social sensor

Project Introduction
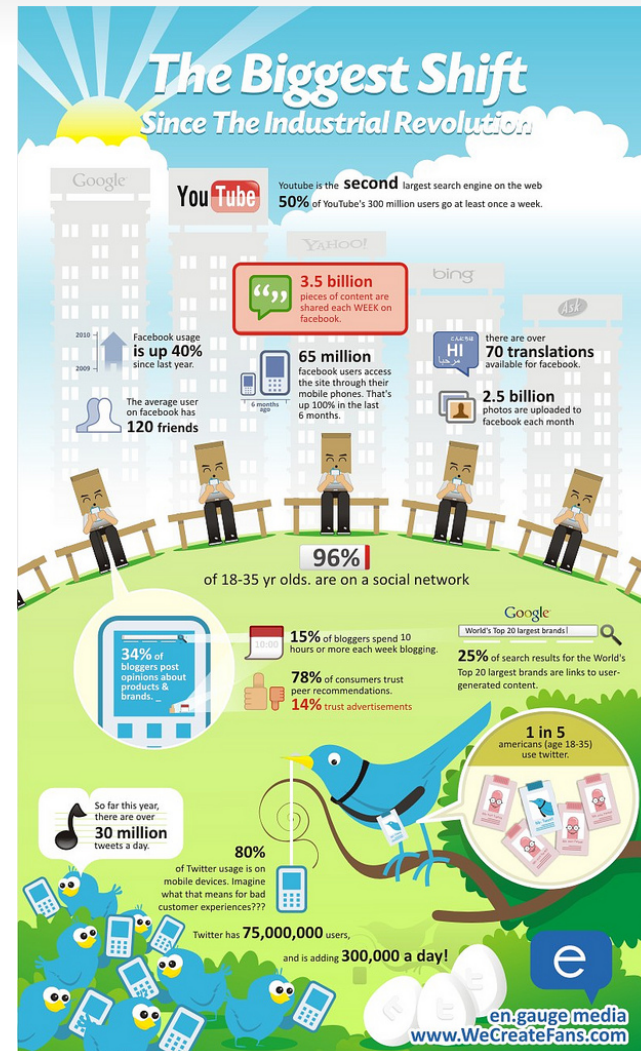
Dr. Yiannis Kompatsiaris (CERTH)

AV Media Search Cluster Meeting

Brussels, 13 December 2011

Social networks and media

- Users upload, tag, share, connect and search

- Emphasis is on uploading, visualization of results and interfaces

- Single media item analysis

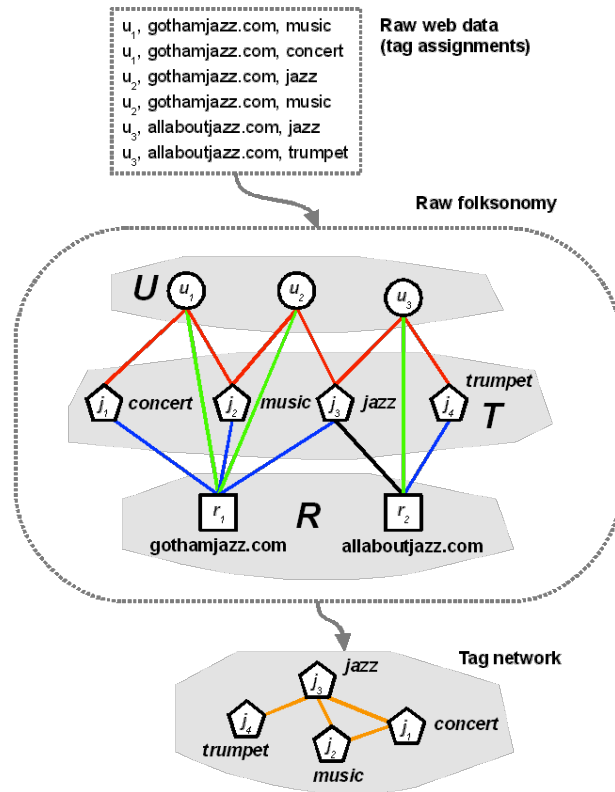- Limited usage of the **Collective and Multimedia nature** of Social Networks
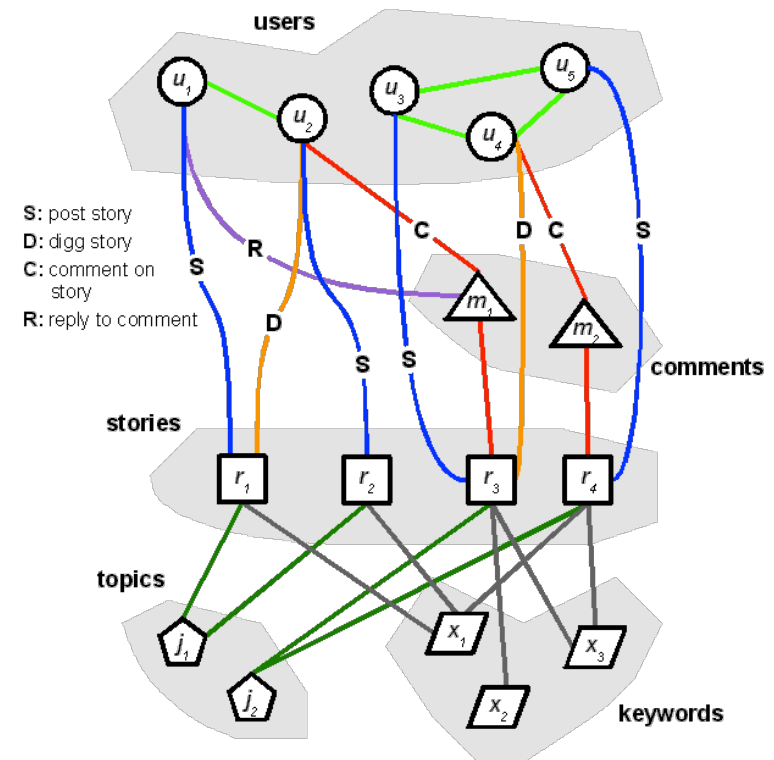
Social Media + Context

Social Graph + Context

Folksonomy (Delicious)

Mika, P. (2005) Ontologies Are Us: A Unified Model of Social Networks and Semantics. Proceedings of the 4th International Semantic Web Conference (ISWC 2005), Springer Berlin / Heidelberg, pp. 522-536

MetaGraph (Digg)

Lin, Y., Sun, J., Castro, P., Konuru, R., Sundaram, H., and Kelliher, A. (2009) MetaFac: community discovery via relational hypergraph factorization. Proceedings of KDD '09, ACM, pp. 527-536

Social Networks as Sensors

- Social Networks is a data source with an extremely dynamic nature that reflects events and the evolution of community focus (user's interests)

- Web 2.0 data consists of **individually rare** but **collectively frequent** events and topics

- Potential for much more if we mine the data and their relations and exploit them in the right context

  - Scalable approaches taking into account the content and social context of social networks

- Search and Discovery of meaningful topics, entities, points of interest, social connections and events

- Rather than search for **isolated** or directly connected **social media items**

Example Applications

Xin Jin, Andrew Gallagher, Liangliang Cao, Jiebo Luo, and Jiawei Han. ***The wisdom of social multimedia: using flickr for prediction and forecast,*** International conference on Multimedia (MM '10). ACM.
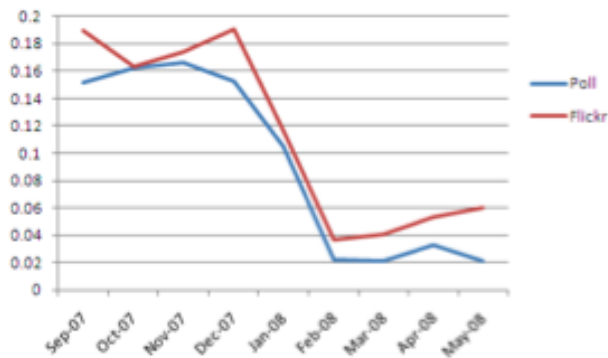

Figure 7: Reuters/Zogby Poll v.s. Flickr. Y-axis denotes the percentage of popularity for candidate Edwards.

Federal Emergency Management Agency ***plans to engage the public*** more in disaster response by sharing data and leveraging reports ***from mobile phones and social media***

Gogobot: Travel Discovery Goes Social And Visual "The service raised $4 million in funding (Google CEO Eric Schmidt is one of the investors)…This is a $100 billion a year industry in the U.S. It's something like $350 billion worldwide."

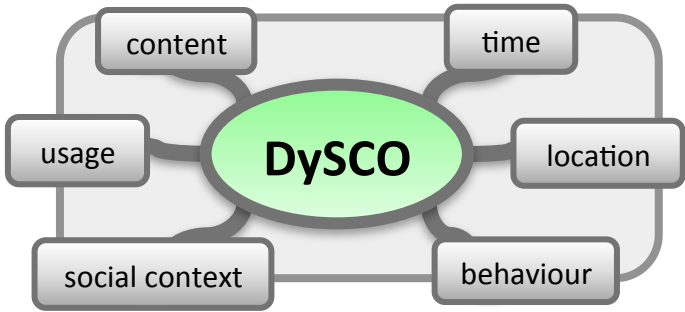# SocialSensor: Objective

Novel approach for mining and making searchable massive amounts of media from multiple social sources in **real-time.**

Social media mining
Aggregation & indexing

content   time

usage   **DySCO**   location

social context   behaviour

Massive social media
and unstructured web

Personalised access
Ad-hoc P2P networks

News - Infotainment

social sensor

Use Cases

- ## News

  - UGC news material immense & fast-paced:

    - Japan earthquake: 1.2K tweets/min → 1.7M tweets/day
    - US elections prediction using Flickr, equivalent to official polls

      Discover trends and topics and create articles with **high coverage** of social content **on time**

- ## Infotainment

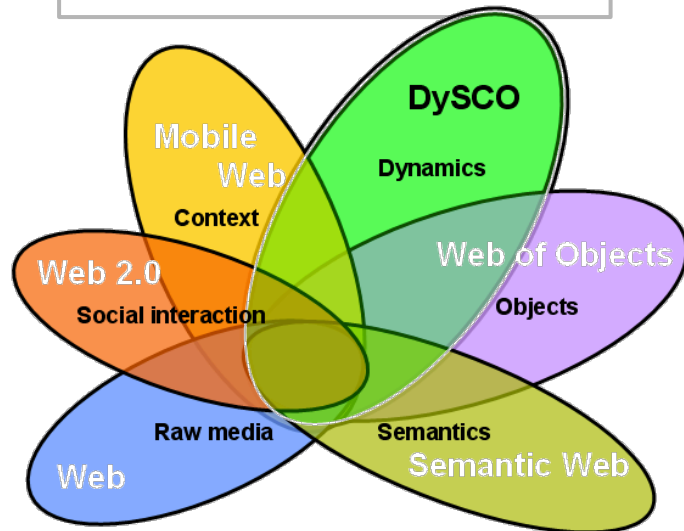 - Large events attract huge numbers of people:

    - Sziget music festival (Budapest) > 350K people, 5K YouTube videos
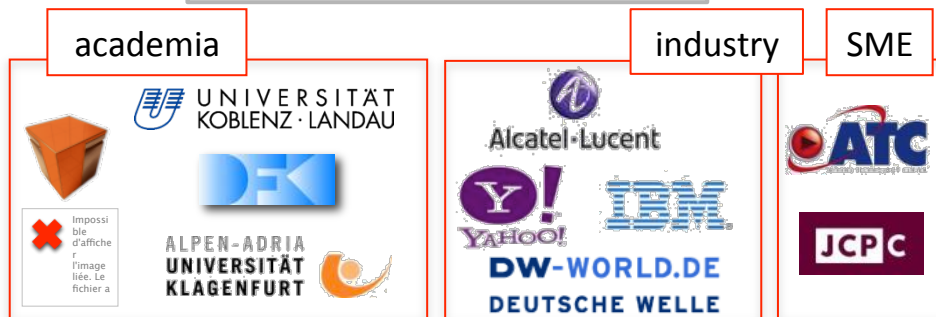   - CeBIT (Hannover) > 330K visitors, 57K Flickr photos

      Satisfy **personalised** media search and recommendation needs in **ad hoc P2P** networks
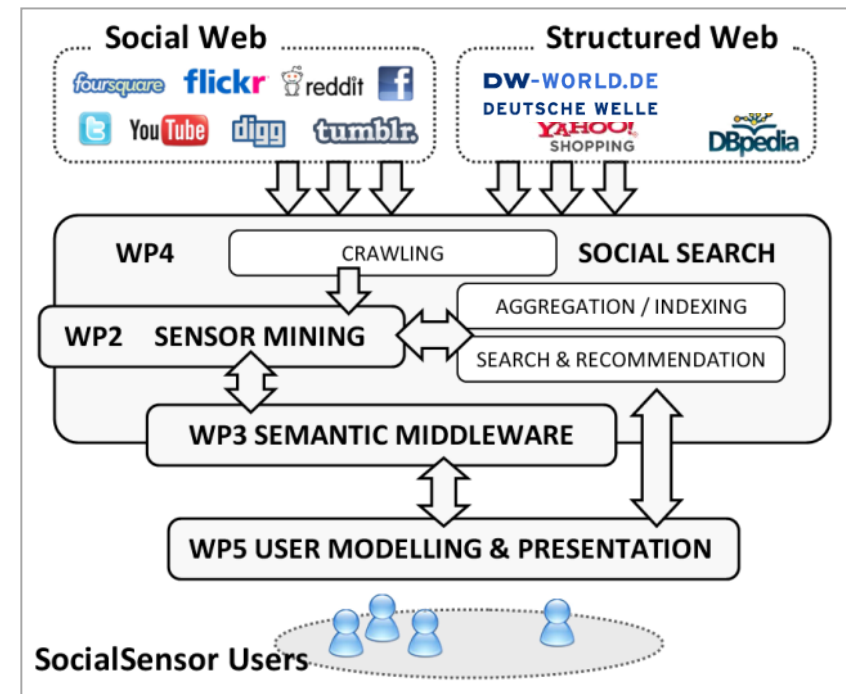
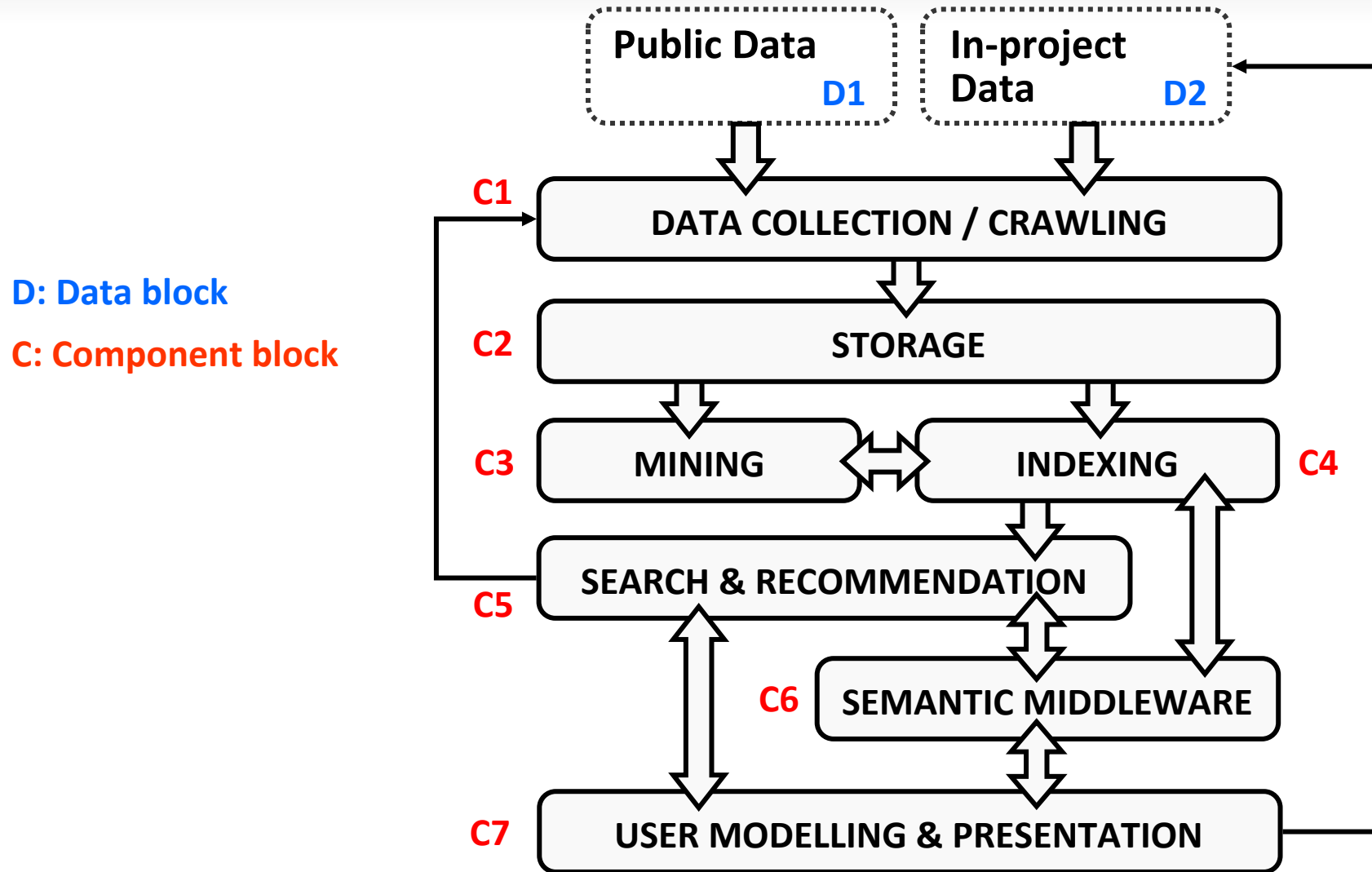# Project Setup and Technical Approach

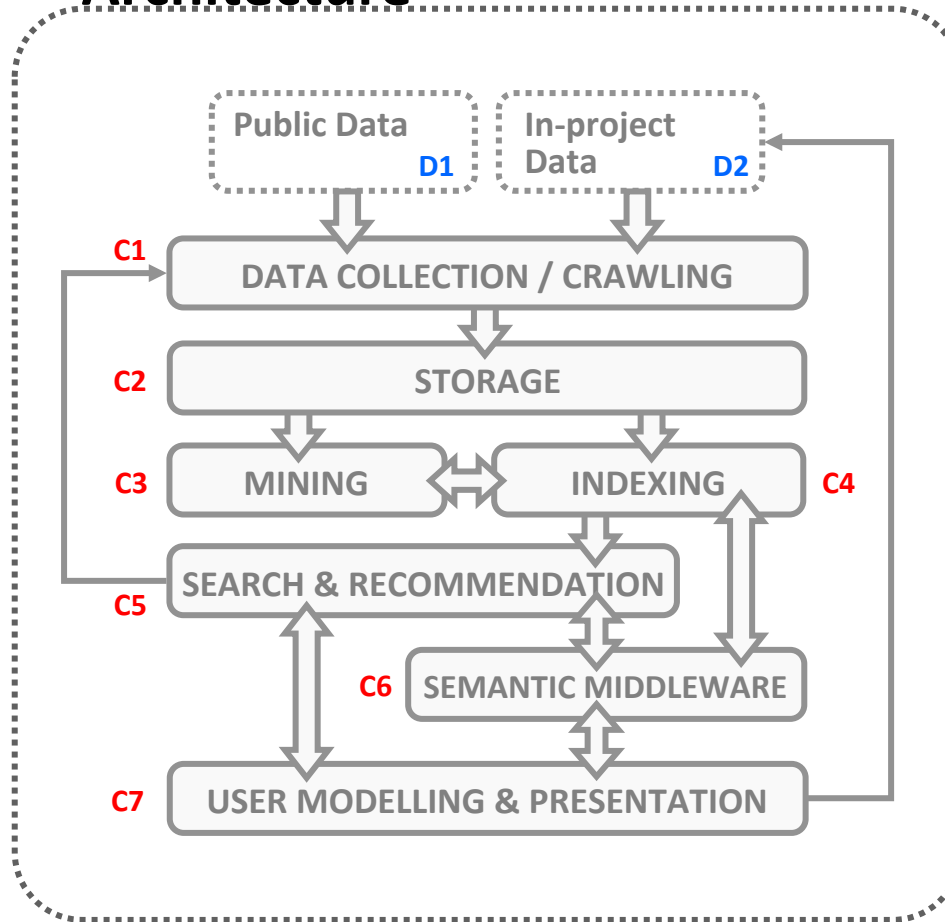SocialSensor environment

Consortium: 11 partners

academia

industry

SME

WP structure &
conceptual architecture

Conceptual Architecture

D: Data block

C: Component block

Public Data **D1**

In-project Data **D2**

C1 DATA COLLECTION / CRAWLING

C2 STORAGE

C3 MINING

C4 INDEXING

C5 SEARCH & RECOMMENDATION

C6 SEMANTIC MIDDLEWARE

C7 USER MODELLING & PRESENTATION

Conceptual Architecture

Architecture

Architecture

- Public Data **D1**
- In-project Data **D2**

C1 DATA COLLECTION / CRAWLING

C2 STORAGE

C3 MINING ⇄ INDEXING **C4**

C5 SEARCH & RECOMMENDATION

C6 SEMANTIC MIDDLEWARE

C7 USER MODELLING & PRESENTATION

Use Cases

Application development
UI & app logic
Domain-specific functionalities

NEWS

Application development
UI & app logic
Domain-specific functionalities

INFOTAINMENT

Sensor mining

- **The data model of the whole system with the first building blocks containing the necessary information for feeding the other components**

- The first level for building the DySCOs objects

- Data analysis for knowledge extraction from data streams
  - Topic detection and evolution
  - Social knowledge
  - Opinion mining
  - Information quality and trust computation

- Leveraging social networks: both content, structure, and temporal

- Linkage of social networks to other sources (traditional Web like Wikipedia, news sites: DW, BBC, WorldNews, digg.com…)
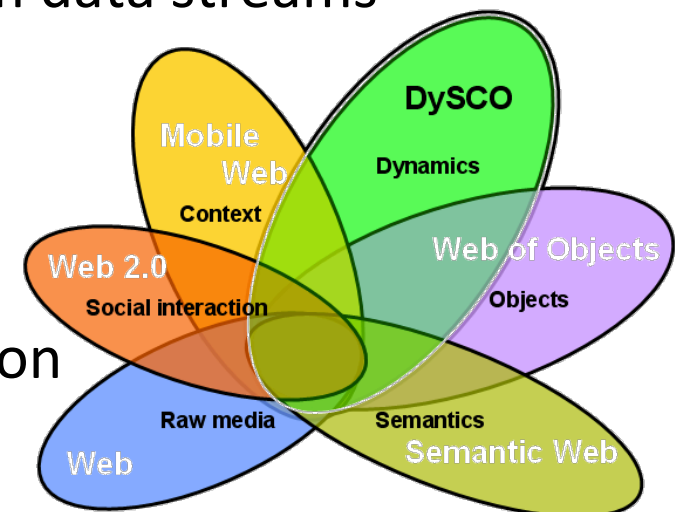
Semantic Middleware

- Semantic p2p service coordination
  - Peer-to-peer discovery and composition of semantically relevant data in changing contexts
    - Semantic data and service discovery
    - Semantic service composition

- Continuous semantic querying:
  - Semantic inconsistency handling
  - Context-based semantic query answering

- Context aware media caching and prefetching
  - Disconnection prediction
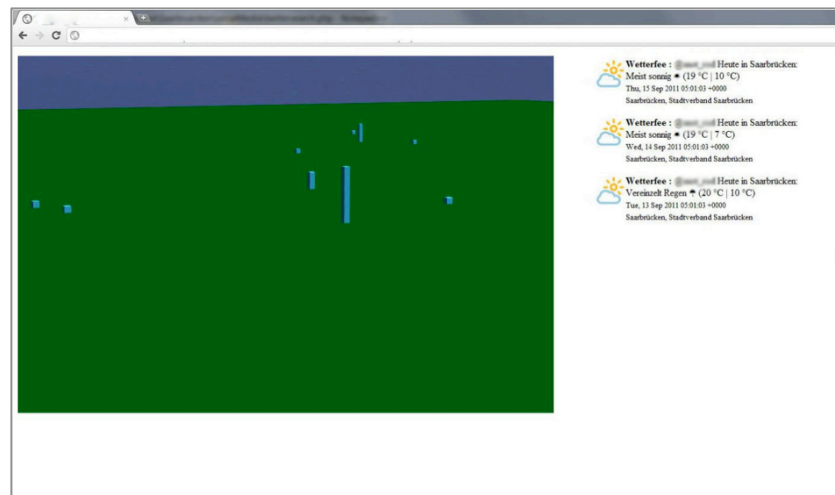 - Prefetching based on context and access point
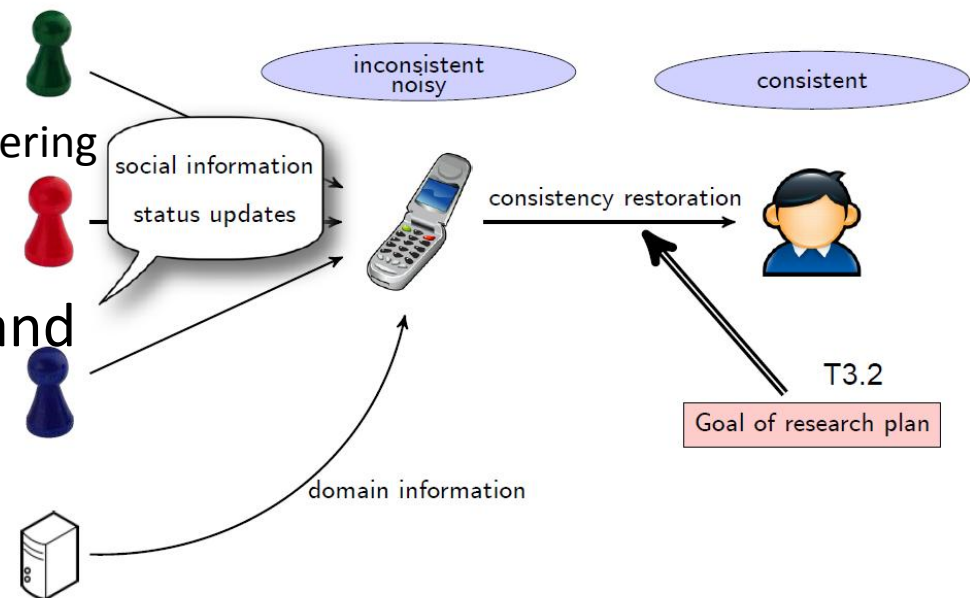
Social Search

- Crawling of social and Web sources
  - Social crawler
  - Real-time aspects (handle updates and user profiles)
- Web analytics:
  - Pattern extraction from query logs
- Context-aware search
  - Use context modelling
- Multimedia preprocessing and feature extraction
  - Fast extraction of sufficiently descriptive features from media
- Aggregation and indexing
  - Social Indexing
- Multimedia retrieval and filtering
  - Images/video and social media content around them

User Modelling and Presentation

Community snapshot

User snapshot

- User and context modeling
  - Preference reasoning and elicitation
  - User profiles

- Personalization by user profiles and
social relations
  - User profiling and preference prediction
  - Mobile exploration of social media

- Intelligent presentation and user
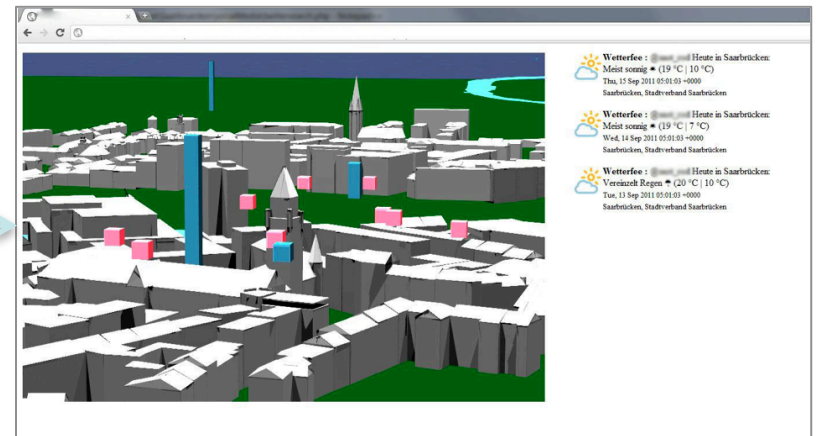interaction
  - 3D virtual environments

GeoNames +

+ eventful +

DBpedia +

+flickr

upcoming.org

+ …

???

???
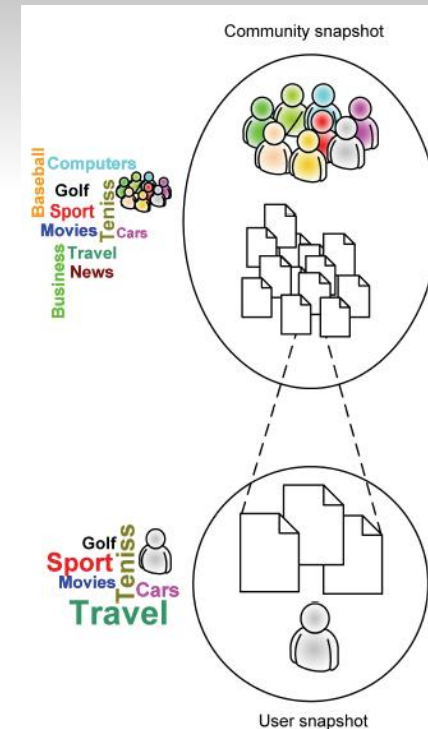
3D Interaction example

... integrated in a XML3D city model with Flickr images

... integrated in a XML3D city model with 2D weather data

Geo located data of Twitter
–> here number of tweets on a specific location

social sensor

Image based search "within" social networking sites (flickr)

http://www.clusttour.gr

clusttour   blog   about   data   faq   disclaimer   Sign In

La Sagrada Familia
Barcelona

Location

expand   Map Satellite
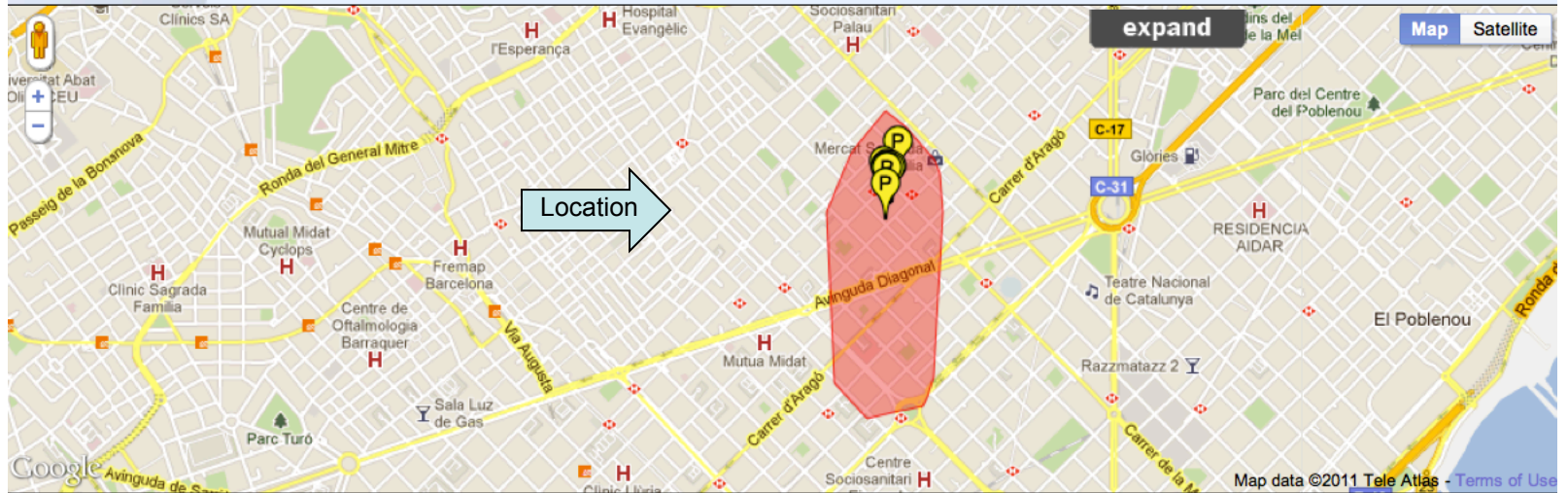
Details

Area Name
La Sagrada Familia

Total Photos
12872

Total Users
1476

Most popular Months
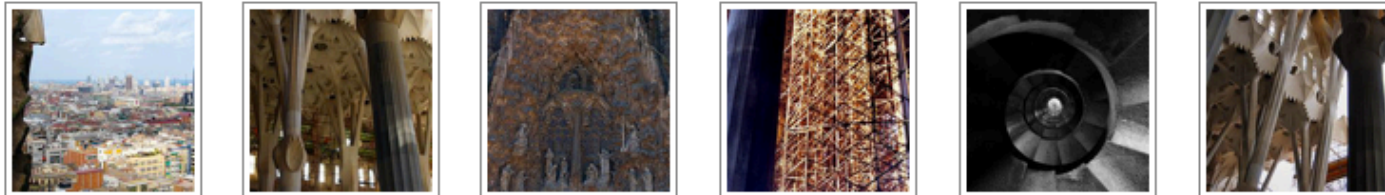May, Aug

Most popular season
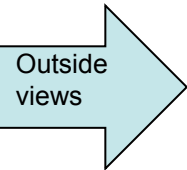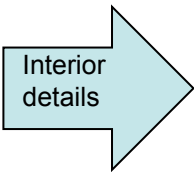Summer

Social

Time

Top Clusters

1. Sagrada Familia

social sensor

SEVENTH FRAMEWORK PROGRAMME

Top Clusters

1. Sagrada Família

2. Sagrada Família

Interior details
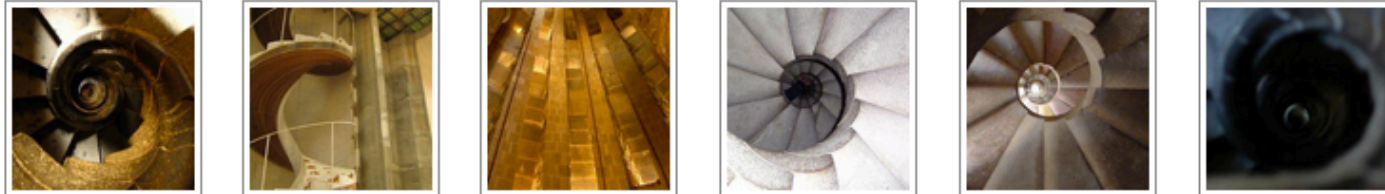
3. Sagrada Familia

4. Sagrada Familia

Outside views

social sensor

ClustTour   By CERTH-ITI

Open iTunes to buy and download apps.

Free

Category: Travel
Released: Dec 13, 2011
Version: 1.0.0
Size: 1.9 MB
Language: English
Seller: Centre for Research and
Technology Hellas
© 2011 Centre for Research
and Technology Hellas
Rated 12+ for the following:
Infrequent/Mild Sexual Content
or Nudity
Infrequent/Mild Realistic
Violence
Infrequent/Mild Horror/Fear
Themes
Infrequent/Mild Cartoon or
Fantasy Violence
Infrequent/Mild
Mature/Suggestive Themes
Infrequent/Mild Alcohol,
Tobacco, or Drug Use or
References

Description

ClustTour is a better way to search, discover and browse interesting city areas, POIs and events. Whether you are planning a trip or just want to check out how a place looks like, ClustTour offers a large collection of photos, maps and descriptions. You may wonder, what's new in this? ClustTour is not based on "official" guides and "experts" but

CERTH-ITI Web Site ▸   ClustTour Support ▸   ...More

iPhone Screenshots

THESSALONIKI INTERNATIONAL FILM FESTIVAL

One of the top 20 most important film festivals

77% average venue capacity
91% for the International Competition section

80,000 viewers / 100,000 visitors in 10 days
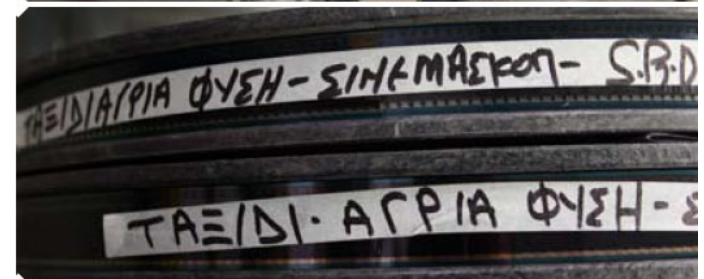
150 films
350 screenings
Screenings in 24 cities around Greece
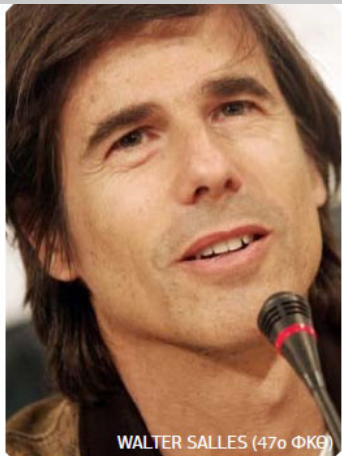Parallel events (exhibitions, workshops, etc.)

250 Staff members
350 Volunteers (applications: 1,300)
900 Accredited Guests

FRANCIS FORD COPPOLA (46ο ΦΚΘ)

WALTER SALLES (47ο ΦΚΘ)

JULIETTE LEWIS (47ο ΦΚΘ)

LILI TAYLOR (47ο ΦΚΘ)

WIM WENDERS (47ο ΦΚΘ)

COSTA GAVRAS (47ο ΦΚΘ)

VALERIA GOLINO (47ο ΦΚΘ)

JOHN MALKOVICH (48ο ΦΚΘ)

JOHN SAYLES (48ο ΦΚΘ)
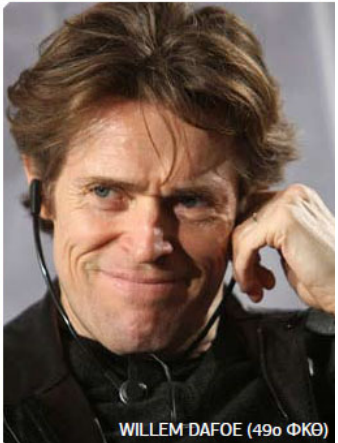
WILLIAM KLEIN (48ο ΦΚΘ)
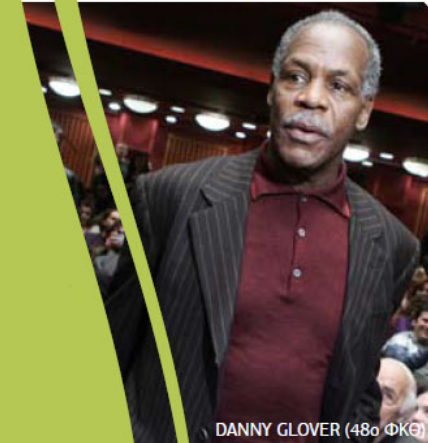
DANNY GLOVER (48ο ΦΚΘ)

SAM ROCKWELL (48ο ΦΚΘ)

LUC DARDENNE (49ο ΦΚΘ)

TAKESHI KITANO (49ο ΦΚΘ)

WILLEM DAFOE (49ο ΦΚΘ)

Social media presence

facebook.com/filmfestival.gr

twitter.com/filmfestivalgr

News Use Case Requirements

- **AWARENESS**

- **FACTS**

- **TRUTH & ACCURACY**

- **VERIFICATION**

- **CLEVER**

Challenges

- **Multi-modality**: e.g. image + tags, image + video

- **Rich Social Context**: spatio-temporal, social connections, relations and social graph

- **Inconsistent quality**: noise, spam, ambiguity

- **Huge volume**: Massively produced and disseminated

- **Multi-source**: may be generated by different applications and user communities

  - Also connected to other sources (e.g. LOD, web)

- **Dynamic:** Fast updates, real-time

Measures of Success and Exploitation

- Impact to our users
  - SocialSensor system part of daily workflow in DW and other news agencies
  - Applications used by TIFF and other large events
  - Number of end users participating in the use cases
- Research and Technological impact
- Economic impact
  - Provide added value to news and infotainment sector
- Exploitation by partners and joint ventures

Collaboration with other projects

- Exchange of data sets
  - Use CHORUS+ wiki to share datasets

- Access to users and use cases

- Common technology approaches
  - Multimedia analysis and features
  - Crawling
  - Storage, architectures for large data

- Bencmarking
  - MediaEval, Social Event Detection Task

Intl. Workshop on Social Media Applications in News and Entertainment

SMANE 2012

16 April 2012, Lyon, France

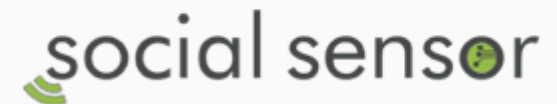## Intl. Workshop on Social Media Applications in News and Entertainment

Monday April 16th, 2012, Lyon, France

The aim of this workshop is to encourage discussion and sharing of ideas and research results on social media research, techniques, and applications, in two crucial areas: News and Entertainment. The workshop offers an opportunity to promote interdisciplinary research and exchange of ideas in this area, not only between industry and academia, but also between different fields (e.g., computer science, journalism, psychology, sociology, economics, history, cultural anthropology, business, etc.). These two application areas have received the most impact from social media technologies, and yet, there are many technical and social challenges.
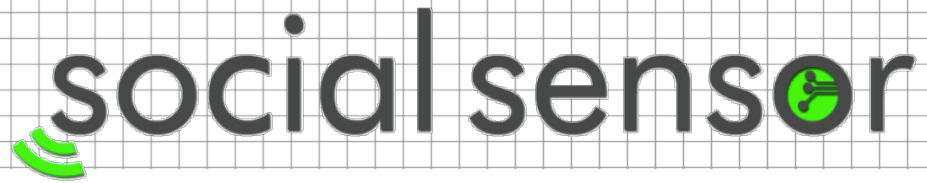
Co-located with

WORLD WIDE WEB
2012
LYON - FRANCE

Supported by

social sensor

- http://smane2012.socialsensor.eu
- Deadline for papers: February 1st, 2012

social sensor

SEVENTH FRAMEWORK
PROGRAMME

Thank you!

http://www.socialsensor.eu