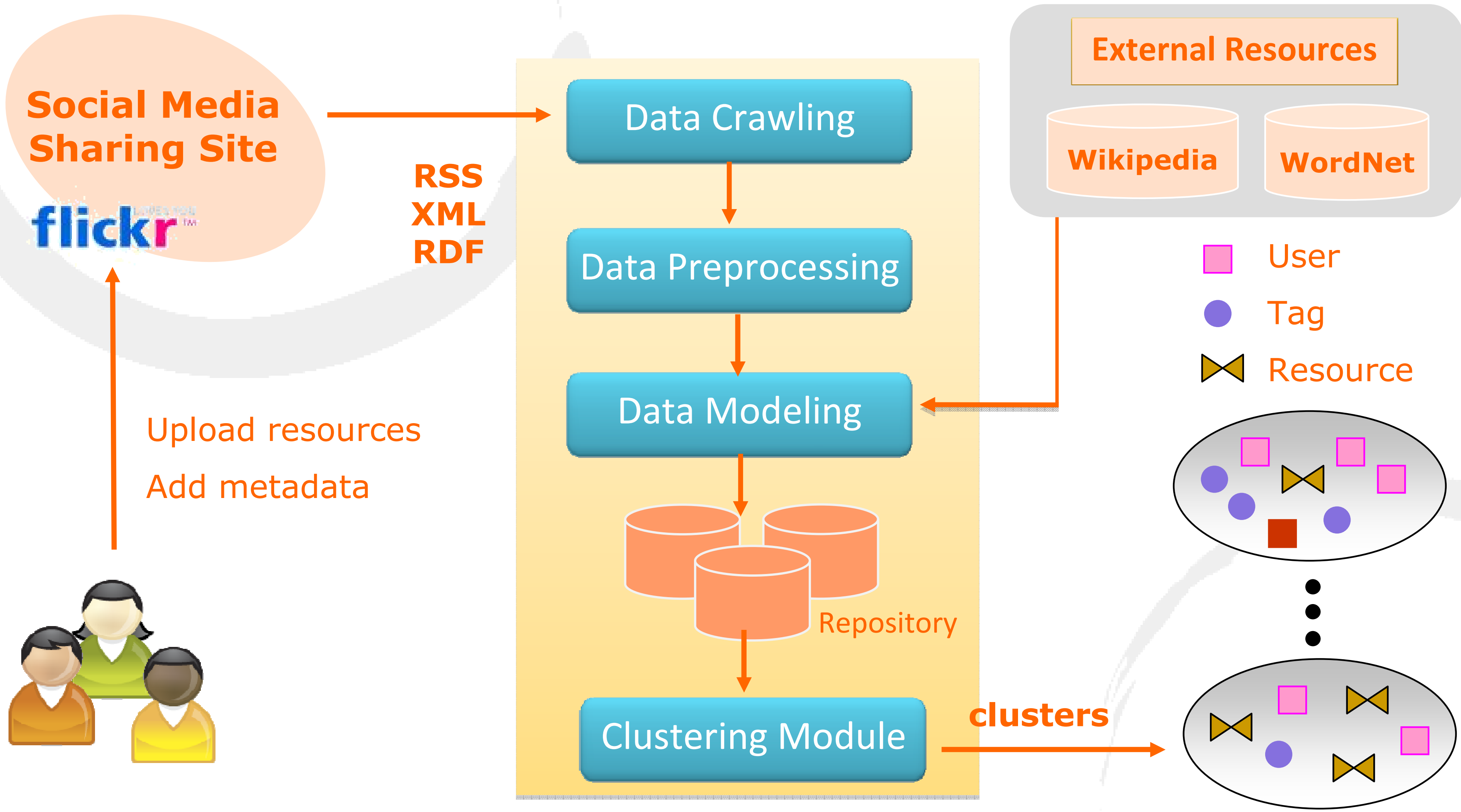
Poor IR (lack of structured information, tag polysemy/ ambiguity)

Questionable tag validity

Problem Formulation

Exploit knowledge hidden in social media sharing sites through clustering.

Cluster structure in folksonomies



Clustering resources

Vector Space Model

- Each resource (picture) is projected onto a space defined by the most prominent tags.
- Projection takes place by using:

  Semantics (WordNet)

  Usage context (co-occurrence)
- Some standard clustering scheme is employed on the $N \times D$ matrix (N resources, D dimensions): K-means, Hierarchical, COBWEB

Experimental results

Buildings

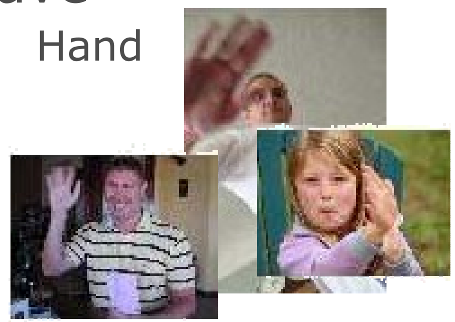
Cars

Festival

Wave

Sea

Hand



Rock

Stone

Music

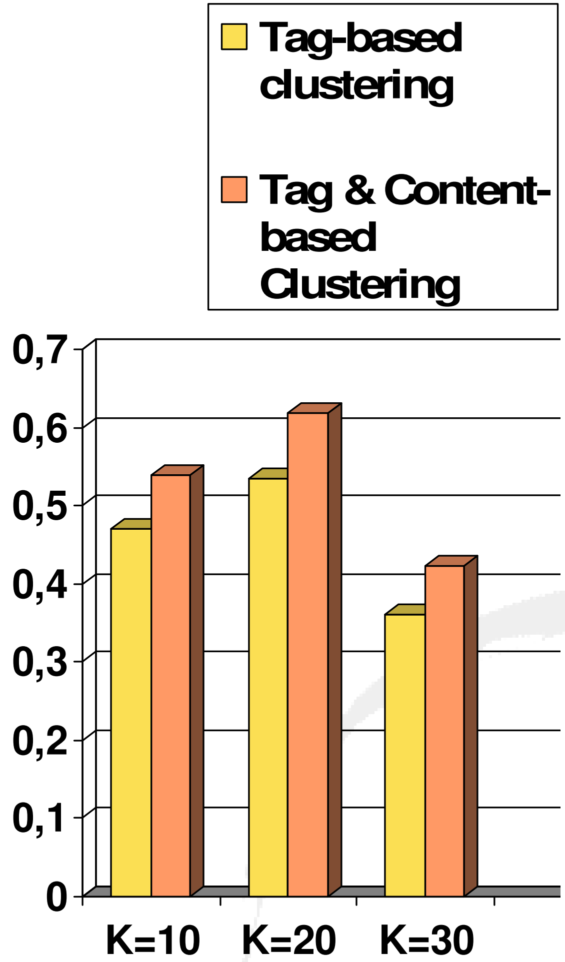
Refining clustering by visual cues

MPEG-7 descriptors

Scalable Color

Color Structure, etc.

- Carry out 2nd order clustering based on visual features → break-apart polysemous clusters
- Outlier detection and removal

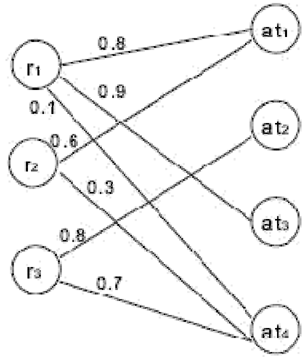


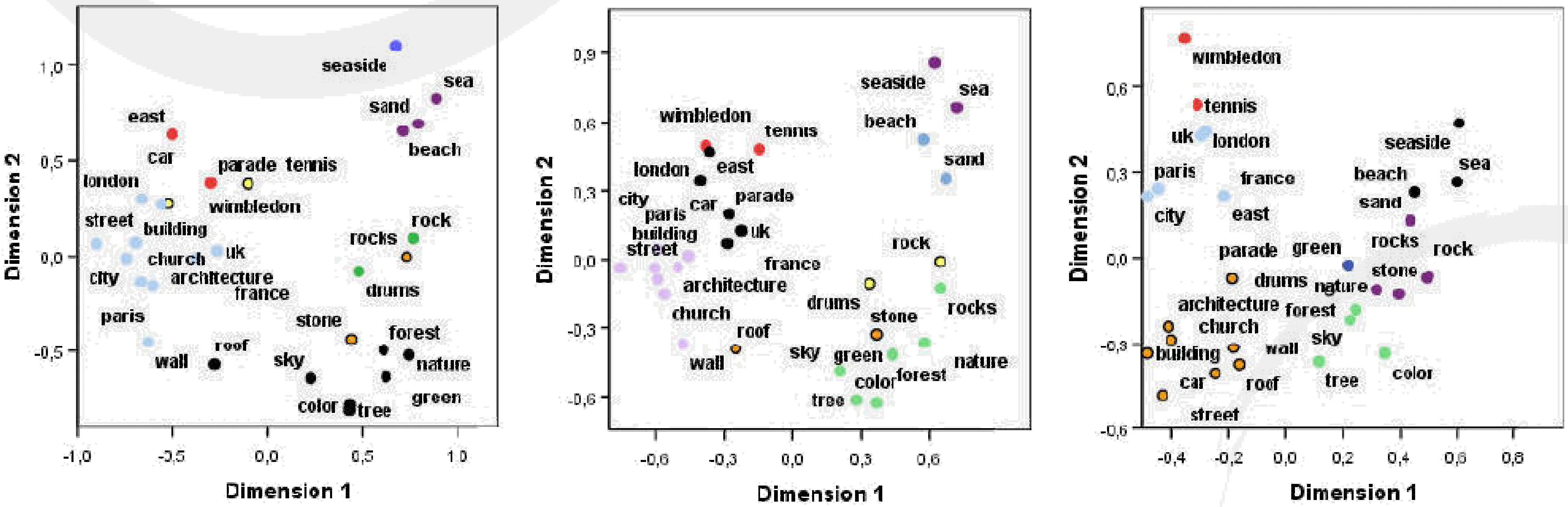
Co-clustering tags & pictures

Graph spectral methods and SVD →

reduced dimensionality

clustering problem



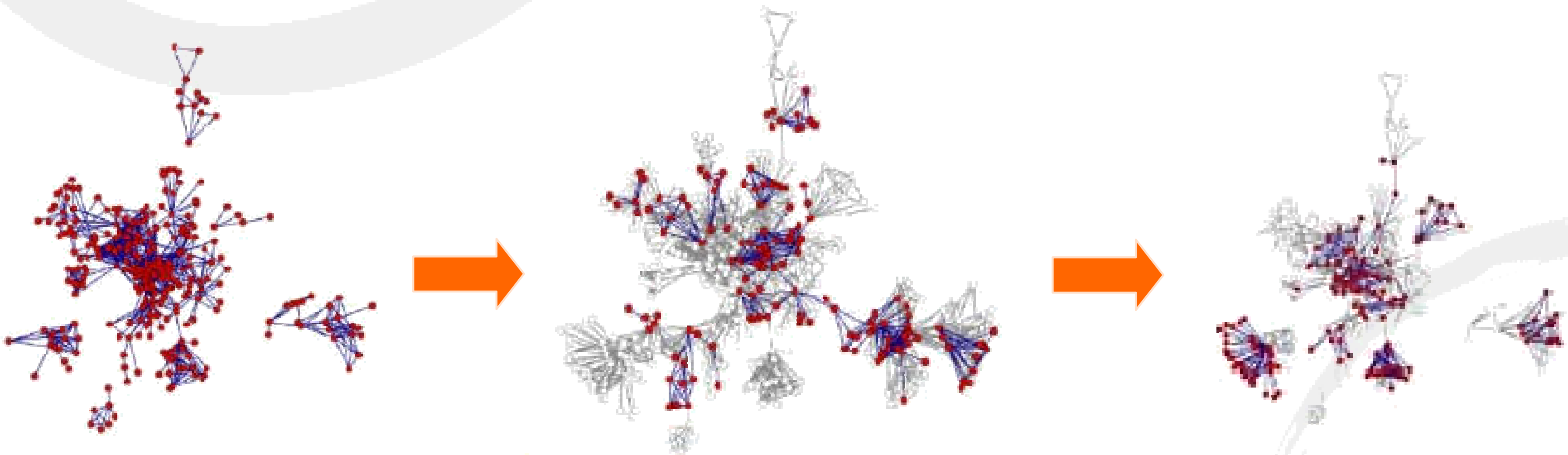


(a) $w = 0.2$ (b) $w = 0.5$ (c) $w = 0.8$

Clusters over time

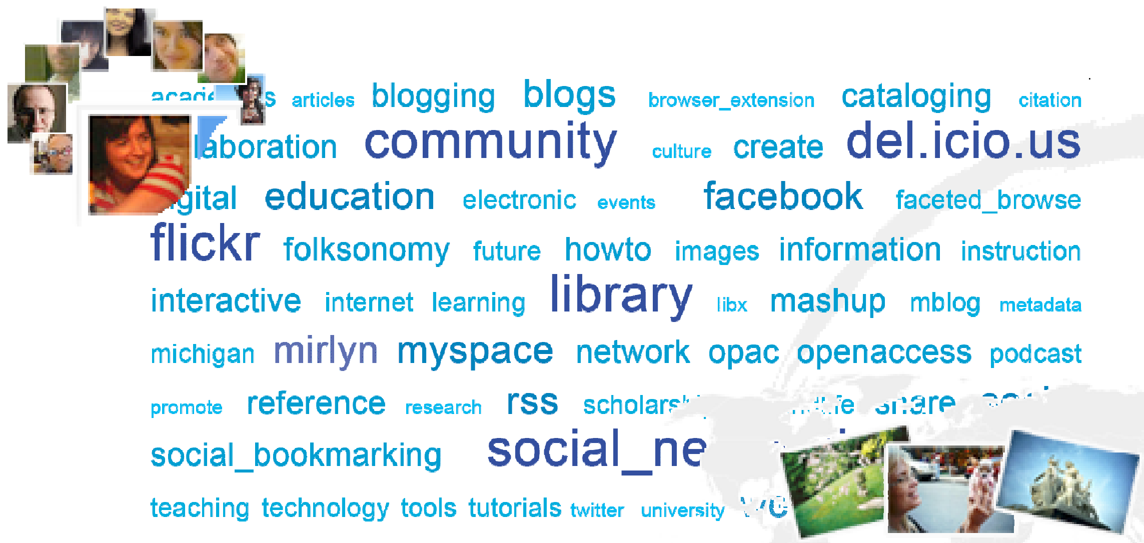
Increase relevance of clusters:

- Temporal relations between users & tags.
- Dominating topics per time slot.



Social Web for improved Machine Learning

Exploiting social annotations for the automatic derivation of concept detectors



Spiros Nikolopoulos

Social annotations → Object Detection

Problem

Manual annotation of content → laborious and time consuming

Object detection schemes require region-detail annotations → hard to create concept classifiers

Solution

Social media sharing (flickr) → image corpora consisting of user tagged images

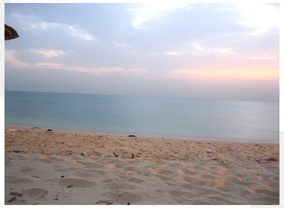
Framework → Training corpus from weakly annotated (tagged) images

Framework concepts

Tagged images

sand, wave, rock, sky

sea, sand

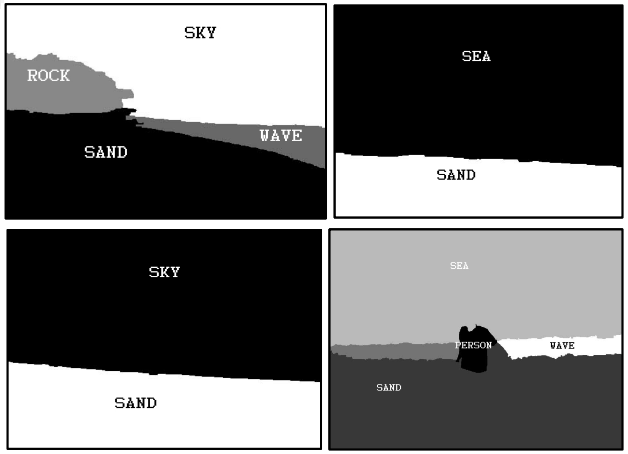


sand, sky

person, sand, wave, see

Social information + **Computer Vision**

Region-detail annotated images

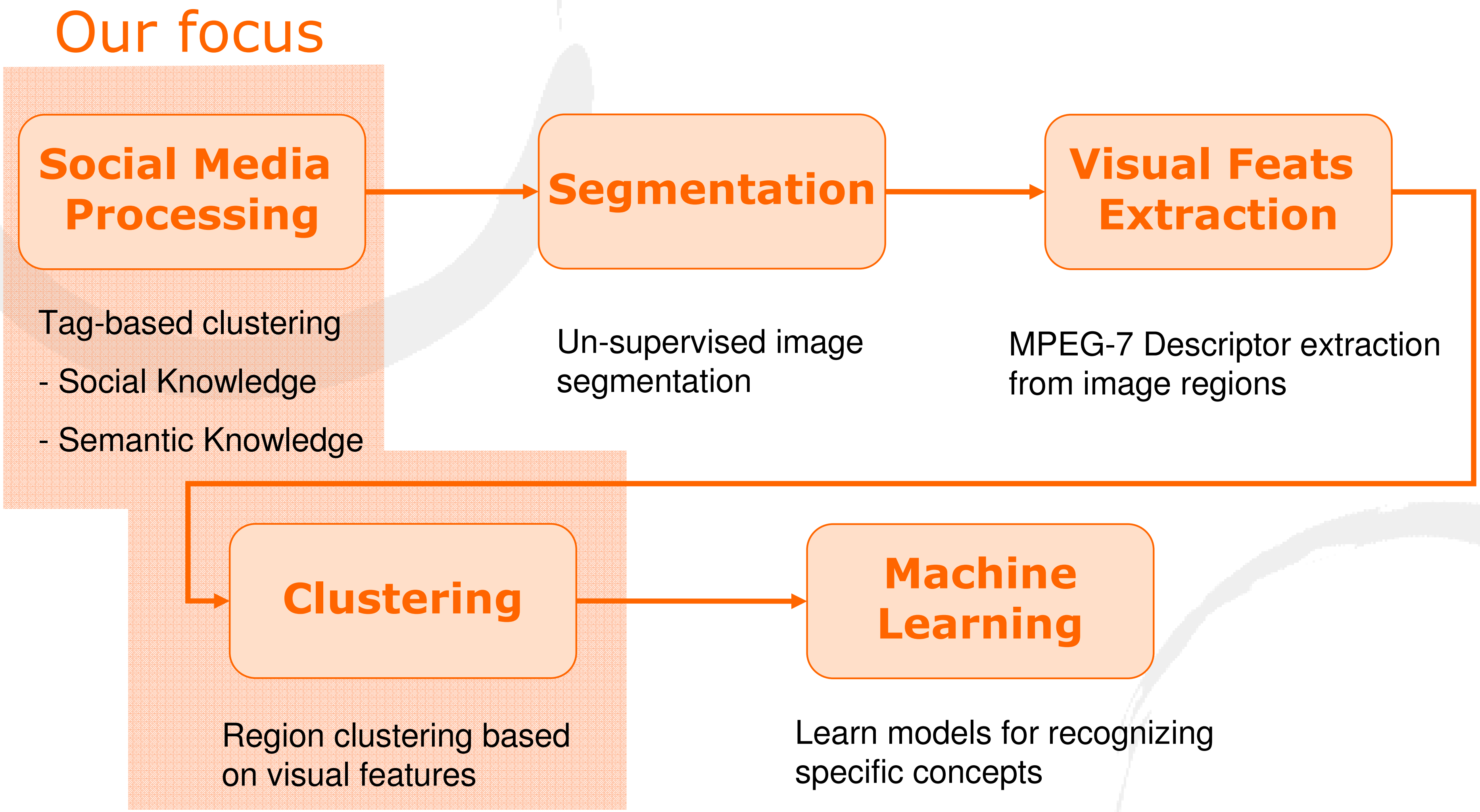


Machine Learning

Models recognizing concepts

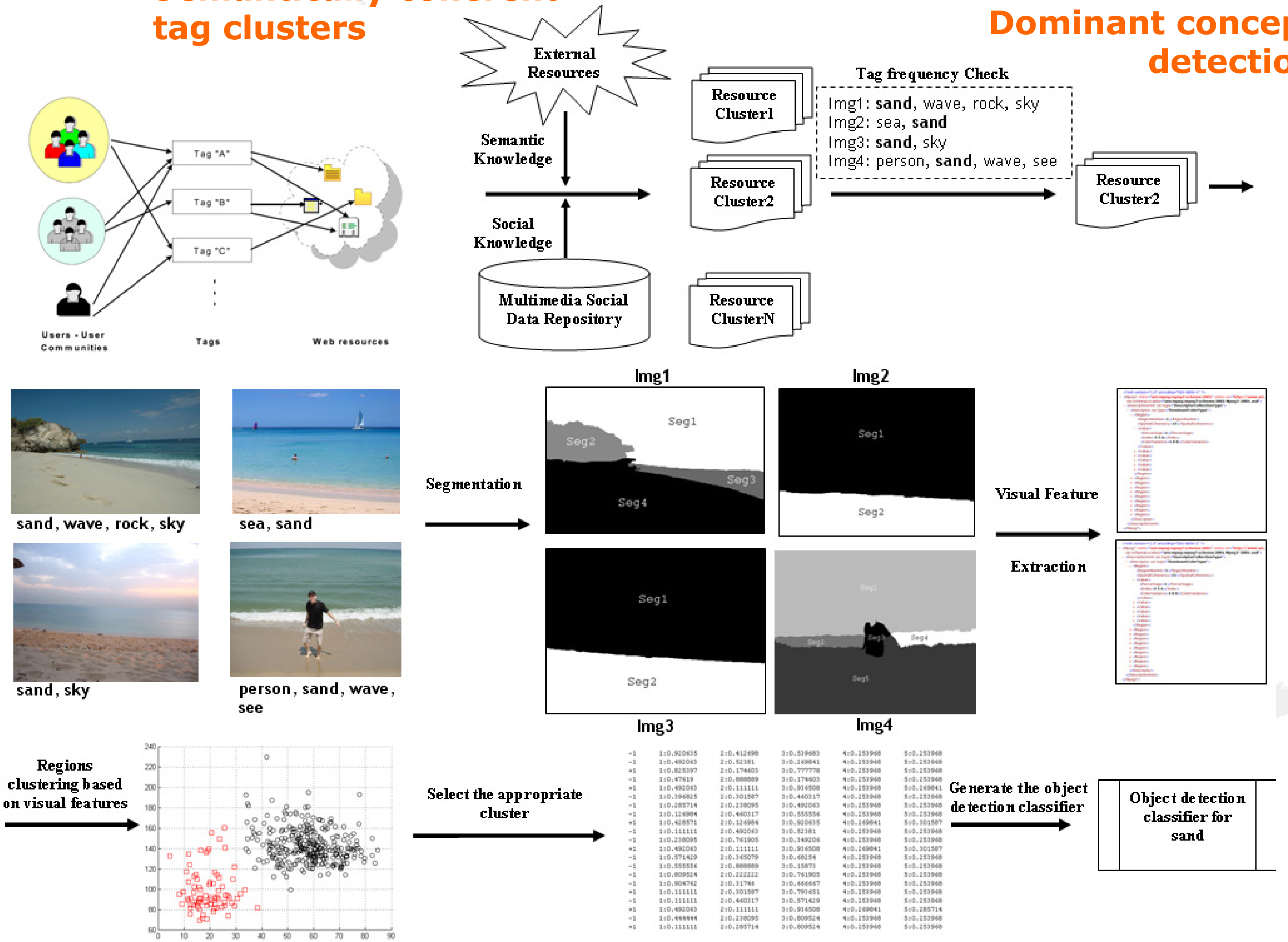
Object Detectors

Analysis Pipeline



Semantically coherent tag clusters

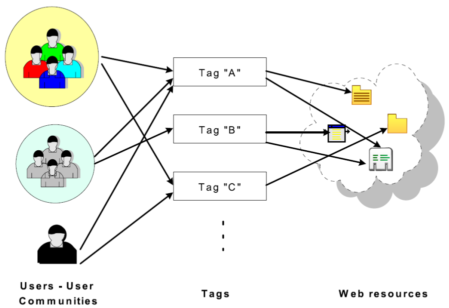
Dominant concept detection



 # Clusters

Cluster selection

Assumption for tag-based processing



Resource cluster of images

C1: Dominant tag

R: # of images

Probabilities that selected image contains a given concept

P1 → C1

P2 → C2

P3 → C3

Assumption: P1 > P2 > P3 >….

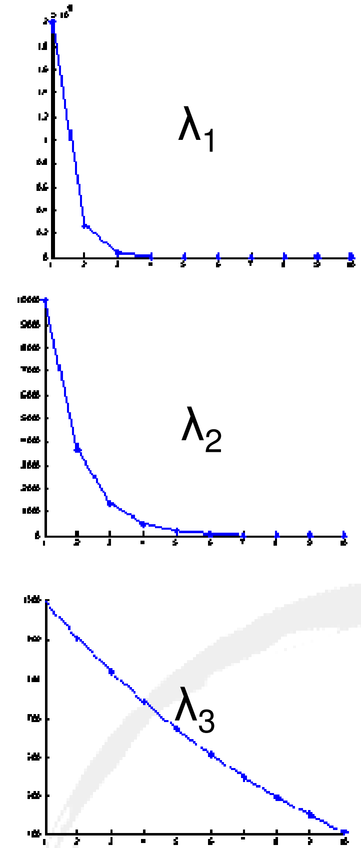
As the volume of the analyzed dataset grows arbitrary big, R increases too

As R increases the regions distribution converges to exponential distribution:

$$f(x) = \lambda e^{-\lambda x}$$

Distribution of regions depicting concepts

C1, C2, C3,…..

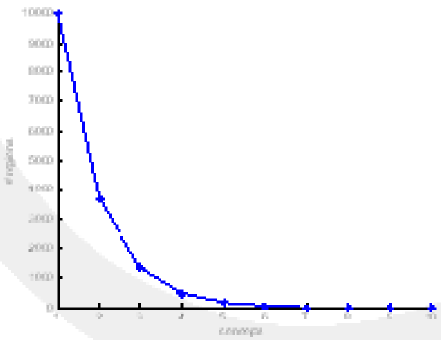


$\lambda_1 > \lambda_2 > \lambda_3$

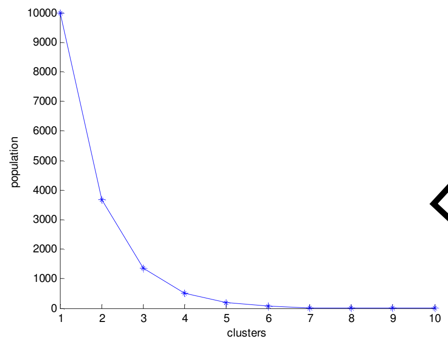
↑**R**

Theoretical Interest → Clustering

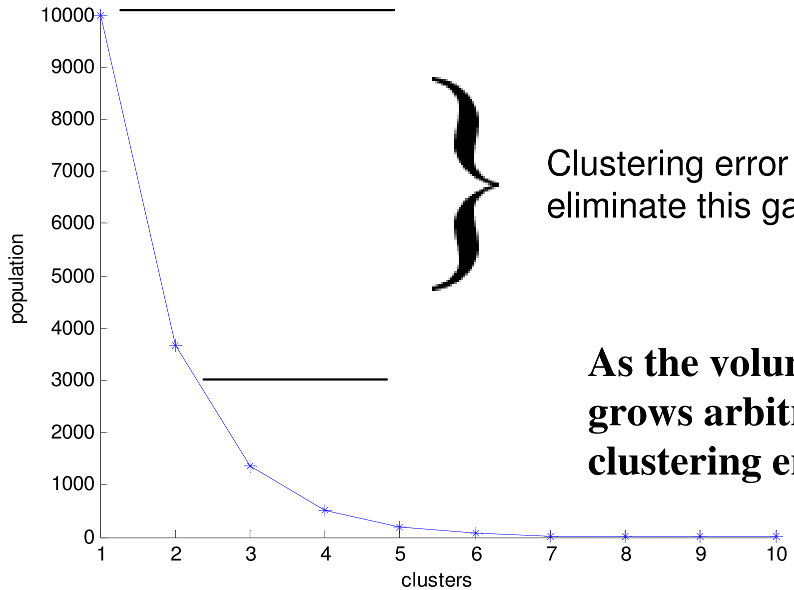
Distribution of regions depicting concepts C1, C2, C3,…..



Distribution of population for clusters CL1, CL2, CL3,…..



When clustering works perfect… these distributions coincide



Clustering error is likely to eliminate this gap

As the volume of the analyzed dataset grows arbitrary big, the effect of clustering error is minimized.

Goal

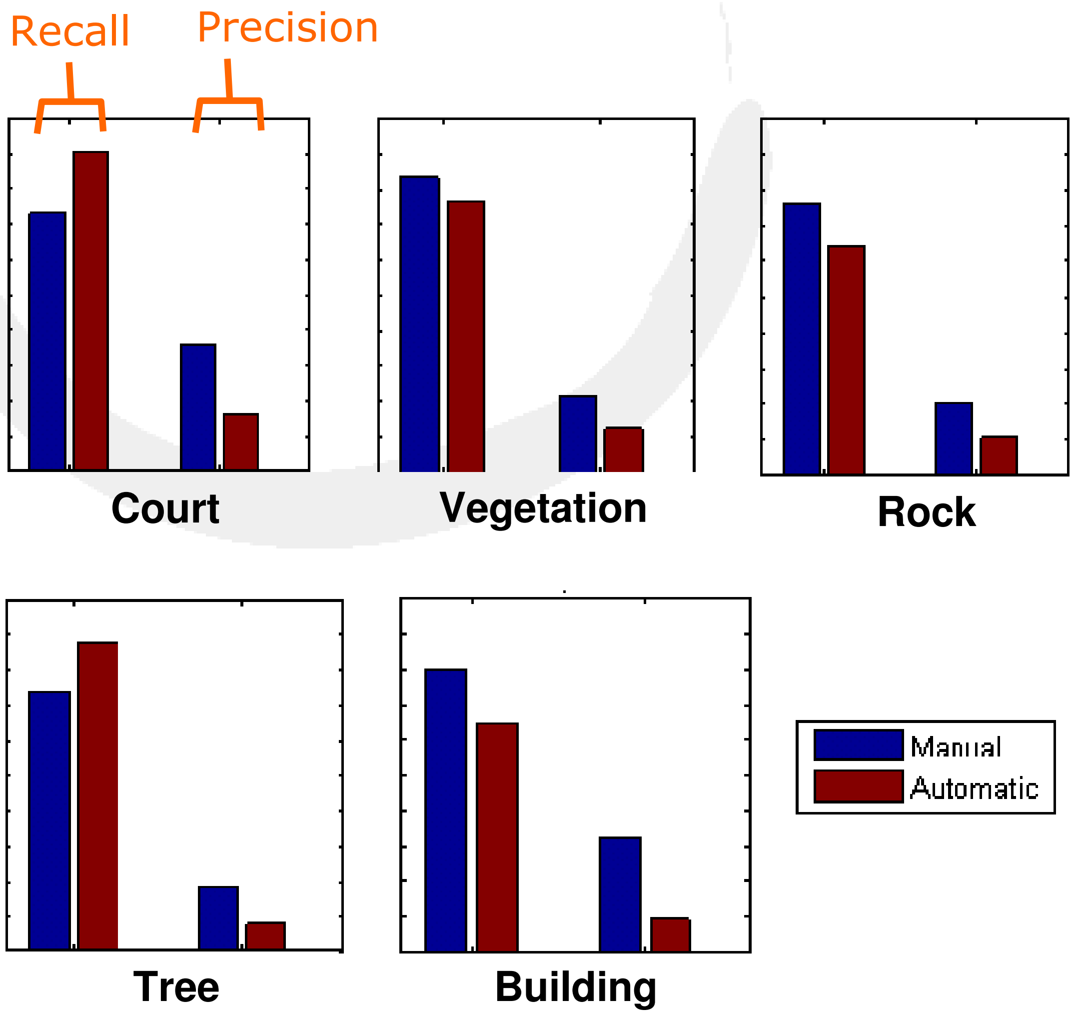
Clustering efficiency → λ and R

Function connecting the population of two clusters i, j with λ:

$$\frac{Pop(i)}{Pop(j)} = \frac{\mathrm{Re}(i)}{\mathrm{Re}(j)} \frac{\mathrm{Pr}(j)}{\mathrm{Pr}(i)} e^{-\lambda(i-j)}$$

Derive useful conclusions regarding the required clustering efficiency with respect to the utilized R

Experimental Results



Setting

- Train:
  - Manual: 300 annotated pics
  - Auto: 3000 tagged flickr pics
  - Test: 300 ground truth pics
- Build detectors using train
- Build detectors using Flickr
- Use test set for evaluation

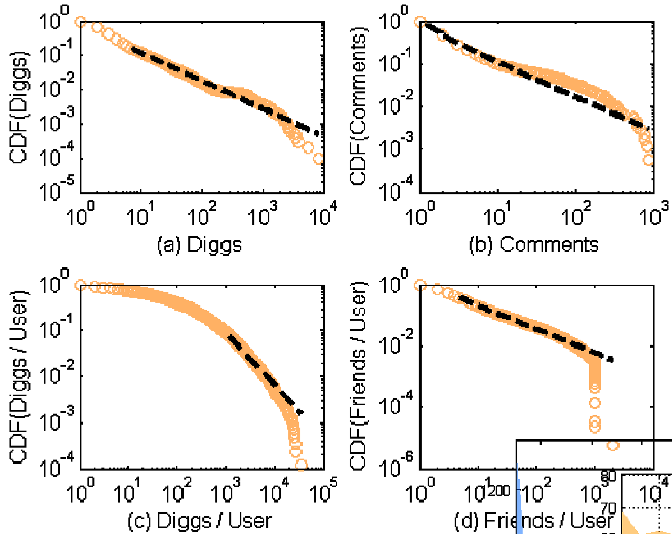
Conclusion

Comparable performance across a range of concepts

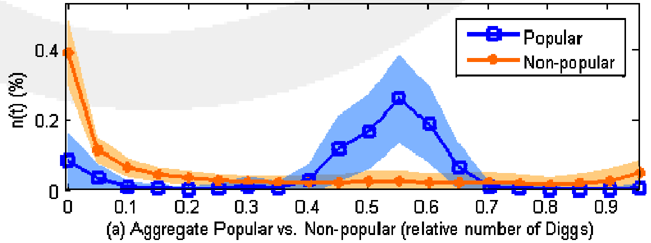
Applications @ MKLab
Case studies & prototypes

Content Popularity: Digg

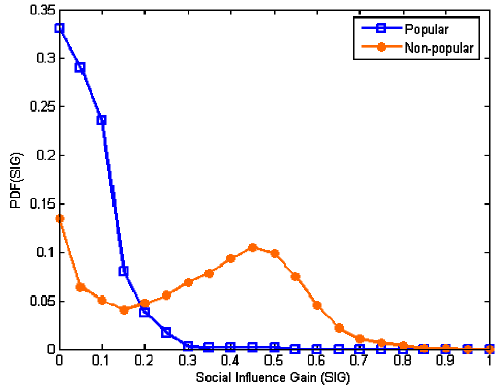
Power-law phenomena

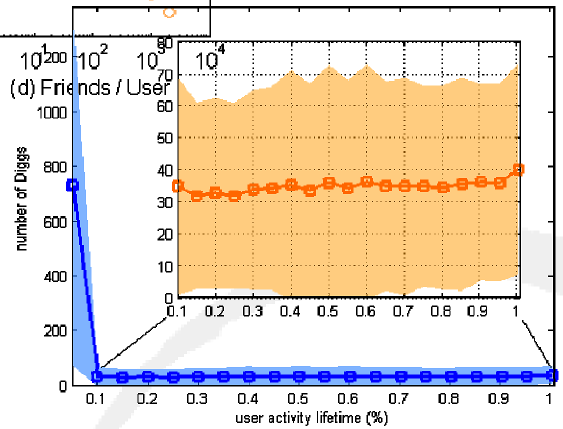


Temporal evolution



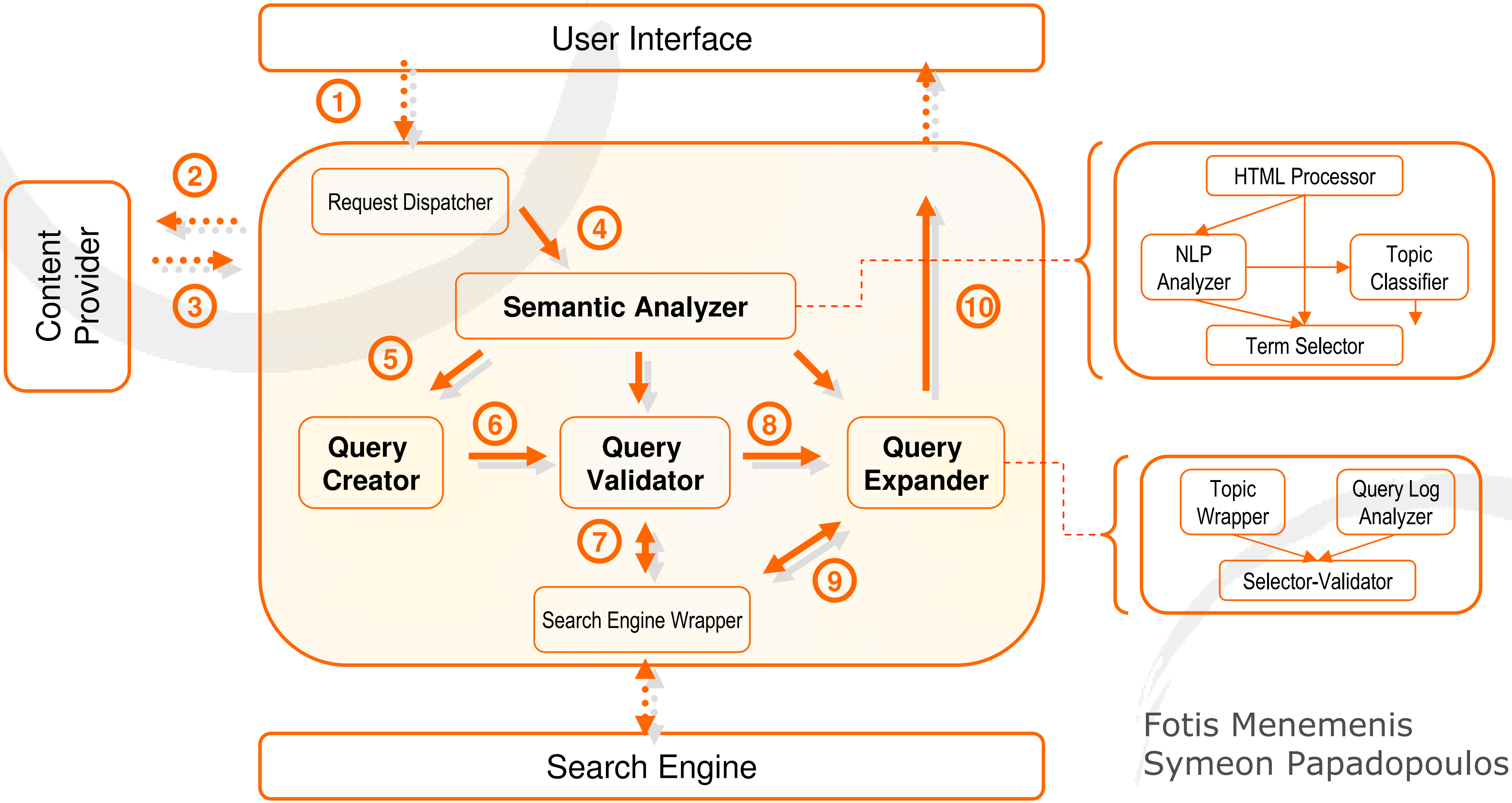
Social Influence





Symeon Papadopoulos
prof. Athena Vakali

Context-based Query Formulation

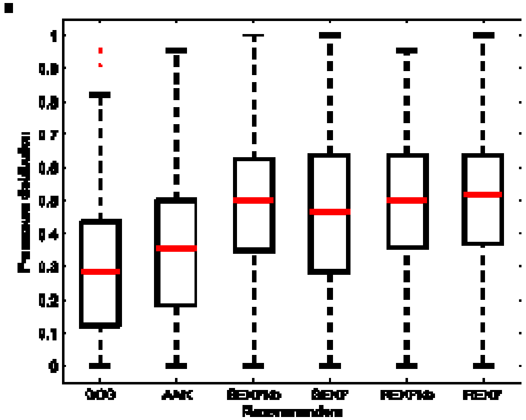


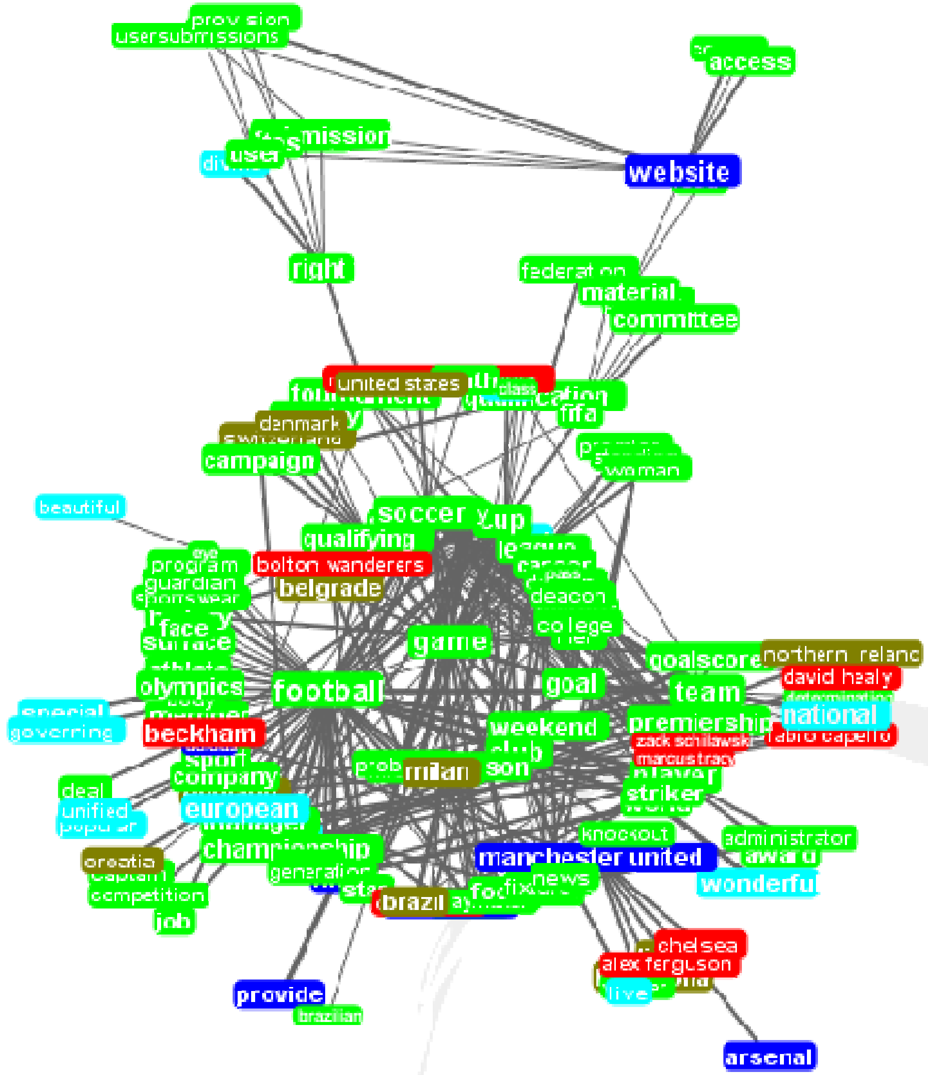
Fotis Menemenis
Symeon Papadopoulos

Advertisement Recommendation based on Lexical Graphs

Vocabulary impedance coupling by means of lexical graphs.

Exploit lexical affinities on the web for ad vocabulary expansion.





Symeon Papadopoulos
Fotis Menemenis

Semantic Retrieval: Reach

Cultural Heritage Collections

- Paintings, inscriptions, coins

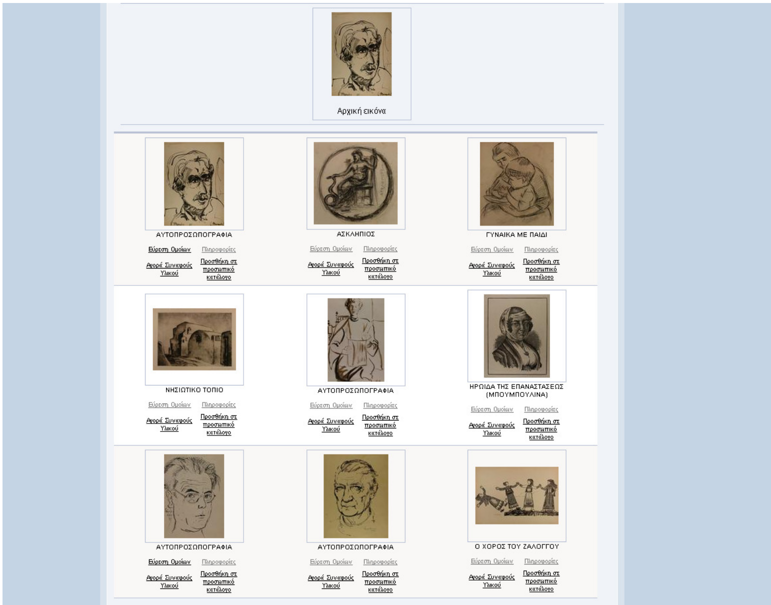
Content-based retrieval

Ontology-based retrieval

- Facets, Semantic Links

Hybrid retrieval

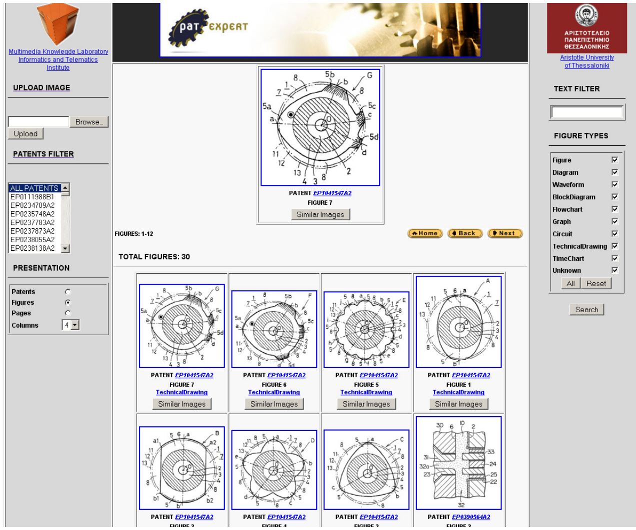
- Item recommendations



http://195.251.117.128/reach/search.html

Stefanos Vrochidis

PAT-MEDIA: Patent Image Retrieval @ PATExpert



Automatic Image extraction

- OCR
- Page Segmentation
- Page Orientation Detection

Patent Images

- Flowcharts, technical drawings
- Binary Content-based retrieval

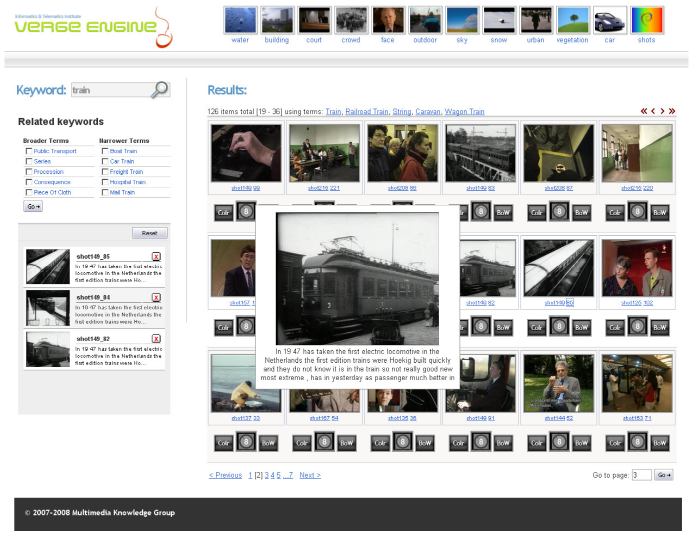
- Text-based retrieval
- Ontology-based retrieval

http://mklab-services.iti.gr/patexpert/

Stefanos Vrochidis
Symeon Papadopoulos

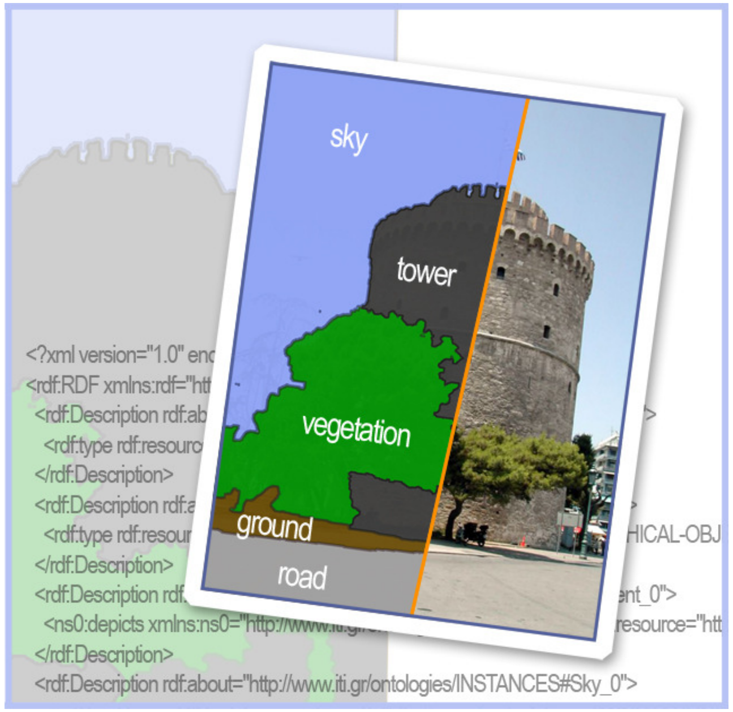
VERGE: Video search engine

- TRECVID collections
- Key frame selection
- Content-based retrieval
- Text-based retrieval
- High level concept retrieval
- Basket for results storing
- Participation
  - TRECVID
  - VideOlympics



http://mklab-services.iti.gr/trec2008/

Stefanos Vrochidis



Thank you!

http://mklab.iti.gr