



Search Computing: Business Areas, Research and Socio-Economic Challenges

Yiannis Kompatsiaris, Spiros Nikolopoulos

Informatics and Telematics Institute, CERTH--ITI

NEM SUMMIT

Torino-Italy, 28th September 2011

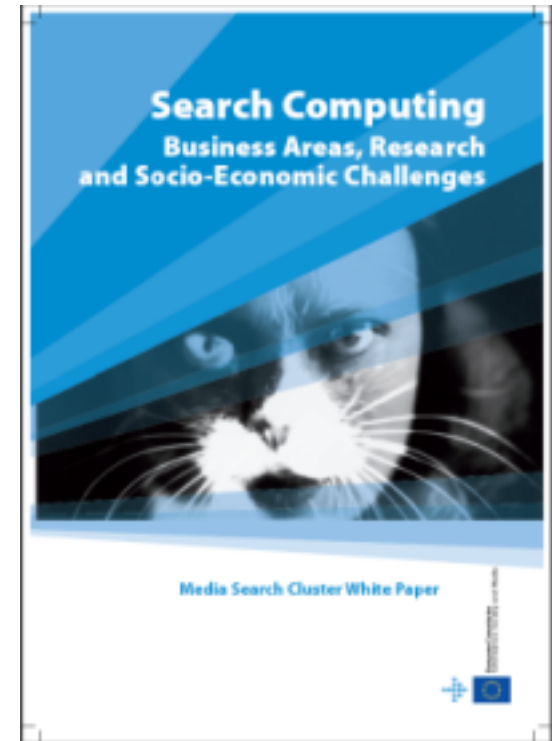
Media Search Cluster “Search Computing” White Paper

Document evolution

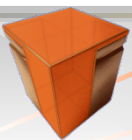
- Draft list of topics and structure:
Media Search Cluster meeting,
14/04/2011, Trento
- First Draft Circulated, 21/04/2011
- Final version available by the EC,
15/09/2011
- 10 versions

Contributors

- 23 individuals
- 12 R&D projects



Avmediasearch.eu



Search Computing

Business areas

- Mobile
- Social
- Enterprise
- Music

Research chal.

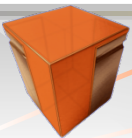
- Multimodal
- Affective
- Large-scale
- Real-time

"Search Computing"

Data not only "found" but also "acted upon"

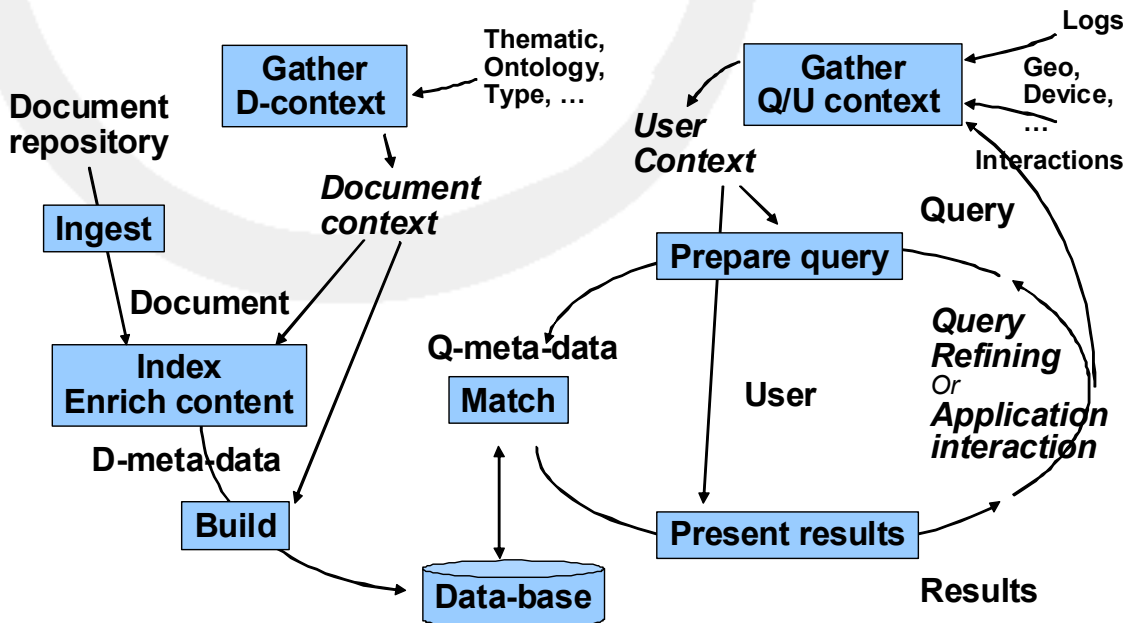
Soc-eco. aspects

- Business models
- Benchmarking
- Innovation
- Legislation



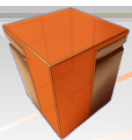
Search engines are changing ...

Search engines have been the primer knowledge broker to the abundant availability of information in the Web



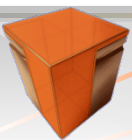
Functional breakdown of a search engine, CHORUS Final Report 2009

- Search engines are becoming both **multimedia** and **metadata** savvy.
- The advances in Enterprise, Social, Mobile and Music search suggest a **fundamental change** of the users' needs in the way they **search** and **consume information**.
- The big **research challenge** is how to progress from today's commercial text and language-based search engines to **multimedia search engines that do more with less**.



Where we are now?

- Text-based
 - Web-scale
 - Fast indexing and retrieval (efficient parallelization)
 - Efficient ranking for web applications
- Image, Music and Video content-based search
 - Content-based similarity (stock photos, mobile search, music)
 - Image processing in web-based search
 - Concept detection
- Mobile
 - Location-based search
- Enterprise
 - Knowledge extraction, semantic analysis
 - Facet results
- Social
 - Search within social networks



more than
7 million
images for sale

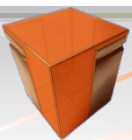
Try our image search!

and find the right stock photos really fast

type keyword or image url

Search

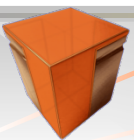
or try searching for stock photos by [uploading an image](#)



www.impalacore.com



using vision TAGS



bing™

roses



Images

Web News Images More▼

Select View: Large Medium Small | SafeSearch: Moderate▼

569,000 results

Clear All

SIZE

LAYOUT

COLOR

- All
- Full color

Black & white

Specific color

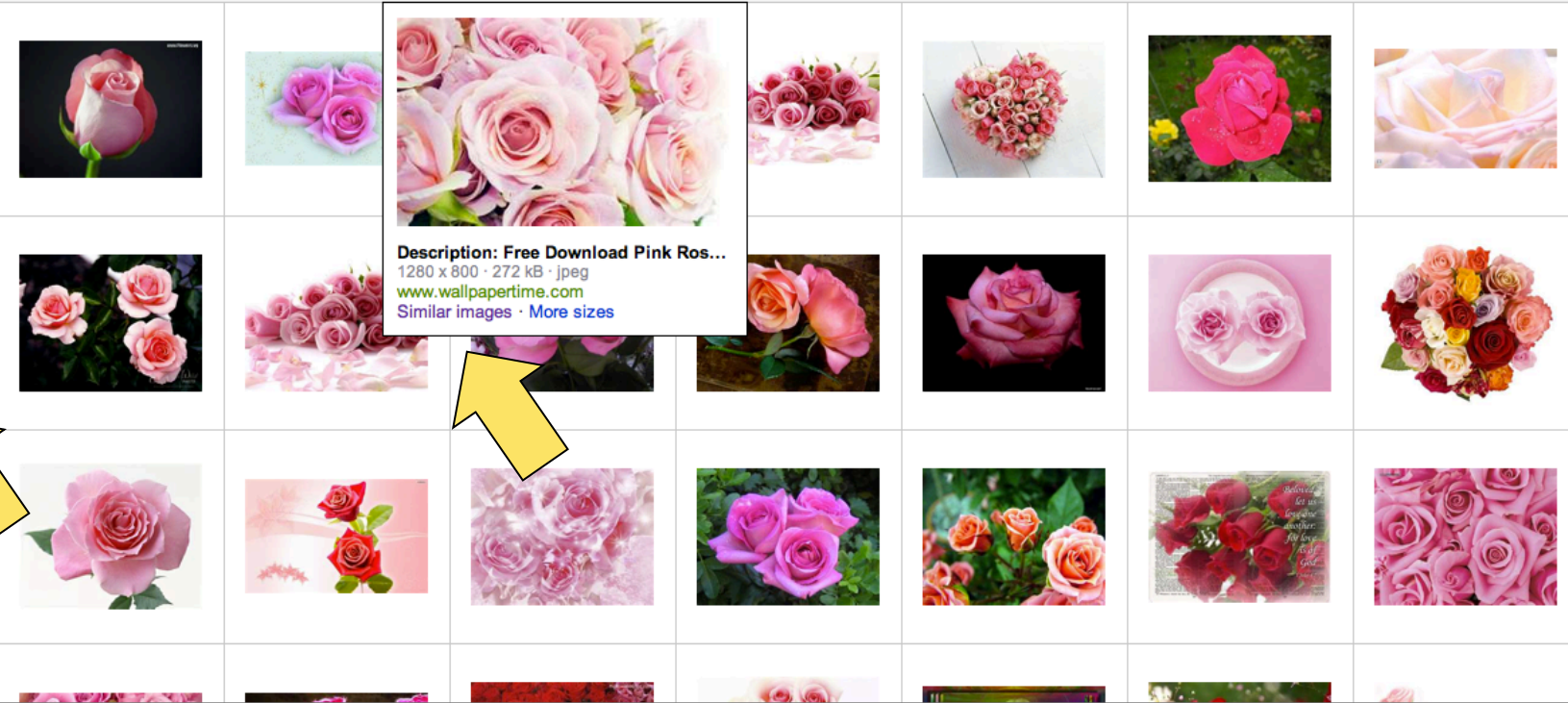


STYLE

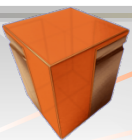
- All
- Photograph
- Clipart

PEOPLE

- All
- Just faces
- Head & shoulders

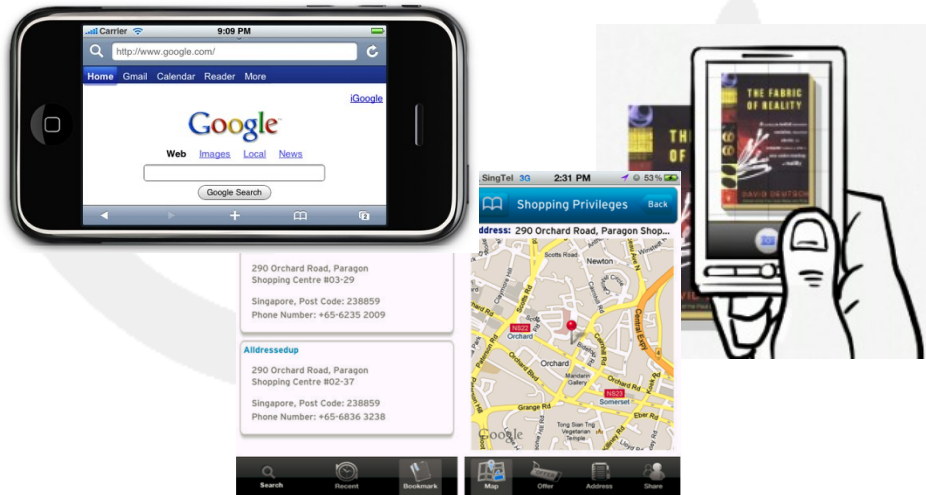


Description: Free Download Pink Ros...
1280 x 800 · 272 kB · jpeg
www.wallpapertime.com
[Similar images](#) · [More sizes](#)



Mobile Search

Mobile search is not just a simple shift of PC web search to mobile equipment

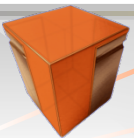


- Limited processing capabilities
- Communication constraints
- Memory restrictions
- Clock and power limitations

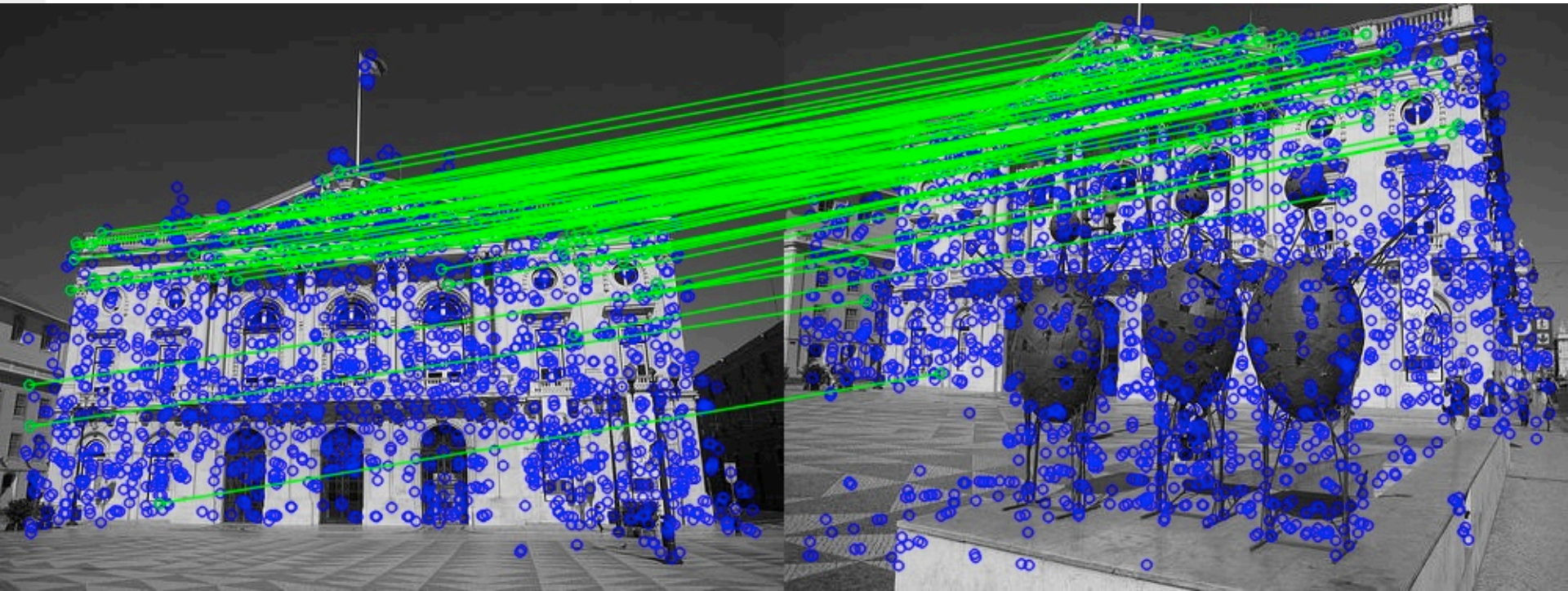
- Visual-based search
- Voice-search
- Location-based services
- Visualization

- 15% of the 1,275 million units were smartphones in 2008
- Mobile image search is moving mainstream and gaining momentum

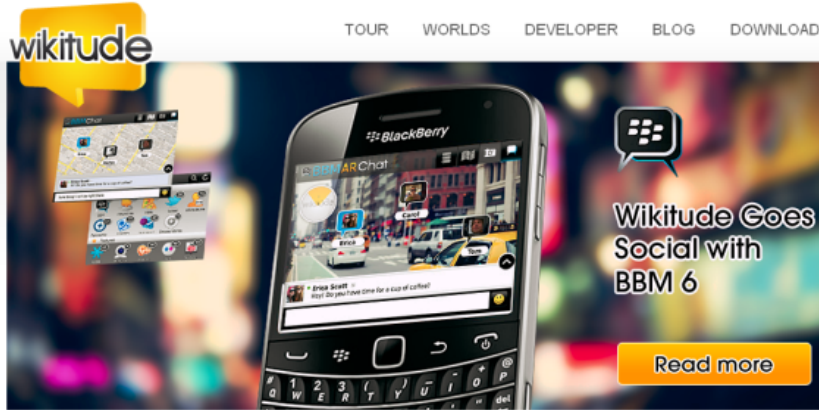
mobileacuity



Key-enabler technology: Similar Image Search



AR: GPS + accelerometer

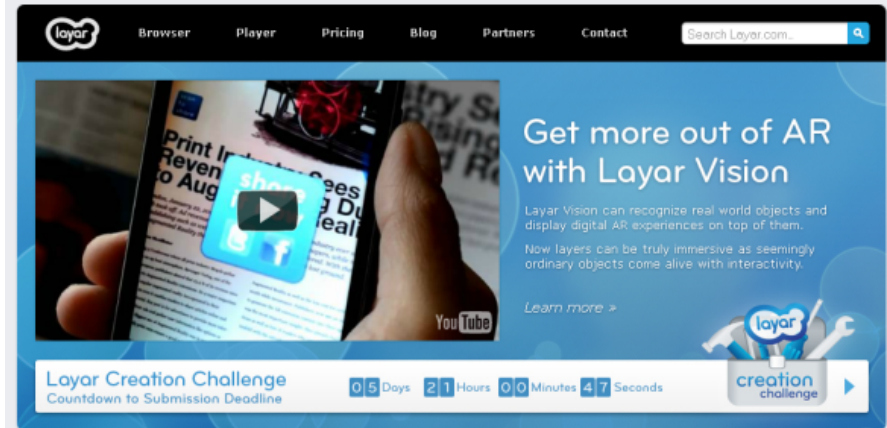


wikitude

TOUR WORLDS DEVELOPER BLOG DOWNLOAD

Wikitude Goes Social with BBM 6

Read more



layar

Browser Player Pricing Blog Partners Contact

Search Layar.com

Get more out of AR with Layar Vision

Layar Vision can recognize real world objects and display digital AR experiences on top of them. Now layers can be truly immersive as seemingly ordinary objects come alive with interactivity.

Learn more >

Layar Creation Challenge

Countdown to Submission Deadline

05 Days 21 Hours 00 Minutes 47 Seconds

creation challenge



juniaio 3.0 takes Augmented Reality another step forward, making the *Augmented World* around us come to life. The new *SCAN* feature now scans everything: images, QR codes and bar codes.

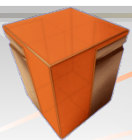
Read the Press Release

Watch the Video!

juniaio 3.0

Compatible with iPhone, iPod Touch, & iPad2. Requires iOS 4.0 or later.

Available in the Apple App Store



Social Search

Social search takes a radical new shape incorporating new dimensions of similarity: Facebook's open graph, Timestamps, geo-location, tag co-occurrence



- Current search engines largely ignore the social context, and solely leverages the textual annotations
- A user searching for multimedia, might very well appreciate relevant photos of a user in her social network
- Ranking is driven by parameters derived from the social network itself and not so much by the content



















source: <http://trusthop.net/blog/tag/social-search/>

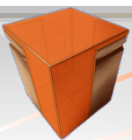
Image based search “within” social networking sites (flickr)

Search | [Photos](#) | [Groups](#) | [People](#)

Everyone's Uploads **SEARCH** [Full Text](#) | [Tags Only](#)
[Advanced Search](#)

Sort: **Relevant** | [Recent](#) | [Interesting](#) View: **Small** | [Medium](#) | [Detail](#) | [Slideshow](#)

 From carlos_seo	 From queropere	 From Bearzito	 From marmalade...	 From Frans Schmit	 From juanma.ca...
 From cabrianno	 From L@Scimmia	 From MattDeanUK	 From Hans Bouman	 From Hans Bouman	 From Hans Bouman
					





flickr®

Deutsches Eck from Ehrenbreitstein
Fortress, Koblenz, Germany



by [schaengel](#)

121 comments 69 faves

Tagged with [koblenz](#), [ehrenbreitstein](#) ...
Taken on [November 15, 2009](#), uploaded
[November 17, 2009](#)

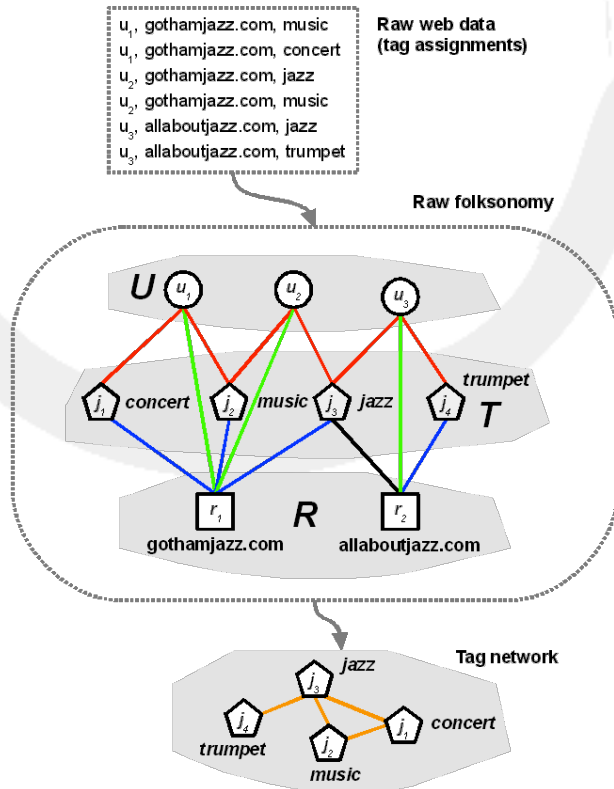
See more of [schaengel](#) photos, or visit
his [profile](#).



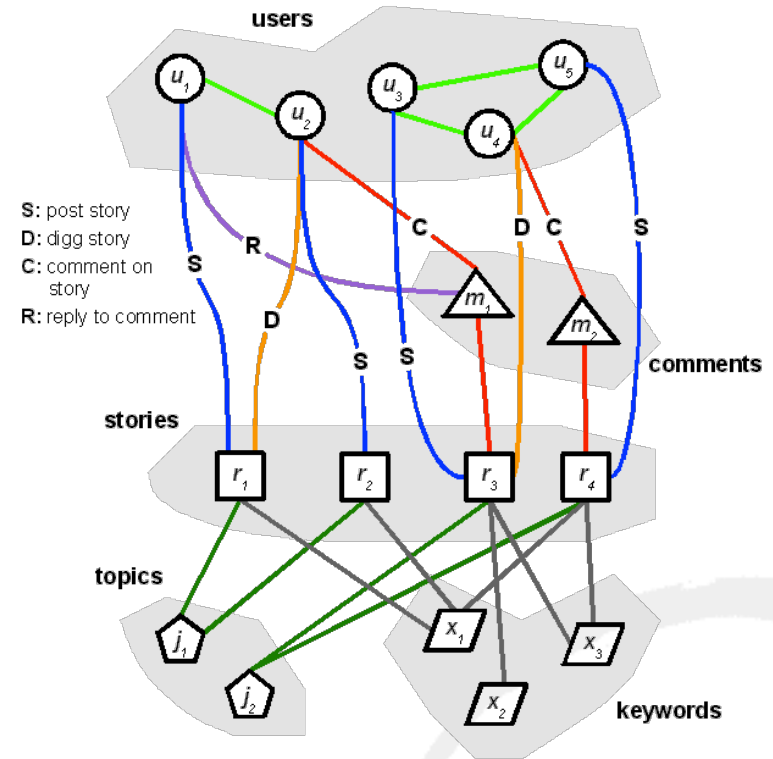
at Ehrenbreitstein Fortress you can get a
ending area.

Examples of Social Media networks

Folksonomy (Delicious)



MetaGraph (Digg)

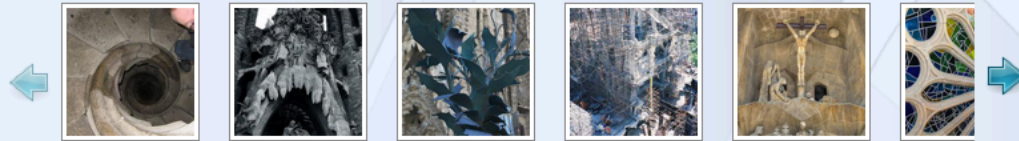
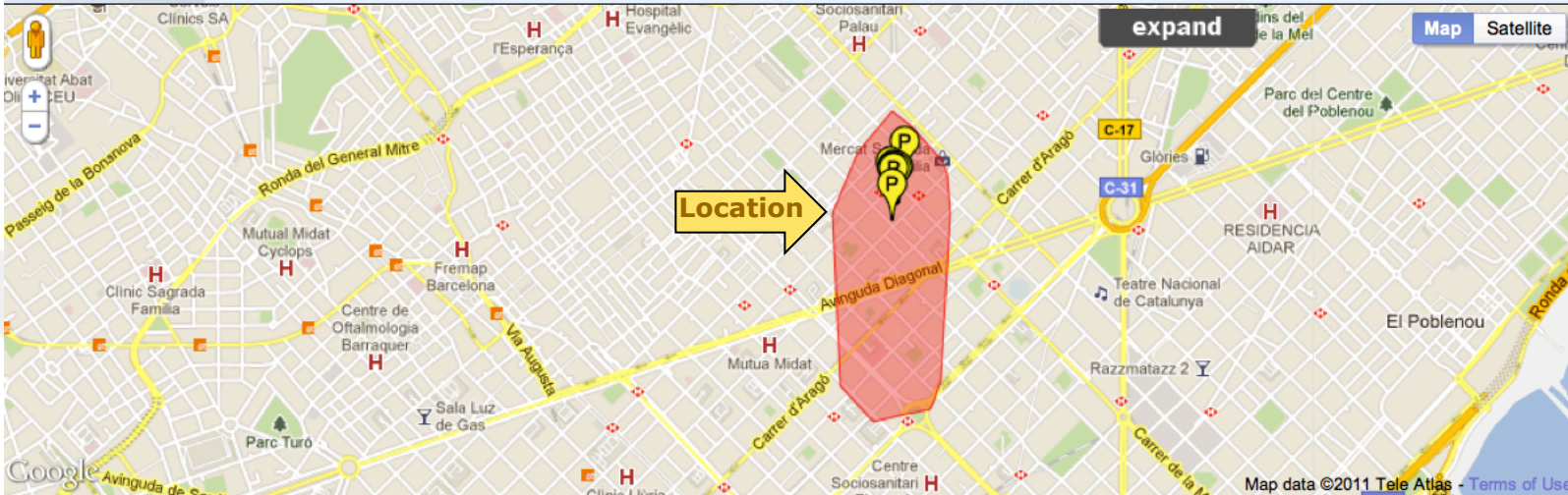


Mika, P. (2005) Ontologies Are Us: A Unified Model of Social Networks and Semantics. Proceedings of the 4th International Semantic Web Conference (ISWC 2005), Springer Berlin / Heidelberg, pp. 522-536

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

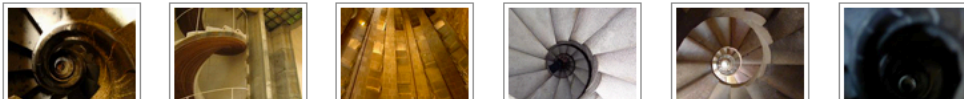
La Sagrada Familia

Barcelona



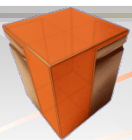
Top Clusters

1. Sagrada Familia



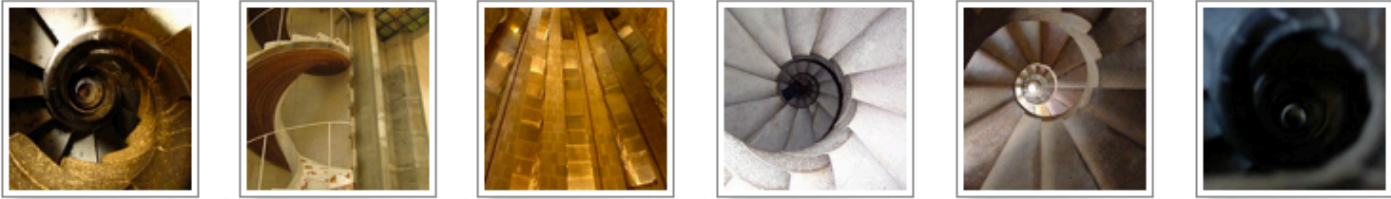
Details

Area Name	La Sagrada Familia
Total Photos	12872
Total Users	1476
Most popular Months	May, Aug
Most popular season	Summer

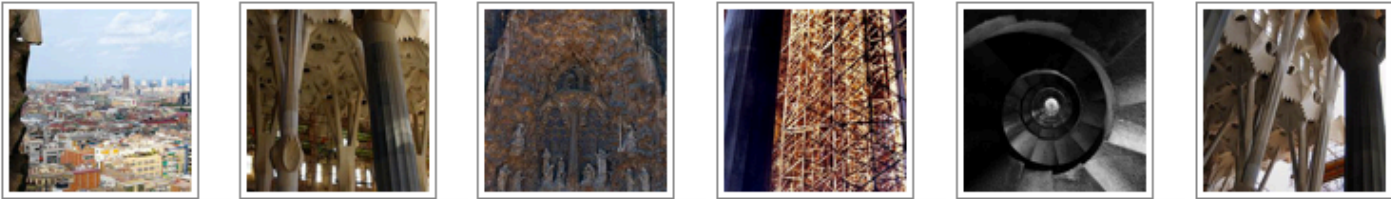


Top Clusters

1. Sagrada Familia



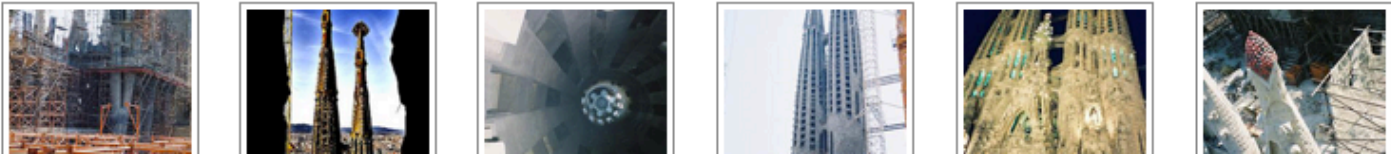
2. Sagrada Familia



3. Sagrada Familia

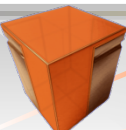


4. Sagrada Familia



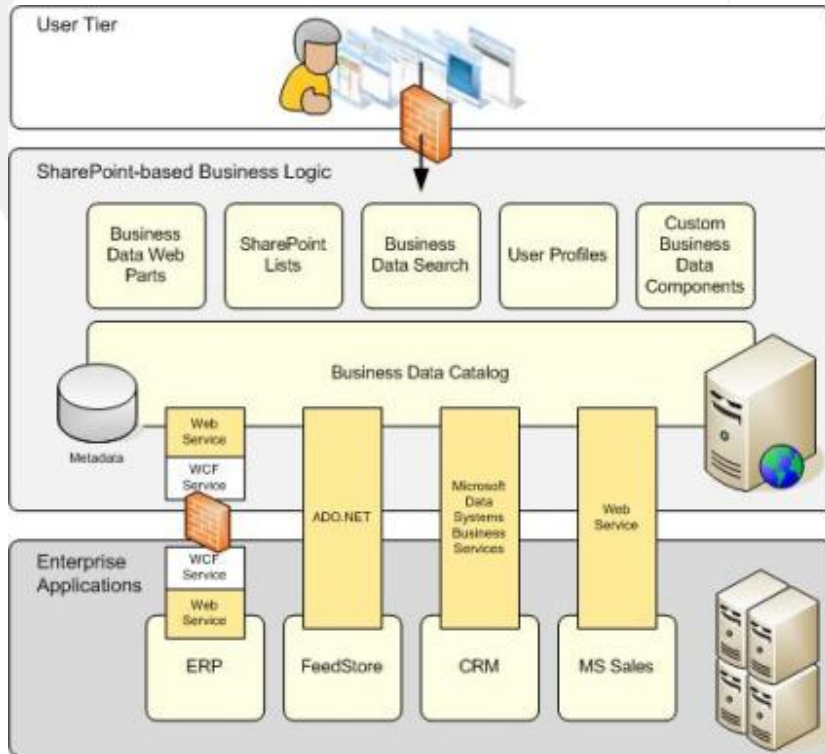
Interior details

Outside views



Enterprise Search Business

Enterprise search is fundamentally different from Web-based search



- Enterprise search engines cannot rely on a popularity based page rank algorithm
- This has led this sub-market to put more emphasis on semantic analysis and faceted results
- Other differentiating factors include security, diversity, accuracy, etc
- All these factors opens up an opportunity for competition against Google where Europe can play a significant role (e.g. Autonomy, Exalead)

source: <http://www.microsoft.com/casestudies/>

Enterprise Search

CHORUS+ Network of Audio-Visual Media Search

AVmediasearch.eu is the hub where you will find all useful information related to events, technologies, resources... linked to audio-visual search.



CHORUS+ ACTIVITIES



■ CHORUS+

CHORUS + ACTIVITIES



"EXPLORING THE FUTURE OF ENTERPRISE SEARCH" - AN EXPERT WORKSHOP & THINK-TANK EVENT

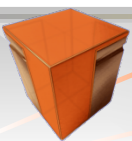
13-14 October 2011, Seville, Spain

which will contribute to gain insights into the techno-economic trends in enterprise search!

The aim of this workshop is to gain insights into the techno-economic trends in enterprise search and to study how they will impact the European economy and society.

The workshop aims to contribute to a better understanding of the following issues:

- **Market Dynamics:** by painting the landscape of enterprise search, including the current and future business models of providers of enterprise search solutions and services (e.g. strengths, entry barriers, differences amongst them, etc.).
- **Future Prospects:** by identifying emerging techno-economic trends, discussing likely developments and the market structure of tomorrow in the domain of enterprise search.
- **SWOT Analysis:** by exploring the strengths, weaknesses, opportunities and threats (SWOT) for the EU with respect to enterprise search. The discussion will focus on the economic drivers and challenges influencing the future of search engines, as well as any other impediments (of regulatory, technical, economic, or social nature) that may hamper successful deployment in Europe.



Music Search Business

The need for retrieval of (very specific) musical content makes more apparent the limitations of traditional text-based search

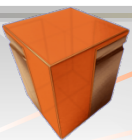


This photo, is copyright (c) 2011 by photosteve101 and made available under a Attribution-Noncommercial-Share Alike 2.0 license

- Private users are primarily interested in matching their own taste
- Professional users, on the other hand, aim at production and recommendation
- Text search is inadequate because:
 - the exact search criterion is not known
 - music needs to be annotated with an enormous manual effort
- The sound-based retrieval and recommendation is gaining popularity

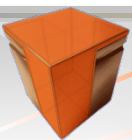
TrackID™

What's that tune?



Where we want to go?

- Understand data and **content**
 - Transform implicit relations, connections, interactions to useful information
 - Provide structure to poorly structured information
- Understand the **user**
 - Context and user-aware search
 - User experience and interfaces
- **Big Data**
 - The size of the data becomes itself the problem
 - Real-time



Where we want to go (1/2)

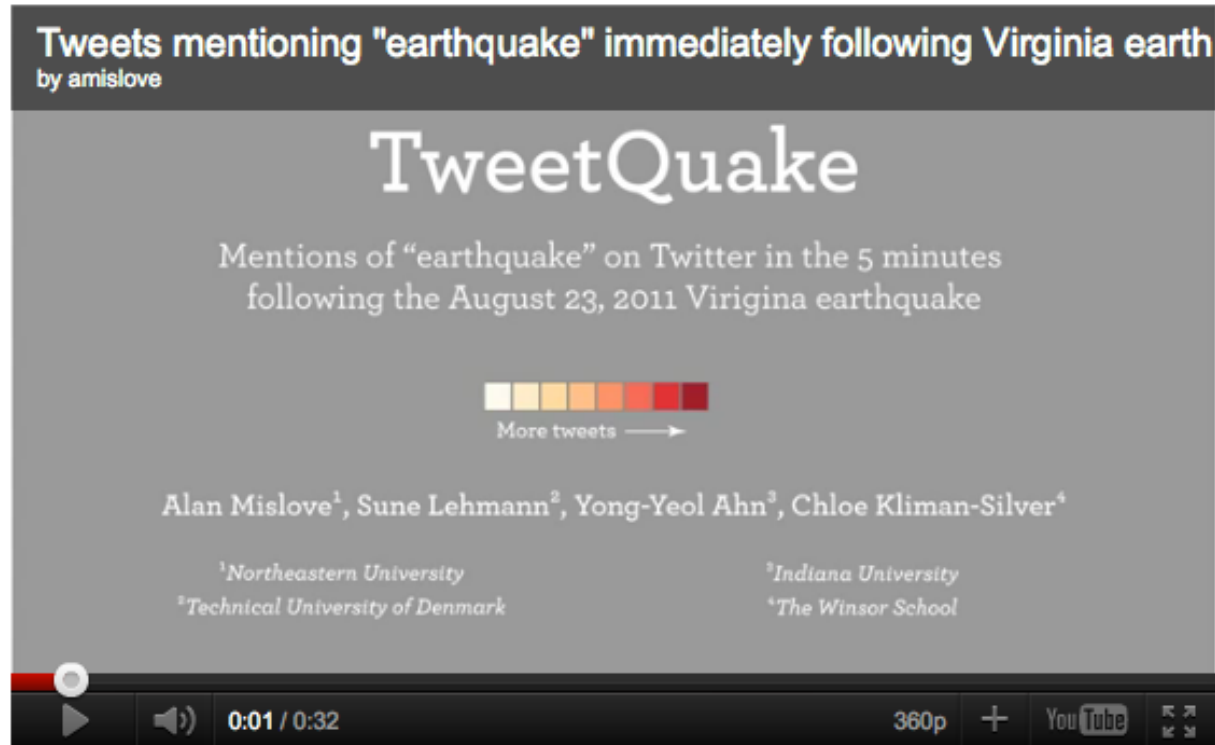
Domain	Challenge
Multimodal content	Exploit the multimodal nature of multimedia content when searching for relevant items
Affective user	Capture and use the feeling of the user along the search/browsing process or in reaction to a search result
Event-based content	Use events as the primary means for organizing and indexing multimedia content
User experience user	Develop interfaces capable of coping with the content explosion and the need for advanced search methods such as visual search interfaces, augmented reality applications
Large scale indexing big data	Enrich over time the “traditional” indexing methods and use sociality to achieve large scale indexing at the level of the Web



Where we want to go (2/2)

Domain	Challenge
Network nodes content	Develop content-aware nodes that will be able to answer questions such as “where” and “how” in addition to the original “what”, that a search engine normally replies
Real-time content big data	Find updates, handle real time indexing of the content, execute real time matching and ranking algorithms, and calculate statistics and trends of the multimedia content
Content diversity content	Develop diversity-aware methods and tools for effective design by harnessing, controlling and using the effects of emergent knowledge properties
Aggregation and mining content big data	Handle the huge scale of sensor data available, its uncontrolled nature, its distribution mechanism, and the potential to combine the analysis results with information from different modalities
Standardisation	Rely on standardisation and global cooperation in order to provide full interoperability of “Search Computing” solutions

Social Media as real-time Sensors



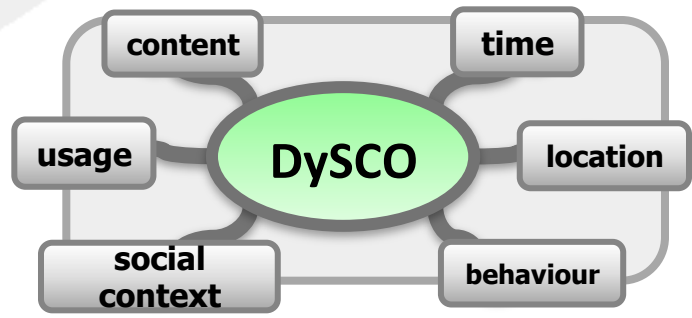
“...if you're more than 100 km away from the epicenter [of an earthquake] you can read about the quake on twitter before it hits you...”

Novel framework for mining and aggregating massive amounts of media from multiple dynamic and quickly evolving social sources.



Massive social media and web

**Social media mining
Aggregation & indexing**

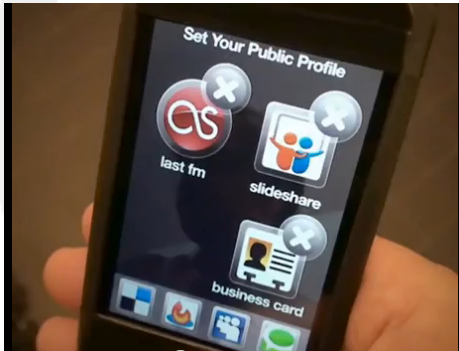


**Personalised access & search
Ad-hoc P2P networks**

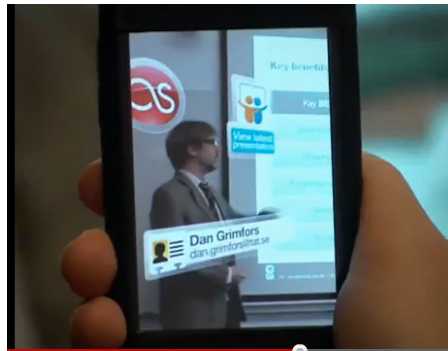


**News -
Infotainment**

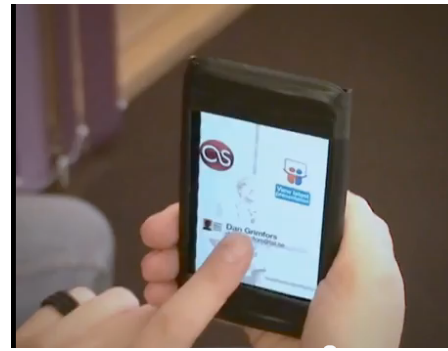
Augmented ID – concept



Select your profile



Allow people to find information about you, through their mobiles



Link them with additional content (Linked in, slideshare)



Business models

- Advertising based
 - Google ads, merchandising, product placement, user profiling
- Packaging search with some other good or service:
 - Packaged with: a) mobile operator, mobile handset, touristic pack, etc.
- Premium service
 - Charge premium functionality, value-added, pay-as-you-go, subscription

Key Challenges

- Include content ownership, copyright and licensing especially in cases where user generated content is involved.
- Handle the fragmentation with respect to regulations and contacts with local businesses and the intellectual property of the technology

Open innovation

- Business ecosystems

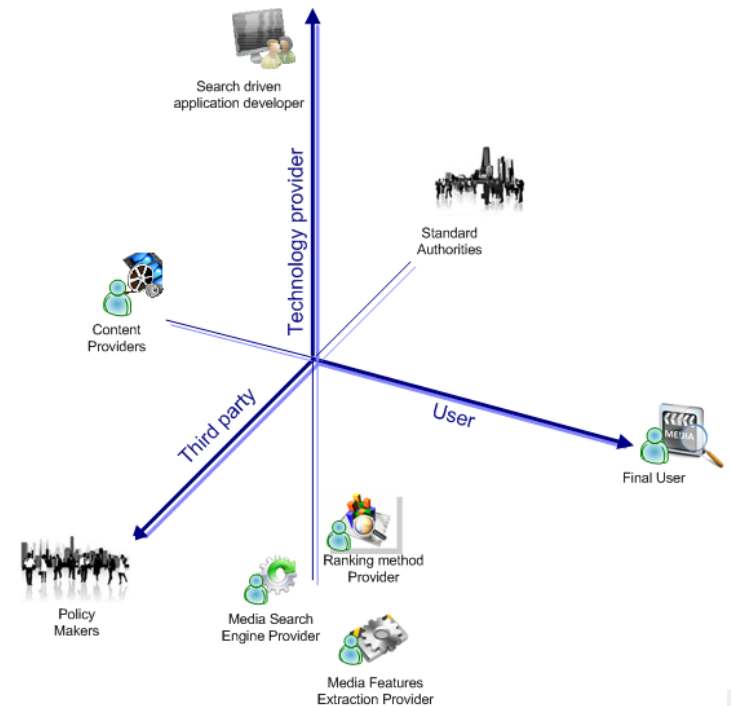
- A sustainable community of enterprises and institutions, sometimes collaborating and sometimes competing, but in both cases creating value for end users, themselves and each other.

- Search Ecosystem at European Level

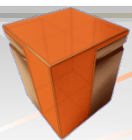
- Adopt a European level strategy for building the business ecosystem

- User generated innovation

- Active users communities/social networks are involved in strategic decisions through moderation tools for large-group dialogues and with tools and services for community and crowd management.



A possible search computing Business Ecosystem

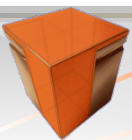


Benchmarking

- Value
 - Streamlines research by eliminating redundancy
 - Enables direct performance comparison between algorithms
 - Increases efficiency by sharing resources between research sites
 - Allows researchers to interact in a productive mixture of competition and collaboration
- Reinforcing Europe's competitiveness
 - MediaEval, PetaMedia, ImageClef, PASCAL, etc.

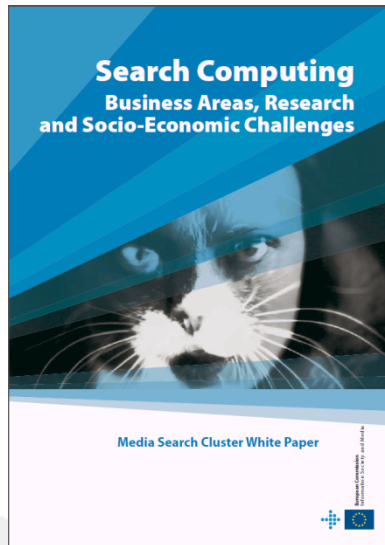
Key Challenges

- Avoid blocking innovation by placing too much importance on pure performance and not on novel technologies.
- Continuously change and modify the tasks, so that plenty of room is available for participants to innovate.



Conclusions

- Search plays an important role in a number of very active sub-markets and business areas ranging from **mobile devices** and **social networks** to **enterprise** and **music** search.
- While a lot of research approaches have been applied to various aspects of “Search Computing”, **the problems are far from being solved and new challenges arise.**
- Key **research and business directions** are identified towards these problems
- The research and socio-economic challenges that fall in the area of “Search computing” should be used as guidelines in defining the **future research agendas and programmes.**



Questions

?

Thank you!

<http://mklab.itι.gr>



CHORUS+
<http://www.ist-chorus.org/>



GLOCAL
www.glocal-project.eu



Cubrik
www.cubrikproject.eu



I-SEARCH
<http://www.isearch-project.eu>



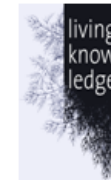
WeKnowIt
<http://www.weknowit.eu/>



SocialSensor
www.socialsensor.eu



COAST
<http://www.coast-fp7.eu/>



Living Knowledge
<http://livingknowledge-project.eu/>



PHAROS
www.pharos-audiovisual-search.eu



PetaMedia
<http://www.petamedia.eu>



ETSI
<http://www.etsi.org>



Networked Media and Search Systems Unit of DG INFSO
http://cordis.europa.eu/fp7/ict/netmedia/home_en.html

