INFSO.E2
Technologies for Information Management

Annual Public Report 2010

European Commission
Information Society and Media

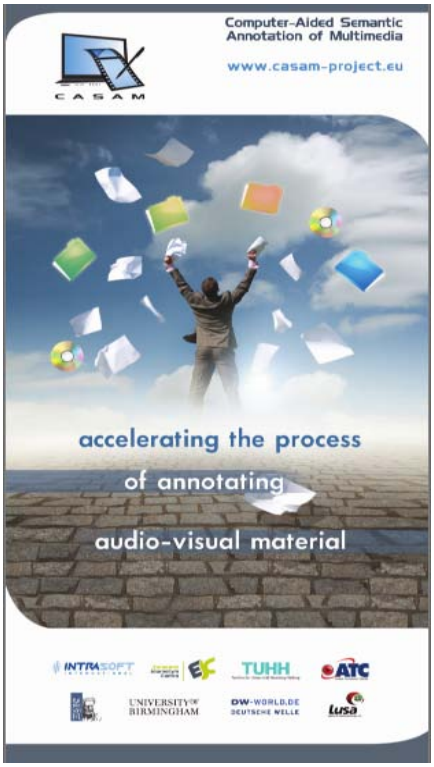
CASAM: "Computer-Aided Semantic Annotation of Multimedia"

3rd Annual Report (D1.4.3)

:: This report covers the period of December 2009 to November 2010 ::

www.casam-project.eu

Project Description



The need for annotated multimedia content is more demanding than ever before. The fully automated annotation has made recently large progress but still faces limitations due to unsolved technical issues. Thus, the machine-only annotation appears immature although presenting promising achievements. On the other hand, manual (human) annotation is very cost-prohibitive.

CASAM project is focusing on facilitating the synergy of human and machine intelligence to significantly speed up the task of human-produced semantic annotation of multimedia content. The project research focus lies in the domains of Reasoning for Multimedia Interpretation (RMI), Knowledge-Driven Multimedia Analysis (KDMA) and Human-Computer Interaction (HCI).The project is developing an annotation tool that will augment machine knowledge with human input with the target of minimizing user effort.

The annotation tool will be able to function within the modelled domain of **news production of News Agencies and Broadcasters**. However, the methods that will be developed will not be bound to the chosen domain, but will be also applicable for the annotation of multimedia documents in a variety of contexts, ensuring generality of the system's usage.

Summary of Activities

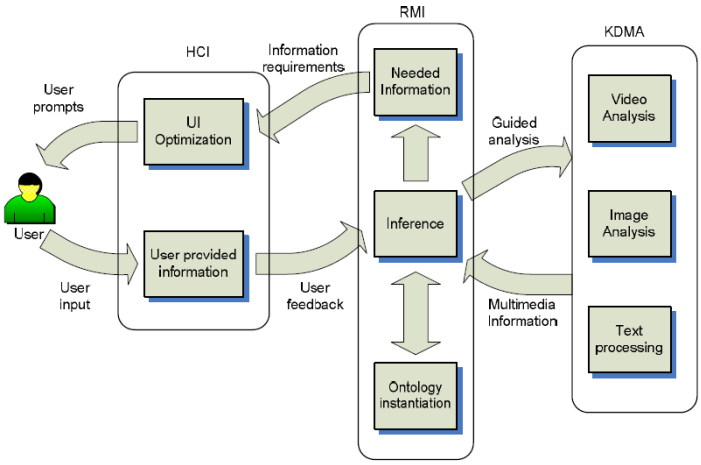
CASAM project is already well within its third and final year and all scheduled activities and research aspects are progressing as anticipated and coordinated according to the predefined project time plan.

The execution - with considerably promising results - of the initial end-user evaluation process based on the initial integrated CASAM prototype, along with the finalization of the implementation of the second generation core components of CASAM system, resulted into the successful completion of: (a) the meta-level reasoning engine (RMI), (b) the machine learning module (ML), (c) the revised versions of the multimedia analysis toolkit (KDMA) and the user interface (HCI) integrated prototype, (e) the improved integration and orchestration platform (IP), and (f) the new additional features related to the semantic search facility and the user - content management facility are the **latest major project achievements**.

Furthermore, reports related to the meta-level reasoning for disambiguation and preference elicitation, evaluation of learning algorithms, adaptive methods for multimedia document analysis, usability evaluation of revised user interface, integration with third-party systems, and Quality of Service (QoS) and recovery management have been produced. Finally, actions towards the integration and testing of the final CASAM prototype are towards the end of their realisation.

During the forthcoming critical period, the project aims to achieve the successful completion of the final integrated CASAM prototype, the measurement of the effectiveness and acceptance of the final prototype through the last round of user evaluation procedures, and finalise the business plan and exploitation strategy for the potential commercialization of the project.

CASAM Prototype Evaluation and End-User Acceptance



The results of the initial prototype evaluation provided a clear and overwhelming approval of the CASAM concept and intention. The system was deemed useful and valuable for the daily activities performed within the multimedia and broadcast environment. The user interface and general usability was appreciated at large resulting at the same time into rich recommendations and additional improvements that could be featured into the final CASAM integrated prototype.

The initial CASAM prototype was supposed to “prove the feasibility of the approach” during evaluations and at the same time provide feedback on functionality and usability to help guide the remaining development process. Both these objectives were fully achieved. Even more, end-users clearly confirmed that CASAM’s simplicity and quality of annotation process are already now at a significantly improved level

over typical approaches used in practical nowadays multimedia working environments.

In the final stages of the project, work has focused on the improvement of the annotation quality as well as the system's response time. Efforts were also put towards an approach to convincingly showcase business and integration opportunities.

Definition of User Interface, Interface Intelligence and User & Domain Models

A final version of a fully functional prototype of the HCI component (comprising of the User Interface and its low-level implementation to support interoperability with the rest of the CASAM components) has been adapted and implemented based upon feedback from user studies and the initial end-user evaluation of project's piloting phase. One of the major extensions featured in the final graphical user interface is the incorporation of disambiguation queries produced by the meta-level reasoning engine.

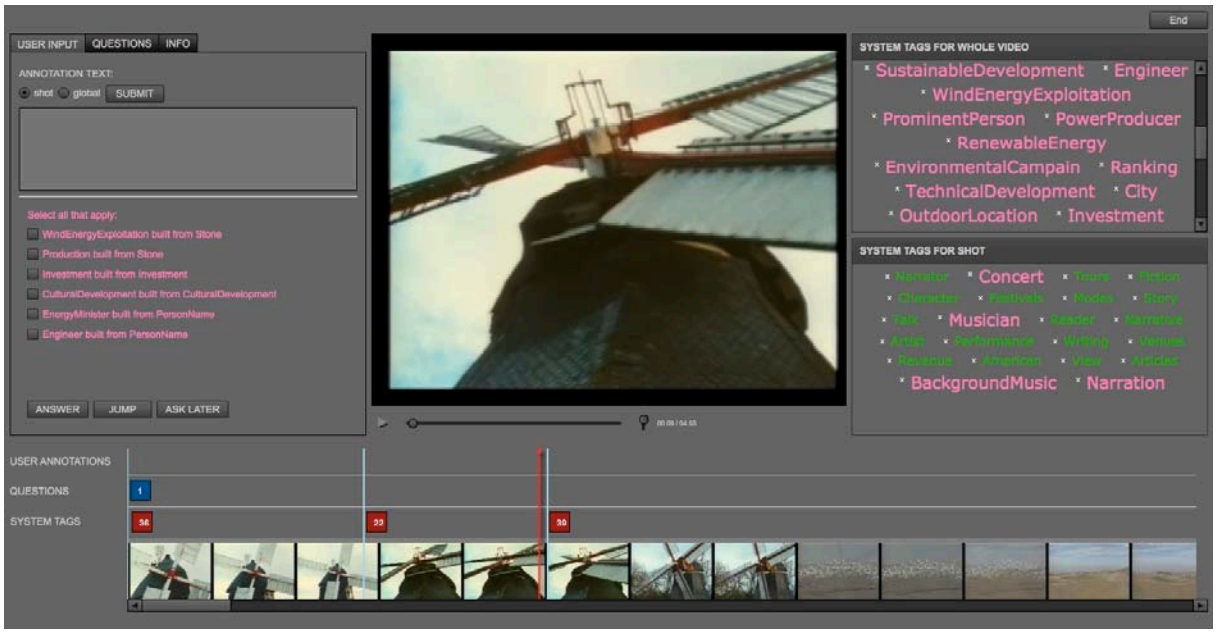


Figure 1: CASAM User Interface final prototype

Describing the updated user interface of CASAM, there are **five** distinct areas appearing on the application's window:

- In the top-left section of the interface, in a tabular section there is a **query area** that displays disambiguation queries produced by the system and handles user interaction towards these queries, and a **user annotation area** allowing users to enter free text annotations,

- a **tag cloud** (top-right) that dynamically changes depending on the content currently being displayed within the video, featuring the annotations automatically produced by the system both at a global as well as video shot level. Users can review and edit current annotations that they have made themselves, or that the system has autonomously generated,

- a **video playback area** (top-middle) which enables users to watch the video that is currently being annotated, and finally

- a **video timeline** (bottom) with shot boundaries identified which is a visual timeline that contains representative key frames for different shots and helps users in navigating to important sections of video for review, and generally in getting a sense of the video's contents.

Furthermore, a series of usability studies have been conducted in order to evaluate and test the updated version of the CASAM HCI prototype with respect to the design, intuitiveness and ease-of-use. Additional evaluations of the adaptive aspects of the HCI prototype will also be undertaken in order to ensure the efficiency and adaptability potentials of the module.

Technological Overview and Innovation Aspects

The final prototype of CASAM's integrated system design includes a conceptual process design following the SOA paradigm reflecting at the same time the technological approach for the implementation of the integration platform. The integrated architecture of CASAM system is described at a conceptual and functional level, focusing on: (a) the role and the functionalities performed by each of the constituting components, (b) the interactions among them along with their relevant processing workflows, and (c) the domain model providing the data types that may be exchanged within these workflows.

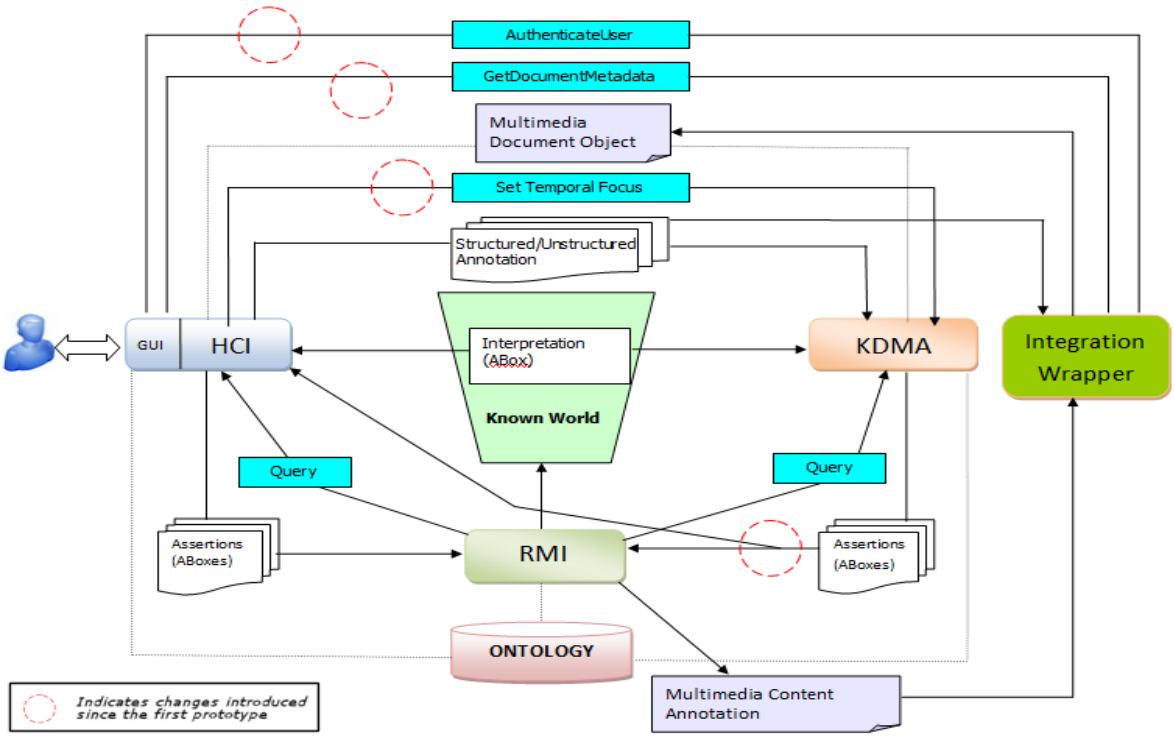


Figure 2: Component interactions within CASAM – A conceptual approach

Following the interoperability concept, CASAM architecture conforms to SOA paradigm supporting BPEL Process Orchestration, Asynchronous Service Invocation, etc., that facilitate the operational requirements of the components, as well as the desired functionalities that have been specified for the integrated system. Each module is regarded as ready-to-use "service" offering specific functionalities

according to their nature. The task of the integration platform is to interconnect the three core subsystems of CASAM (namely HCI, RMI and KDMA) within an infrastructure that can enable their cooperation for the execution of the required usage scenarios. The three subsystems are in fact three independent, reusable software platforms. Each one is designed and developed independently, using the design patterns and technologies that best suit the functional requirements of each component.

Compared to the basic prototype of the integrated CASAM system, the following modifications and additions were introduced into the final system: (a) user authentication and authorization, (b) IPTC metadata are now accompanying each multimedia document, (c) Machine Learning functionality which aims at enhancing the annotation process, (d) user and multimedia content management facility, and (e) a semantic search module which enables users to perform high level search operations within the metadata document corpus and retrieves a fully ranked list of multimedia segments matching the search criteria.

In the framework of Knowledge-driven multimedia analysis, an improved version of the respective toolkit (KDMA) is now available. This tool is organized in a number of sub-modules and a controller component. The controller component handles the flow of information between sub-modules and acts as a wrapper among the KDMA's sub-modules and the rest of the CASAM components. The core analysis modules are responsible to handle audio, video and text data respectively, while fusion module augments the extracted information by jointly considering the information extracted by the other modules. With respect to the initial version of KDMA, progress has been made mainly with respect to the adaptability of the multimedia analysis to user input as well to the adaptability of one medium analysis to the other. Moreover, a number of other features have been added, such as automatic video classification, speaker clustering, integration of new background textual knowledge sources and a new video-oriented OCR engine.

Regarding CASAM knowledge representation and reasoning for multimedia interpretation, the improved version of the respective module (RMI) incorporates a mechanism for meta-level reasoning with the aim to disambiguate interpretation alternatives. This is achieved by generating queries from a set of interpretation alternatives and forwarding this data to HCI for presentation to the user.. Queries themselves are ranked by a weight value, listed in order of importance for the disambiguation process. Furthermore, a machine learning mechanism has been incorporated in order to improve the learnability of the platform and improve produced annotation over time.

User Involvement, Promotion and Awareness

In general, CASAM project continues to receive a lot of interest from the **targeted project users** (DW, EJC and LUSA), who acclaimed the results of the project. Major user involvement is foreseen during the following months of the project when the final integrated CASAM prototype will undergo its final end-user evaluation phase. With reference to the final end-user evaluation phase, the updated evaluation plan has been drafted, outlining the employed theoretical approach and methodology to be followed.

Co-operation activities in the context of CASAM, have been enhanced significant particularly over the last few months. CASAM targeted EU-funded projects and other

related initiatives with similar objectives and targets.. Furthermore, certain actions have been initiated with nominating members of IPTC in order to explore standardization opportunities through CASAM. Features with potential standardization contribution are the CASAM ontology (for IPTC's News Codes) and the CASAM annotation file metadata fields (for IPTC's NewsML-G2).

To **raise the profile and visibility** of the project, primarily the web site of CASAM has been enhanced substantially, via refining the content featured within the web site, with particular attention to the news and events related to the project, publications to various conferences and journals, etc. Two of the major new features are the launch of the CASAM LinkedIn Group and the publication of a stylish multimedia simulation demonstration video of the initial prototype. Statistics provided by Google Analytics portray an increase in the interest of a wider audience towards the progress and developments of the CASAM project.

In the broader academic community, numerous scientific research papers have been submitted and presented in relevant to the field of our research conferences and journals.. Furthermore, CASAM project had an active presence into the EU ICT 2010 event - Europe's most visible forum for ICT research and innovation. The presence of CASAM within the exhibition has been quite successful, raising a lot of interest regarding the projects' objectives and outcomes. Finally, CASAM successfully organised a special session at the 5th International Conference on Semantic and Digital Media Technologies (SAMT 2010). All details about the project publications are available via the project website at http://www.casam-project.eu/?Page=publications.

To further raise awareness of the project, the consortium has provided a lot of additional dissemination material including: (a) Pens carrying the CASAM project logo, (b) Books for notes, (c) USB Flash Drives, (d) T-Shirts, and (e) CASAM Poster. Finally, at the end of the second project year (March 2010), an updated project presentation has been created in the context of the Project Showcase.

Future Work or Exploitation Prospects

The following months of CASAM project are very important as the final prototype of the integrated system will be finalised and the final end-user evaluation process will be conducted. Important outcomes and results from this evaluation phase will prove to be vital for future research initiatives related to multimedia annotation. A summary of the most **important future work objectives** includes:

- Finalisation of the final Integrated CASAM system (second prototype)
- End-user Evaluation Report for the final prototype
- Final report on HCI methods
- Test cases and overall system evaluation results
- Final project activity report, project presentation, and project showcase

Moreover, the final **exploitation prospects** of the project have been presented in the intermediate "Business and Exploitation plan" which describes and analyzes the CASAM business idea, CASAM business model, the project exploitable assets (tangible and intangible) and the partners individual exploitation plans. It analyzes the ICT and IT markets, addresses the potential of the market trends relevant to CASAM, the main competitors as well as the CASAM Strengths, Weaknesses, Opportunities and Threats. Nevertheless, market trends and competition will be monitored for all the duration till the project end and the commercial strategy will be refined accordingly.

Further Information

Further information can be obtained directly from the web site of the project. On the project web site, www.casam-project.eu, you can download the following:

Project Presentation:

http://www.casam-project.eu/files/CASAM_D2.6b.pdf

Project Factsheet:

http://www.casam-project.eu/files/CASAM_Project%20Fact%20Sheet.pdf

CASAM leaflet:

http://www.casam-project.eu/files/CASAM_Leaflet_Oct2008.pdf