MULTISENSOR: Mining and Understanding of multilinguaL contenT for Intelligent Sentiment Enriched coNtext and Social Oriented inteRpretation

Stefanos Vrochidis[1], Ioannis Kompatsiaris[1], Vladimir Alexiev[2], Gerard Casamayor[3], Enric Staromiejski[4], Ioannis Arapakis[5], Reinhard Busch[6], Michael Jugov[7], Nicolaus Heise[8], Teresa Forrellat[9], Anastasia Moumtzidou[1], Maurice Grinberg[2], Leo Wanner[10, 3], Alan Mas Soro[4], Iris Miliaraki[5], Vera Aleksic[6], Mirja Eckhoff[7], Tilman Wagner[8], Sara Pedraz[9]

[1] Centre for Research and Technology Hellas, [2] Ontotext, [3] Pompeu Fabra University, [4] Everis, [5] Barcelona Media-Yahoo! Labs, [6] Linguatec, [7] pressrelations, [8] Deutsche Welle, [9] PIMEC, [10] ICREA

stefanos@iti.gr (Corresponding author)

Abstract. This paper presents an overview of MULTISENSOR project, which deals with semantic content integration for sentiment enriched context and social oriented reasoning-based interpretation.

Keywords: multilingual, semantics, reasoning, multidimensional, sentiment, context, decision support, media monitoring, SME internationalisation

1 MULTISENSOR motivation and goals

During the past decade, the rapid development of digital technologies has led to a great increase in the availability of multilingual and multimedia content worldwide, which is repetitive, complementary, contradictory and even unreliable across political, cultural, or linguistic borders. The consumption of such content regardless of its reliability and cross-validation can have important consequences on the society.

To break this isolation, we need technologies that provide unified access to multilingual and multicultural economic, news story material across borders. To this end MULTISENSOR envisages at an integrated view of heterogeneous resources sensing the world (i.e. sensors), such as international TV, newspapers, radio and social media. The approach of MULTISENSOR will build upon the concept of multidimensional content integration (Fig 1) by considering the following dimensions for mining, correlating, linking, understanding and summarising heterogeneous material: language, multimedia, semantics, context, emotion, as well as time and location. The overall goal of MULTISENSOR is to develop a unified platform with a view to providing services such as media monitoring, and decision support for SME internationalisation.

Towards the development of a unified platform that allows for multidimensional content integration, MULTISENSOR confronts the following scientific objectives:

1. Mining and content distillation of unstructured heterogeneous and distributed multimedia and multilingual data including semantic concept extraction.
2. User- and context-centric analysis of multimedia and multilingual content.
3. Semantic integration and interpretation over the spatiotemporal and psychological dimension of heterogeneous multimedia and multilingual data.
4. Development of hybrid reasoning techniques (forward and backward chaining, multi-threaded, temporal, geo-spatial reasoning) for intelligent decision support.
5. Context-aware multimodal aggregation and multilingual summarisation and adequate presentation of the information to the user.
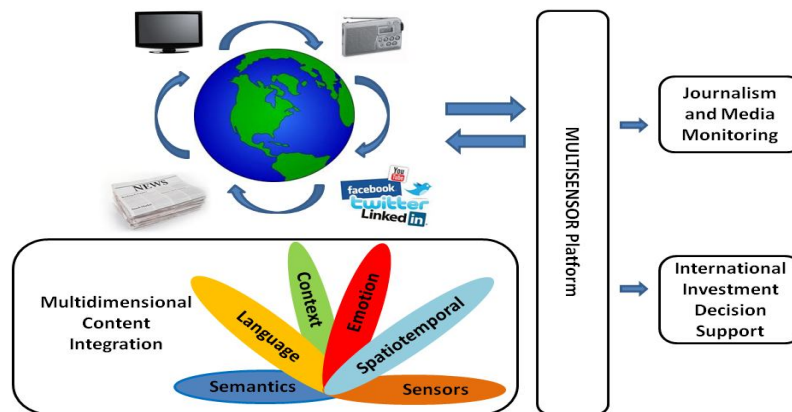

Fig. 1. The MULTISENSOR concept

2   Dissemination and collaboration objectives in ESWC 2014

MULTISENSOR was initiated on November 2013. The current phase of the project includes the definition of the use cases and the elicitation of user and technical requirements that will lead to the development of the platform architecture.

In ESWC 2014, MULTISENSOR aims at presenting the scientific objectives, the use cases and the developments foreseen by the project, in order to identify areas for collaboration with existing European projects focusing on the semantic dimension. Especially, MULTISENSOR is interested in developing synergies and coordinating the developments between concurrent projects in the areas of semantic content integration, reasoning for decision support, semantic concept extraction and semantic interpretation of heterogeneous multilingual and multimodal content. These synergies will include knowledge, tools, datasets and technology sharing, as well as identification of activities and goals, which can form the basis for common research proposals, researcher exchange and joint organization of events and participation in standards.