The MULTISENSOR project – Development of Multimedia Content Integration Technologies for Journalism, Media Monitoring and International Exporting Decision Support

Dimitris Liparas[1], Stefanos Vrochidis[1], Ioannis Kompatsiaris[1], Gerard Casamayor[2], Leo Wanner[2], Ioannis Arapakis[3], David García Soriano[3], Reinhard Busch[4], Boris Vaisman[4], Boyan Simeonov[5], Vladimir Alexiev[5], Andrey Belous[6], Emmanuel Jamin[6], Nicolaus Heise[7], Tilman Wagner[7], Michael Jugov[8], Mirja Eckhoff[8], Teresa Forrellat[9] and Martí Puigbó[9]

[1]Centre for Research and Technology Hellas, [2]Pompeu Fabra University, [3]Eurecat, [4]Linguatec, [5]Ontotext, [6]everis, [7]Deutsche Welle, [8]pressrelations, [9]PIMEC

Abstract. The rapid development of digital technologies has led to a great increase in the availability of multimedia content. The consumption of such large amounts of content regardless of its reliability and cross-validation can have important consequences on the society and especially on journalism, media monitoring and international investments. In this context, MULTISENSOR has researched and developed tools that provide unified access to multilingual and multicultural economic, news story material across borders, that ensure its context-aware, spatiotemporal, sentiment-oriented and semantic interpretation, and that correlate and summarise the content into a coherent whole. The goal of the MULTISENSOR project is to provide a platform, which allows for an integrated view of heterogeneous resources sensing the world (i.e. sensors), such as international TV, newspapers, radio and social media. Three demonstrators have been developed, indicating the potential of the platform and providing end-user services such as journalism, commercial media monitoring and decision support for SME (Small and Medium Enterprises) internationalisation.

1 Introduction

Nowadays, the extensive availability of multilingual and multimedia content worldwide is a result of the advances in digital technologies during the past decade, as well as the low cost of recording media. In the best case, this content is repetitive or complementary across political, cultural, or linguistic borders. However, the reality shows that it is also often contradictive and in some cases unreliable, something that can greatly impact its consumption. An indicative example is the current crisis of the financial markets in Europe, which has created an extremely unstable ground for economic transactions and caused insecurity in the population. The consequence is an extreme uncertainty and nervousness of politics and economy on the one side, which makes national and international investments really risky, and on the other side, the inability of journalism and media monitoring to equally consider all the media resources leaves the population in each of these encapsulated areas in its own perspective–without the realistic opportunity to understand the perspective developed in the other areas in order to adjust the own.

To break this isolation, there is a need for technologies capable to capture, interpret and relate economic information and news from various subjective views as disseminated via TV, radio, newspapers, blogs and social media. On top of this, semantic integration of heterogeneous media, including computer-mediated interaction, is required to gain a usable understanding based on social intelligence, while a correlation with relevant incidents with different spatiotemporal characteristics would allow for extracting hidden meaning.

In the MULTISENSOR (**M**ining and **U**nderstanding of multilin-gua**L** conten**T** for **I**ntelligent **S**entiment **E**nriched coo**N**text and **S**ocial **O**riented inte**R**pretation) project[1], we have developed a unified platform for enabling the multidimensional content integration from heterogeneous sensors, with a view to providing end-user services such as journalism, commercial media monitoring and decision support for SME (Small and Medium Enterprises) internationalisation. More specifically, potential investors can benefit from integration and context-aware interpretation of complementary and contradicting multilingual and multimedia information and get decision support for international investments. Media companies and archives can also benefit from the spatiotemporal integration and sentiment-oriented interpretation of heterogeneous content both for media monitoring and for journalism purposes. Finally, the European public can benefit from this integration and context-aware interpretation in the sense that it learns and comes to understand the views, fears and worries of the citizens all over Europe and get support for forming an objective opinion with respect to the state of affairs.

The approach of MULTISENSOR builds upon the multidimensional content integration concept (Figure 1) by considering the following dimensions for mining, linking, understanding and summarising heterogeneous material: language, multimedia, semantics, context, emotion, time and location.
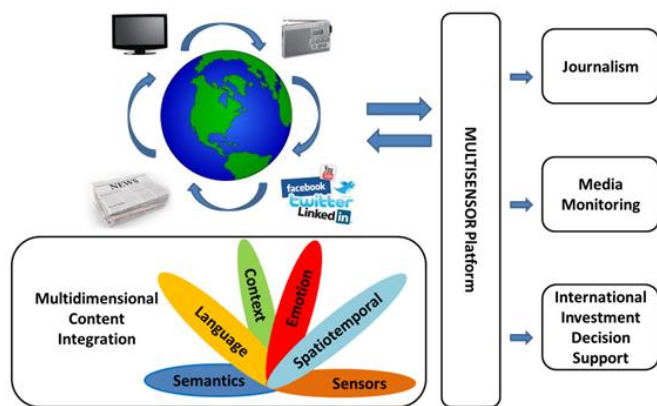

Figure 1. The MULTISENSOR concept

2 Project objectives

In the context of MULTISENSOR, the following scientific objectives with respect to the individual research areas of the project are addressed:

- **Mining and content distillation of unstructured heterogeneous and distributed multimedia and multilingual data**: In this objective, MULTISENSOR attempts to facilitate the data mining from several international resources, including news articles, audiovisual content (TV, radio), blogs and social media and provide intelligent mechanisms for the distillation of information. This objective includes low- and high-level content analysis.

- **User- and context-centric analysis of heterogeneous multimedia and multilingual content**: Here, the focus is on analysing content from the user perspective to extract sentiment and context, analysing computer-mediated interaction in the web and specifically in social media, as well as generating high-level information based on the outcome of the previously mentioned objective. The aim is to develop and integrate into the MULTISENSOR platform research techniques for context extraction, sentiment extraction and social media mining (influential user detection and community detection).

- **Semantic integration and context-aware interpretation over the spatiotemporal and psychological dimension of heterogeneous and spatiotemporally distributed multimedia and multilingual data**: This includes multidimensional content correlation and alignment based on reasoning techniques, as well as on multimodal vector-based representation and topic-based modelling. The multimodal integration is performed on top of the low- and high-level content extracted in the two aforementioned objectives.

- **Semantic reasoning and intelligent decision support services**: The purpose here is to make sense of very large amounts of heterogeneous data by providing diverse analytics, contextualised decision-making support for different situations to enable view of the information from multiple perspectives. In this context, MULTISENSOR has researched and developed advanced reasoning techniques that abide to requirements for scalability and usability.

- **Context-aware multimodal aggregation, multilingual summarisation and adequate presentation of the information to the user**: This objective also includes context-aware interpretation of news by examining their impact on the news consumers in the light of cultural aspects, user experience and engagement, as well as impact on its condensed presentation along with the content summary.

3 User perspective

Within MULTISENSOR, three pilot use cases (UC) were defined and specific requirements were extracted for each one of them:

UC1: **Journalism**: Journalists need to master large heterogeneous amounts of multimedia and multilingual data when writing a new article. On the basis of a market analysis that was conducted and from a journalistic point of view, MULTISENSOR should be able to provide an automatic summarisation of heterogeneous and multilingual digital information. The platform should also automatically suggest related content and information that allows journalists to enrich their coverage of a specific topic.

UC2: **Commercial media monitoring**: Professional clients of media monitoring portals require direct access to comprehensive and targeted business and consumer information. This could include information on consumption habits, competitors and opinions. From a media monitoring point of view, it is important that the MULTISENSOR system follows the usual workflow for the creation of a media analysis. In a first step, the user needs to define the sources and time frame that is to be monitored, along with the search terms he wants to use. In a second step, the search results need to be curated and validated. The MULTISENSOR system should present the results of these queries in different output formats and visualisations.

UC3: **SME (Small and Medium Enterprises) internationalisation**: This UC deals with SME internationalisation, which refers to small or medium-sized companies that want to start or are in the process of expanding from a regional or a national market to a new and foreign market in order to increase turnover and profit. This process is of particular importance, as it is often the only option to achieve growth. But it is also aligned with considerable challenges, such as a lack in knowledge about market conditions or the spoken language in the targeted countries. From the aforementioned, in order for the MULTISENSOR platform to be fully helpful in SME internationalisation cases and improve the decision-making process, it should provide information about several related indicators, regarding the condition of the market, the political and financial situation of the countries, potential competitors, consumption habits, etc. Furthermore, two very important requirements from this UC are summarisation (to reduce the amount of information that the internationalisation expert will need to read and study) and automatic language detection and translation.

4 MULTISENSOR framework

The architecture of the MULTISENSOR framework is depicted in Figure 2. In this architecture, a periodic process of content harvesting takes place, which retrieves source material by crawling a set of sources for news, multimedia and social network content. Next, the different components of the framework, as well as the functionality of the modules that they contain and provide are described.

4.1 Multimedia content extraction

This component aims at extracting knowledge from multimedia input data and presenting the extracted knowledge in a way that subsequent components can operate on it. It includes the following technologies:
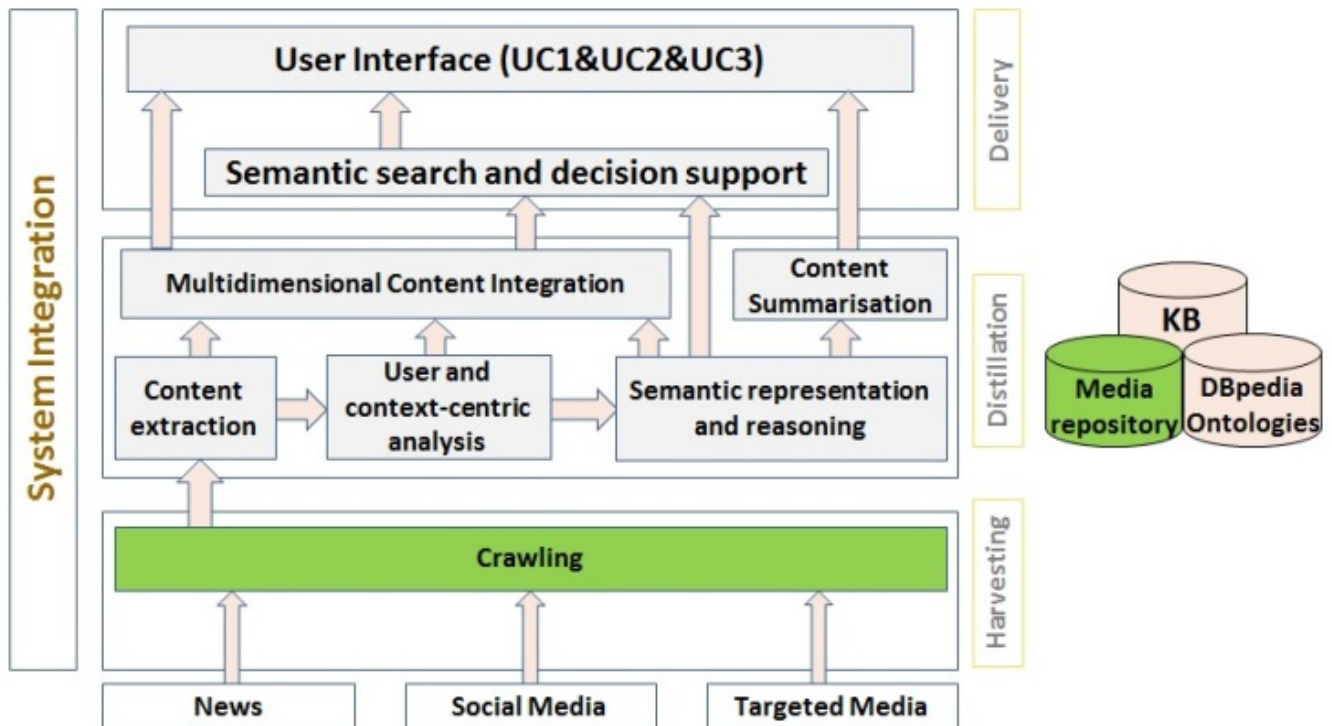
Figure 2. Architecture of the MULTISENSOR framework

1. **Language Identification**: Before a text is stored in the repository, it is analysed in which language it is written and the text is annotated accordingly. The languages considered in MULTISENSOR are English, German, Spanish, Bulgarian and French.

2. **Named entities extraction**: This module aims at identifying names (named entities) in texts. Names are words which uniquely identify objects, like 'Berlin', 'Siemens', etc. The module incorporates two linguistic components that allow all analysis modules to operate on the same input: sentence segmentation and tokenisation.

3. **Concept extraction from text**: Concept extraction starts from the results of the named entities extraction task. The goal of this module is to identify in the text mentions to concepts that belong to the project domains. Candidate concepts are identified through analysis of multilingual corpora. When processing new documents, the module attempts disambiguation of mentions of concepts against relevant ontologies and datasets.

4. **Concept linking and relations**: This module aims at identifying in texts relations between mentions of named entities and concepts. Two relation types are considered: i) coreference relations i.e. several mentions make reference to the same entity, and ii) n-ary relations describing situations and events involving multiple entities and concepts. To this end, a deep dependency parser [1] that delivers deep-syntactic dependency structures from sentences in nature language has been developed. This parser uses the output of an optimised dependency parser [2] as input.

5. **Audio recognition and analysis**: Automatic speech recognition (ASR) is employed in order to provide a channel for analysis of spoken language in audio and video files. The transcripts produced follow the same analysis procedure as the input from other text sources. The languages covered by the ASR component are English and German.

6. **Multimedia concept and event detection**: This module receives as input a multimedia file (i.e. image or video) and computes degrees of confidence for a predefined set of visual concepts. The module performs video decoding (applicable for video files only), feature extraction and classification in order to assign a confidence value for a concept or event existence in an image or video shot [3].

7. **Machine translation**: Automatic machine translation (MT) has two main goals: to provide the translation of the summarisation results in the end of the content analysis and summarisation chain and to enable full-text translation on-demand during the development of language-dependent analysis tools in the project, in case a subset of required languages is not supported by these tools.

4.2 User- and context-centric analysis

The objectives of this component are to model and represent contextual, sentiment and online social interaction features, as well as deploy linguistic processing at different levels of accuracy and completeness.

1. **Extraction of contextual features**: This module provides a set of contextual indicators characterising the content items and a framework for measuring their impact in the context of the use cases. Moreover, it provides representation techniques to be used in effective context-based search.

2. **Polarity and sentiment extraction**: The polarity and sentiment extraction module aims at modelling a robust opinion mining system that is based on linguistic analysis and is applicable to large datasets. Moreover, models that take into account the presence of

named entities in different sentences have been designed within the module.

3. **Social interaction analysis**: The social interaction analysis task involves a set of processes related to analysis of social network data stored into the MULTISENSOR repositories, namely crawled Twitter data. Two modules have been developed in the context of this task, namely the influential user detection and community detection modules. First, a topic-dependent network of contributors based on the mentions in the set of monitored tweets is built and next, retweet probabilities between users in this network are computed. The goal of the influential user detection module is to provide a ranked list of users by decreasing order of influence based on the aforementioned network of mentions and retweet probabilities. The goal of the community detection module is to make use of crawled Twitter posts in order to detect online dynamic communities by means of an appropriate community detection algorithm, which is applied to each graph snapshot defined by the user network of mentions.

4.3 Multidimensional content integration and retrieval

The objective of this component is to achieve integration and retrieval of content along different dimensions.

1. **Multimodal indexing and retrieval**: In this module, a multimedia data representation framework that allows for the efficient storage and retrieval of socially connected multimedia objects is developed. The representation model is called SIMMO (Socially Interconnected MultiMedia-enriched Objects) [4] and has the ability to fully capture all the content information of interconnected multimedia objects, while at the same time avoiding the complexity of previously proposed models.
2. **Topic-based modelling**: In this module, two subtasks are considered: a) category-based classification and b) topic-event detection. The module receives as input multimodal features that are created in the multimedia content extraction component and provides as output the degree of confidence of a number of categories for a specific content item (for category-based classification) [5] or a grouping for a list of content items based on the existence or not of a number of topics / events (for topic-event detection) [6].

4.4 Semantic representation and reasoning

MULTISENSOR includes a semantic layer in order to represent in a unified way heterogeneous content. The following technologies are involved:

1. **Semantic representation**: This representation includes a number of ontologies that are integrated in a common framework, such as DBpedia, GeoNames and FreeBase.
2. **Ontology alignment**: The ontology alignment module discovers candidate semantic correspondences between heterogeneous information descriptions and terminologies and verifies the correctness and consistency of the discovered mappings in an automatic way.
3. **Content alignment**: This module deals with the semantic processing of the multimodal content, in order to identify near duplicate and contradictory information relying on semantic technologies.
4. **Hybrid reasoning and decision support**: In the hybrid reasoning and decision support module, four reasoning techniques have

been developed: hybrid reasoning (consists of a combination of forward and backward chaining), multi-threaded reasoning (parallel inference calculation), temporal reasoning (inference based on temporal entities and sequence in time) and geo-spatial reasoning (ability to reason based on latitude, longitude and altitude of a given location). Additionally, a reasoning-based recommendation system with two main functionalities has been developed: firstly, it determines relevant facts by navigating the graph and secondly, it advises the user by interpreting these facts through the use of the aforementioned hybrid reasoning techniques and the assignment of relevance weights for each selected fact [7].

4.5 Content summarisation

The content summarisation component implements procedures for producing multilingual briefings. Two established strategies in the field of text summarisation are considered in MULTISENSOR:

1. **Extractive summarisation**: Text-to-text summarisation, where the relevance of sentences in the original documents is assessed based on shallow linguistic features in order to decide on its inclusion of a summary. A module following this strategy is used in order to establish a basic infrastructure for summarisation services and implement a fall-back method.
2. **Abstractive summarisation**: Documents are analysed and the information extracted from them is used to generate a summary that is not composed of fragments of the original documents, but is generated directly from data. A module implementing abstractive methods operates on the semantic layer in order to select contents extracted from multimedia documents and also coming from other datasets integrated into the MULTISENSOR system. Contents are selected and organised into a text plan that guarantees the coherent presentation of information. A multilingual linguistic generation system renders text plans into the final summaries.

5 Use cases applications

During the three years of the project's lifetime, three applications have been developed based on MULTISENSOR technologies, with each application addressing one of the three use cases considered in MULTISENSOR. The first one provides search and exploratory functionalities for journalists[2], the second one aims at supporting a media monitoring company to monitor specific profiles for their clients[3], while the third one provides decision support for SME internationalisation[4].

5.1 Journalism use case application

The journalism use case demonstrator is an application that assists media professionals (e.g. journalists, media experts) in finding relevant information in different formats, coming from different sources, and according to the social activities that were produced around.

Figure3 shows the results section of the application, which displays the results of a search query that the user can make, based on a selection of keywords and filtering criteria. On the left side, search-related entities are displayed. By clicking on an entity it will be added to the search query. Then these entities can be used to extend the search query. On the right side, the following information per article is displayed:

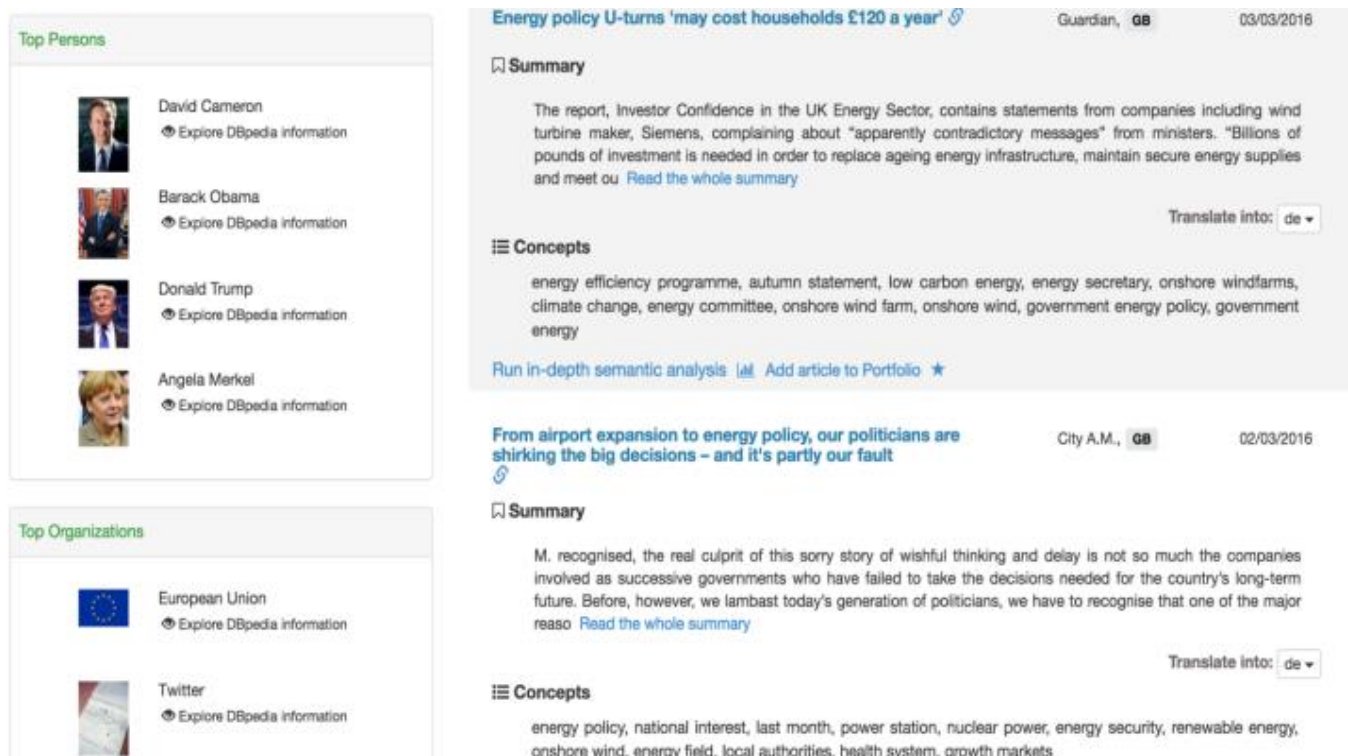
Figure 3. Journalism use case application – Results section

- **Context**: Contextual features per article (title, source, etc.).
- **Summarisation**: Display of the output of the summarisation module.
- **Translation**: The online machine translation service operates on this functionality in order to translate a summary to one of the available languages of the MULTISENSOR project.
- **In-depth semantic analysis**: Displays semantic page view. On this page, more information extracted from the text is displayed (list of named entities, sentiment polarity, cloud of specific concepts and related articles).
- **Article to portfolio**: The link to add an article to the "portfolio" (a folder that contains the user's favourite documents), for further analysis. This analysis generates the aggregated analytical view of the portfolio content.

5.2 Media monitoring use case application

The media monitoring use case application replicates the workflow of a media monitoring professional to execute an analysis for a client. This includes checking articles for relevance by various indicators and saving the relevant articles for a client's profile. The relevant articles can then be analysed, so that conclusions can be drawn from this analysis.

Figure4 depicts the **search** section of the application, where the user is presented with a view to search for articles, based on keywords and language/country filters. Alternatively, the user can select a **profile**, which has settings stored for recurring searches in order to quickly populate the search mask.

In order to evaluate whether an article is relevant for the client, the user can use additional functionalities, such as calling the summarisation and/or translation service. In addition, he can take a look at the entities extracted from the text and read the article's full text. There is also the **analysis** section, where visual results in the form of bar

charts are shown for all articles that have been marked as relevant in the search section. Finally, in the **influencer** section, the information that is extracted from the social interaction analysis modules of MULTISENSOR (influential user detection and community detection) is displayed.

5.3 SME internationalisation use case application

The SME internationalisation use case application supports SMEs in order to start a process of internationalisation with any kind of product. Relevant information related to the countries, the economic situation of the market, the legal information, and the exportation/importation conditions can be retrieved easily to support decision making.

The application supports a number of sectors and products. When a user selects a specific sector, articles about that sector are shown. After the selection of a product, the search will contain specific information about it. Another important functionality of this application is browsing specific information to a certain country for internationalisation support purposes, based on a number of indicators. The considered indicators have been selected and organised by categories to depict the relevant information related to a target country: Politics, Economy, Society and Culture.

Furthermore, the SME professionals are interested in targeting specific countries to establish new commercial activities. For this, the application offers the comparison of several indicators between two targeted countries through the decision support system of MULTISENSOR, which is depicted in Figure5.

6 Conclusions

In this paper, an overview of the successful MULTISENSOR project is provided. The project has developed a platform that supports a)

Figure 4. Media monitoring use case application – Search section

Figure 5. SME internationalisation use case application – Decision support interface

journalists in mastering heterogeneous content in order to prepare articles and identify topics, as well as have access to multilingual summaries; b) commercial media monitoring companies in summarising the opinions of people for specific products and c) SMEs that want to internationalise by providing market analysis, product reports and decision support services. This platform integrates and makes use of innovative modules, which could be separately exploited.

MULTISENSOR technologies will have a big impact from several perspectives. First, they will actively support the journalists (professional and amateurs), commercial media monitoring companies and the international investments by SMEs. Second, the SMEs in the ICT domain will benefit from the open source tools and technologies developed in MULTISENSOR, in order to improve their existing products and offer new services to their clients. Third, the development of such tools will boost the competitiveness specifically in the media monitoring domain and in Europe, since the mobility of SMEs will be facilitated. Finally, the social impact of MULTISENSOR refers to the production of cross-validated news articles and the presentation of news stories from different cultural, political and linguistic perspectives.

7 Acknowledgements

REFERENCES

[1] M. Ballesteros, B. Bohnet, S. Mille and L. Wanner, "Deep-Syntactic Parsing," COLING 2014, Dublin, Ireland, 2014.

[2] M. Ballesteros and B. Bohnet, "Automatic Feature Selection for Agenda-Based Dependency Parsing," COLING 2014, Dublin, Ireland, 2014.

[3] N. Gkalelis, F. Markatopoulou, A. Moumtzidou, D. Galanopoulos, K. Avgerinakis, N. Pittaras, S. Vrochidis, V. Mezaris, I. Kompatsiaris and I. Patras, "ITI-CERTH participation to TRECVID 2014," Proc. TRECVID 2014 Workshop, Orlando, FL, USA, November 2014.

[4] T. Tsikrika, K. Andreadou, A. Moumtzidou, E. Schinas, S. Papadopoulos, S. Vrochidis and I. Kompatsiaris, "A Unified Model for Socially Interconnected Multimedia-Enriched Objects," MultiMedia Modelling, pp. 372-384, Springer International Publishing, January 2015.

[5] D. Liparas, Y. Hacohen-Kerner, A. Moumtzidou, S. Vrochidis and I. Kompatsiaris, "News articles classification using Random Forests and weighted multimodal features", 3rd Open Interdisciplinary MU-MIA Conference and 7th Information Retrieval Facility Conference (IRFC2014), Copenhagen, Denmark, November 10-12, 2014.

[6] I. Gialampoukidis, S. Vrochidis and I. Kompatsiaris, "A hybrid framework for news clustering based on the DBSCAN-Martingale and LDA", 12th International Conference on Machine Learning and Data Mining, New York, 16-21 July 2016.

[7] B. Simeonov, V. Alexiev, D. Liparas, M. Puigbo, S. Vrochidis, E. Jamin and I. Kompatsiaris, "Semantic integration of web data for international investment decision support", In International Conference on Internet Science, pp. 205-217, Springer International Publishing, September 2016.