

META=NET

Results of Vision Group Media and Information Services

META=VISION

Stelios Piperidis
ILSP-Athena RC, Greece

META-COUNCIL 2010, Brussels



About the Speaker

- ❑ Stelios Piperidis
- ❑ Senior Researcher/Head of Department at ILSP- Athena RC
- ❑ President of ELRA
- ❑ Founding Member of META-NET
- ❑ Convenor Media and Information Services Vision Group
- ❑ Member of META-COUNCIL

The Vision Group

Media and Information Services

- ❑ **Fields:** audiovisual sector, news services, digital libraries, portals and vortals, search engines, social networks etc.
- ❑ **Stakeholders:** audiovisual and media industries, web and search engine providers, archives, etc.
- ❑ **Technologies:** speech processing, subtitling, text simplification, media mining, multilingual data processing, topic identification, content classification and structuring, information extraction, summarisation, multidimensional analytics, authoring, LT-enabled digitisation, mono/ multilingual/ multimedia search, semantic search, unregulated language and discourse processing , etc.
- ❑ **Organizers**
 - Stelios Piperidis (ILSP, Greece)
 - Margaretha Mazura (EMF, Belgium/UK)
- ❑ **Rapporteur**
 - Philippe Wacker (EMF, Belgium/UK)
- ❑ **Meetings**
 1. Paris, 10 September 2010 , 2. Barcelona, 15 October 2010

The Vision Group

Media and Information Services

- Goals:
 - Focus on LT enabled Media and Information Services
 - Language & content processing technologies as communication & information facilitators
 - Requirements that Media and Information related sectors have from LT and the role LT will play.
- A clear view on LT in Media and Information Services will:
 - highlight their importance in a sector which is becoming more and more important from a financial point of view,
 - trigger financial and strategic support (by the EC and the EP, among others), to give Europe the chance of becoming *the* global player.

External members

Name	Organisation	Sector
Nozha Boujemaa	INRIA, France	Multimedia content search, Digital libraries
Marin Dimitrov	Ontotext, Bulgaria	Knowledge management, Semantic Web, Web Services
Christoph Dosch	IRT, Germany	Broadcasting technologies/ Multimedia content search
René van Erk	Wolters Kluwer, The Netherlands	Global information services
Robert Gaizauskas	University of Sheffield, UK	IE, Dialogue modelling
Gregory Grefenstette	Exalead, France	NLP/IE/IR
Christopher Kermorvant	A2iA, France	Document classification, advanced data extraction
Claude de Loupy	Syllabs, France	Information services, semantic search
Alexandre Passant	DERI, Ireland	Semantic web, social software
Sergi Sagàs	MediaPro, Spain	Broadcasting technologies
Alessandro Tescari	Pervoice, Italy	Multilingual Speech
Jakub Zavrel	Textkernel, The Netherlands	Document understanding, web mining, text matching solutions

Name	Organisation
Toni Badia	BM, Spain
Aljoscha Burchardt	DFKI, Germany
Nicoletta Calzolari	CNR, Italy
Roberto Cencioni	EC, Luxembourg
Gil Francopoulo	LIMSI, France
Marko Grobelnik	JSI, Slovenia
Hanna Klimek	EC, Luxembourg
Maria Koutsombogera	ILSP, Greece
Joseph Mariani	LIMSI, France
Margaretha Mazura	EMF, Belgium (Convenor, Rapporteur)
Stelios Piperidis	ILSP, Greece (Convenor)
Georg Rehm	DFKI, Germany
Hans Uszkoreit	DFKI, Germany
Philippe Wacker	EMF, Belgium (Chair)



Domain specific Needs

Demand and Observations

- ❑ Capture of web content [a collection of everything, text, images, videos etc., like archive.org] at a European level
- ❑ Improvement of speech technologies
- ❑ Advances in Audiovisual Search
- ❑ Deeper understanding of text content
- ❑ Large-scale text modelling
- ❑ Genre and text-type based layers in language analysis and generation.
- ❑ Dialogue and Interaction modelling



Domain specific Visions

I need to know now...

- ❑ European citizens need to accurately know about X : <who, what, where, when, why, what others say about X, ...>
- ❑ Information overload handling and Social stream mining
 - An efficient way of exploiting millions of available knowledge bases, social & real time streams is required.
- ❑ Life logging
 - Gather information on a massive scale and exploit it when someone is looking for something
 - Provide the information generated in the 27 EU MS to their citizens in their own language
 - Technologies able to gather all concepts & associated content and/or knowledge related to every discipline need to be implemented

Improved information navigation and presentation

- ❑ Quality, coverage and robustness of text mining technologies, further exploited in search and in information navigation and presentation
 - Robust, wide coverage language analysis (parsing, etc) in all community languages should be pursued.
- ❑ Genre-aware LT applications
 - Able to react and behave in a way that is best suited to a given communication situation.
 - Research advances needed in formalizing and incorporating text type knowledge in LT applications.
- ❑ Semantic Annotation
 - Extend research from sentences to discourses, from documents to dialogues, from artificial to natural interaction.
 - Include multimodal, multifunctional, interpersonal communication, and cross-modal interaction to facilitate accessibility.

Intelligent media channels speaking our language

- ❑ Federated Audiovisual Search
 - Provide intelligent answers to everyday's questions.
 - Innovative technologies in intelligent ways of recognizing and identifying objects, persons and actions are required.

- ❑ Multimedia multilingual subtitling
 - Foster the access to information for ethnic minorities and facilitate information exchange between EU countries.

- ❑ Deliver more efficient online advertising

- ❑ Personalized task-centered interactive information assistants
 - Know or adapt to: what the user knows or has already asked for, the user's language, education and level of expertise.
 - Richer understanding of information seeking should be a priority.
- ❑ Robust speech-enabled executive assistants
 - Automatic minutes production from meetings, automatic indexing for voice search, speech transcription and translation for videos.
 - Research advances in Speech Recognition, transcription and synthesis should be pursued.
- ❑ Voice Control instead of traditional GUIs
 - Simpler in use, provide trustful services .
 - Need to upgrade old-fashioned, complex-to-use GUIs with new functionalities

In a Nutshell: Topics with Visionary Potential

- ❑ Domain specific
 - Information agents
 - Enabling citizens to know accurately and instantly
 - Improved information navigation and presentation
 - Intelligent language aware media channels and related services
 - Personal information assistants

- ❑ Domain independent (not discussed in this short presentation)
 - Eliminate language barriers for consumers and SMEs.
 - Promote the Language Resources ecosystem.
 - Adopt a common infrastructure to ensure interoperability.
 - Cost-effective porting of LT services and solutions across domains.
 - Enable rapid progress in basic technology for semantic annotation and search.
 - Develop synergies among industry and academia.