## 3.1 Publishable summary

### *3.1.1 Project synopsis*

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. This content can be repetitive or complementary across political, cultural, or linguistic borders but can be also contradictive and in some cases unreliable. 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.

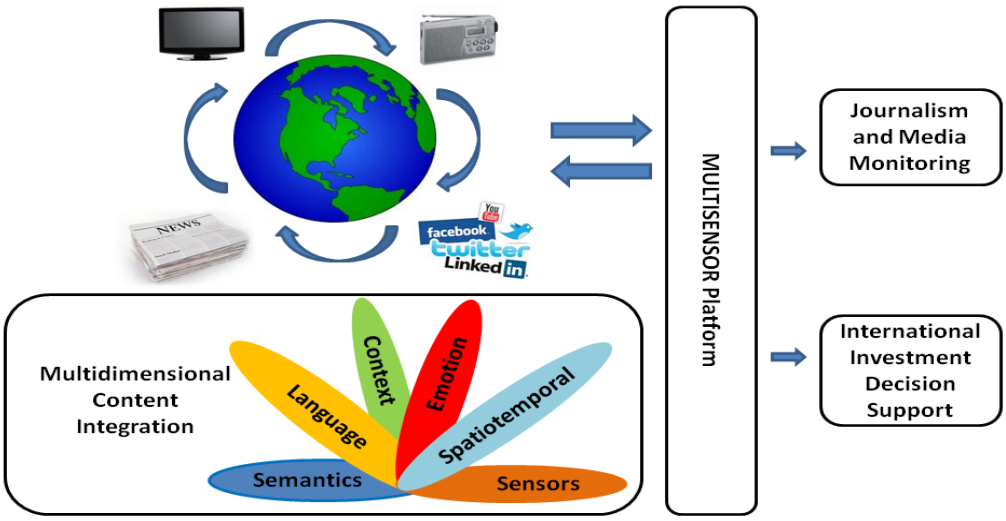


**Figure 0. The MULTISENSOR concept**

In this context, MULTISENSOR aims at providing 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. MULTISENSOR envisages an integrated view of heterogeneous resources sensing the world (i.e. sensors) such as web articles, newspapers, international TV and social media. The approach of MULTISENSOR will build upon the concept of **multidimensional content integration** (Figure 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 research and develop a **unified platform,** which will allow for 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 internationalisation.

To realise this goal MULTISENSOR addresses a number of scientific (SOs), technological (TOs) and business objectives (BOs):

**SO1.** To perform **mining and content distillation of unstructured heterogeneous and distributed multimedia and multilingual data**.
**SO2.** To perform a user- and context-centric analysis of **heterogeneous multimedia and multilingual content**.
**SO3.** To enable the **semantic integration and context-aware interpretation over the spatiotemporal and psychological dimension of heterogeneous and spatiotemporally distributed multimedia and multilingual data** as audio, video, text and social content interaction.
**SO4.** To provide **semantic reasoning and intelligent decision support** services.

**SO5.** To facilitate context-aware **multimodal aggregation and multilingual summarisation and adequate presentation of the information to the user**.

**TO1.** To develop **crawlers** and **data delivery channels** necessary for the collection of data.

**TO2.** To validate the technologies developed in MULTISENSOR by implementing a number of representative **scenarios of pilot use cases** that target **international media monitoring** and **international investment decision support**.

**TO3.** To deliver an **operational demonstrator**

**BO1.**To define a **business model for the exploitation of MULTISENSOR results by the partners and the consortium as a whole.**

The MULTISENSOR Consortium consists of nine partners, each of them contributing their competence and experience in one or several fields crucial for the success of the project.

| Participant no. | Participant organisation name | Part. short name | Country |
|---|---|---|---|
| 1 (CO) | Centre for Research and Technology Hellas, Information Technologies Institute | CERTH | Greece |
| 2 | Universitat Pompeu Fabra | UPF | Spain |
| 3 | Fundacio Barcelona Media | BM-Y! | Spain |
| 4 | Linguatec | LT | Germany |
| 5 | Everis Spain SLU | EVERIS | Spain |
| 6 | Pressrelations | PR | Germany |
| 7 | Ontotext | ONTO | Bulgaria |
| 8 | Deutsche Welle | DW | Germany |
| 9 | PIMEC | PIMEC | Spain |

### *3.1.2 Summary of activities during the first year*

During the first year of its lifetime, MULTISENSOR achieved considerable progress with respect to all of its objectives. Thus, empirical studies of the material were performed and the technical infrastructure has been set up for all modules. During this period, the user requirements have been defined, based on which the MULTISENOR technical specifications have been set and the platform architecture has been designed. This architecture was realised by an operational prototype, which integrates the skeleton versions of the research modules, and will serve as fundament for the work in the Project in the second year. In addition, in the context of each individual objective, further targeted actions have been performed.

For **SO1** the developed technologies concerned the name entity extraction module, the deep syntactic parser, the speech recognition and the machine translation module, as well as the concept extraction mechanisms from multimedia.

With respect to **SO2**, the implemented technologies included the context extraction module, the infrastructure for sentiment classification and a module for contributor analysis in social media.

In the context of **SO3**, we have implemented a framework for ontology alignment based on the weighted fusion of existing ontology matchers. Furthermore, a baseline module for topic-based classification has been developed realising multimodal feature weighted classification techniques. Finally, a representation model for socially interconnected and multimedia enriched objects has been introduced.

With respect to **SO4**, the basic semantic representation and infrastructure was defined, covering a variance of ontologies and linked open data loaded in the semantic repository of the infrastructure.

In the context of **SO5**, the developed technologies included the basic summarisation infrastructure and the extractive summarisation pipeline, which integrates a set of Simple Multi-document Summarizer (SUMMA) modules.

In addition, substantial progress has been reported in the technological objectives. In the context **TO1** a prototype social media crawler has been implemented, while the crawler of a commercial media monitoring company (PR) has been integrated to the crawling framework of MULTISENSOR.

Following the workplan, with respect to **TO2,** three main use cases have been designed relevant to journalism, commercial media monitoring and SME internationalisation and based on these the user requirements have been extracted. The main scenarios of the three use cases are the "European energy policy", the "household appliances market" and the "market analysis and regulation" respectively.

Towards the achievement of **TO3**, the operational prototype of the MULTISENSOR platform has been implemented integrating the skeleton of the content analysis services developed in the context of SO1-SO5.

Finally, with respect to **BO1,** a detailed analysis of media monitoring and SME markets has taken place and business models were investigated for the exploitation of the project's results.

During this period, the consortium has been also very active in increasing the awareness about MULTISENSOR and in the dissemination of the MULTISENSOR objectives and its achievements. In total, during the reporting period, fifteen papers related to MULTISENSOR have been published (reported in WP9), while another five have already been accepted for publication after the reporting period. The consortium has been also very active in interacting with the research, the media monitoring and the SME communities, as well as the society by a) participating in more than ten cluster events; b) issuing two pressreleases and c) collaborating with more than five European projects.

Overall the project progresses as planned, with good collaboration and team spirit among its partners. The consortium achieved to have a common goal and understanding of the project objectives, the objectives of the first year were achieved successfully and corrective actions took place to address any deviations.

### *3.1.3 Future work*

In the next two years of its lifetime, MULTISENSOR will carry on the research and development tasks related to the individual scientific, technological and business objectives and the accomplishment of the milestones.

Specifically, for the second year, MULTISENSOR moves forward towards two important milestones: a) the implementation of the first MULTISENSOR prototype (MS3) integrating the initial versions of the research modules and its evaluation; b) the development of the second MULTISENSOR prototype (MS4), which will integrate the advanced versions of the research modules.

Finally, during the third year, the project will focus on the achievement of the fifth milestone (MS5), which is the development of the final MULTISENSOR platform and its evaluation.

### *3.1.4 Expected final results and potential impact*

The expected final result of MULTISENSOR is a platform that will support a) the journalists in mastering large and heterogeneous content in order to prepare articles and identify topics, relevant and contradictory content, as well as have access to multilingual multidocument summaries; b) the commercial media monitoring companies to summarise the opinions of people for specific products and c) the SMEs that want to internationalise by providing market analysis, product reports and decision support services. This platform will integrate innovative modules, which could be

separately exploited. These include modules for name entity extraction, speech recognition, machine translation, concept detection from text and multimedia, topic classification and detection, multimedia retrieval, semantic search, ontology alignment, reasoning and decision support, as well as summarisation.

In its final stage of development, MULTISENSOR is expected to have a big impact from several perspectives. First, as discussed above it will support actively the journalist (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 general in Europe, since the mobility of SMEs will be facilitated. Finally, the social impact of MULTISENSOR will be the production of cross-validated news articles and the presentation of news stories from different cultural, political and linguistic perspectives.

### *3.1.5 Contacts and Information*

For further information visit the project web site http://www.multisensorproject.eu/ or send e-mail to the Project Coordinator Dr. Ioannis Kompatsiaris, ikom@iti.gr or to the deputy Project Coordinator & Scientific Manager Dr. Stefanos Vrochidis, stefanos@iti.gr. You are also welcome to join us in social media:

Facebook: https://www.facebook.com/pages/Multisensor/1481238748826033

Twitter: https://twitter.com/multisensor

LinkedIn: http://de.linkedin.com/in/multisensor