M-ADVANTAGE

Multimedia - Automatic Digital Video & Audio Network Through Advanced Publishing European Service

Andrea de Polo on the behalf of the M-ADVANTAGE Consortium
Affiliation: Fratelli Alinari Photo Archive, Florence, Italy

Abstract: The research and development efforts within the M-ADVANTAGE project described in this paper aim at increasing the competitiveness of Europe's digital content industries by semantic-based services across the content value chain, including personalized delivery. Nowadays, user needs are addressed by costly solutions that require intensive human intervention. The described activities strive at filling the gap in automatic processing of multimedia by creating an intelligent infrastructure allowing considerable productivity gains. For achieving this goal, it is proposed to carry out a tightly integrated research and development activities also in terms of the blend of research, technology, content and user partners involved.

The research and development targets to build a service infrastructure for automated semantic discovery, extraction, summarization, labelling, composition, and personalized delivery of content from heterogeneous multimedia repositories. This will involve foundational research such as data models and ontologies required for merging multiple heterogeneous data types into an integral representation; component-level research for parts of the service infrastructure; as well as development of semantic-based productivity tools. Making use of established Semantic Web, multimedia description and other standards are anticipated to enable a broad uptake of M-ADVANTAGE's open source and non proprietary technologies.

While the project's research & technology partners include leading university centres and industry players (including SMEs), on the content side renowned private as well as public organisations are involved. They hold and provide access to all types of content such as audiovisual, stock image archives, news agency and other content, and strive to develop and market knowledge-based content services.

Key words: service portability, service adaptation, network interoperability, context awareness, semantic ontology, industrial Semantic Web environment

1. STATE OF THE ART

M-ADVANTAGE, being an integrated project, has a wealth of tasks, processes and scientific and technological objectives. Due to the nature of the work, these can be grouped under the umbrella of intelligent multimedia analysis and access with the use of ontological information. Thus, the most relevant state of the art is that related to the development of ontological knowledge representations for multimedia applications as well as those related to multimedia analysis and access approaches.

As far as representation is concerned, the MPEG-7 standard, formally named "Multimedia Content Description Interface", provides a rich set of standardized tools to describe multimedia content. However, in order to make MPEG-7 accessible, re-usable and interoperable with many domains, the semantics of the MPEG-7 metadata terms need to be expressed in an ontology using a machine-understandable language. Additionally, there is an increasing need to allow some degree of machine interpretation of multimedia information's meaning. To this end, several approaches in the literature address the problem of building multimedia ontologies to enable the inclusion and exchange of multimedia content through a common understanding of the multimedia content description and semantic information.

Hunter [1] describes the trials and tribulations of building a multimedia ontology represented in RDF Schema [2] and demonstrates how this ontology can be exploited and reused by other communities on the semantic web (such as TV-Anytime [3], MPEG-21 [4], NewsML [5], museum, educational and geospatial domains). A core subset of the XML-based MPEG-7 specifications together with a top-down approach to generate the ontology is used. The first step is to determine the basic multimedia entities (classes) and their hierarchies from the MPEG-7 Multimedia Description Scheme (MDS) basic entities [6]. The RDF schema semantic definitions for MPEG-7 can be linked to their corresponding pre-existing MPEG-7 XML schema definitions. Additionally, the RDF Schema can be merged with RDF schemas from other domains to generate a single "super-ontology" called MetaNet. Expressed in DAML+OIL [7], MetaNet can be used to provide common semantic understanding between domains. This super-ontology can be used to enable the co-existence of interoperability, extensibility and diversity within metadata descriptions generated by integrating metadata terms from different domains.

The proposed method for building a multimedia ontology has been applied to manage the manufacturing, performance and image data captured from fuel cell components [8,9]. Future work plan of [1] includes the automatic semantic extraction from the MPEG-7 XML schema document as well as the automatic linking of the semantics to the XML schema document. Acknowledging the importance of coupling domain-specific and low-level description vocabularies, a similar methodology for enabling interoperability of OWL domain-specific ontologies with the complete MPEG-7 MDS is described in [10]. The approach is based on an OWL ontology, referred as a core ontology, which fully captures the MPEG-7 MDS. For the development of the core ontology, a set of rules is defined to map particular MPEG-7 components to OWL statements. The integration of the domain-specific knowledge is performed by considering the domain-specific ontologies as comprising the second layer of the semantic metadata model used in the DS-MIRF framework (the first layer is encapsulated in the so called core ontology). For this reason a set of methodological steps is provided. Additionally, rules are provided for transforming the OWL/RDF metadata, structured according to the core ontology and the domain-specific ontologies, into MPEG-7 compliant metadata. Following this approach proves advantageous for MPEG-7-based multimedia content services, such as search and filtering services, since incorporating semantics can lead to more accurate and meaningful results in terms of meeting the user queries.

The most conventional approach is the keyword approach, which is based solely on textual metadata annotation. Such search technologies often exacerbate information overload; although they can identify documents in which a search term appears, they cannot tell how relevant the document is to the subject being researched. They simply look for the occurrence of keywords and are unable to decipher whether the concept represented by a search term is related to the main idea of a document. This approach follows closely the developments in the field of simple text retrieval [11], which has not progressed much since 1999.

On the other hand, semantic indexing aims at finding "patterns" in unstructured data (documents without descriptors such as keywords or special tags) and use these patterns to offer more effective search and categorization services [12]. Semantic indexing techniques are language-agnostic, so data collections do not have to be in English, or even in any specific language. This approach comes closer to bridging the semantic gap, seen as the discrepancy between the capabilities of a machine and that of a human to perceive visual content. In order to understand the content of an image, one necessary step is to identify objects within it. The aim is not absolute recognition of each individual object in the image but to enable similarity search on image parts immersed into various contexts. A possible

route to achieve content understanding is the direct and automated extraction of textual descriptors from visual content directly.

In most auto-annotation efforts, prediction is done at keyword (or concept) level and each concept is predicted independently from the other. It is therefore possible to obtain incoherent predictions such as "space" and "indoor" simultaneously when describing the same image. There is need for a global maintenance of semantic coherence between parts of the annotation. This clearly requires the use of a consistent and normalized multimedia description scheme, which will be defined as a formal structure of digital meta-publication, where digital meta-publication means a set of connected digital objects (text, audio, video, etc.) with a strict hierarchy, advanced metadata information and other sophisticated possibilities.

When it comes to retrieval, semantic multimedia retrieval requires the presence of an already annotated multimedia content. There are several types of semantic retrieval, all of which utilize semantic matching algorithms between the semantic content descriptions. The first type of semantic retrieval is based on direct description / definition of the Semantic Track of the target data by a user. Such process of semantic definition requires a user-friendly interface with features for Ontology browsing. Precise results may be retrieved via specifying the significance of semantic features. A second type of semantic retrieval is based on describing / defining the Semantic Track of the target data by a user indirectly through defining the initial similar multimedia data. Thus, a sample of media file with a set of extracted mathematical features is used as an input query. Precise results may be retrieved via specifying the significance of mathematical features. The third type represents a kind of combination of the previous two semantic retrieval types.

Using a combination of Bayesian Inference and Signal Processing Technology (SPT, Shannon's Information theory), can indeed help in the automatic extraction of key conceptual aspects of any piece of unstructured information (documents, web pages, emails, voice, videos, images ,etc). Bayesian Inference is a mathematical technique for modelling the significance of semantic concepts (ideas) based on how they occur in conjunction with other concepts. By applying contemporary computational power to the concepts pioneered by Bayes, it is now feasible to calculate the relationships between many variables quickly and efficiently, allowing software to manipulate concepts.

Information Theory provides a mechanism for being able to extract the most meaningful ideas in documents, thus leading us to the definition of a "pattern matching" technology. Information Theory is the mathematical foundation for all digital communication systems. Natural languages contain a high degree of redundancy. A conversation in a noisy room can be

understood even when some of the words cannot be heard; the essence of a news article can be obtained by skimming over the text. Information theory provides a framework for extracting the concepts from redundancy. Shannon's theory is that "the less frequently a unit of communication occurs, the more information in conveys". Therefore, ideas, which are rarer within the context of a communication, tend to be more indicative of its meaning.

The Pattern-matching approach has the additional benefits: (a) it is robust to false positive matches and (b) it can determine how similar documents are, without both documents being tagged the same way, or even tagged at all; this is called idea distancing.

In the next sections of the paper we will provide and in depth description of the technical and scientific solution that we would like to achieve within the project.

2. SCIENTIFIC AND TECHNICAL OBJECTIVES

The M-ADVANTAGE project aims at developing an infrastructure capable of delivering multimedia information and content customized to the needs of end-users. It focuses on building some specific components to provide the functionalities necessary to facilitate the construction of advanced multimedia content applications and the use of structured and unstructured multimedia information.

The goal of the M-ADVANTAGE approach to the "delivering multimedia information and content customized to the needs of end-users" is based on three ambitious deliverables:

- M-ADVANTAGE is able to automatically integrate heterogeneous multimedia content. Since the integration is automatic as a result the M-ADVANTAGE infrastructure is highly scalable and will be able to expand the current 6 content flows to an unlimited number, simply by adapting the hardware infrastructure, accordingly.
- 360° Technology Approach: M-ADVANTAGE infrastructure is based on the more up-to-date technology approaches for managing unstructured information: Keyword, Semantic and Statistical (through a pattern matching system).
- Develop specific application services to deliver the content managed by the M-ADVANTAGE back-end infrastructure

These features will enable the utilization of digital content delivery systems distributed across the computer network and will process the information stored within these archives in order to find dependencies, links and similarities between various pieces of information. This will allow to

automatically manage and customize the available content for the needs of end-user applications built on top of the M-ADVANTAGE infrastructure.

From the scientific point of view, the following contributions are expected from the M-ADVANTAGE project to the research community:

- Automated multimedia (semantic) discovery, which concerns both retrieval, i.e. search for multimedia files; and extraction, i.e. more focused search for specific structural components of the multimedia: episodes, frames, images (focuses), etc.
- Advanced video summarization, i.e. content of the whole video clip can be browsed quickly.
- Advanced techniques for semantic labelling, i.e. propagation of labels through hierarchical database structures.
- Automated multimedia integration / composition: real power is in composition of different structural elements (episodes, frames, focuses) extracted from heterogeneous multimedia files in a coherent track.
- Semantic personalized delivery: based on semantic interactions of user activities / actions on content and user's explicit preferences; proactive supply to the user of relevant multimedia.
- Interoperability between heterogeneous (web-) services and multimedia: this is possible following Semantic Web's recommendations about common (upper-) ontology or managing mapping between semantic concepts from different ontologies.

From the technical point of view, the M-ADVANTAGE platform aims at creating a state-of-the-art cutting edge technology that is going to serve public and business sector in the Knowledge Management for multimedia content. In this respect:

- Statistical search will be used as a super set of the conventional methods in the sense that where these should fail it will always be possible with this methodology to grasp concepts embedded in images, text and videos together and deliver a complete content overview on somebody's search.
- The combination of semantic / ontology methodologies and the statistical one will offer users the possibility to have a much more precise and to the point interaction with the KB. Users will be profiled and grouped into communities according to their previous interactions with the KB.
- Content will be classified automatically according to concepts that undoubtedly identify it. It will also be possible to split any video content into it's fundamental scenes using "scene detection" and "object extraction" techniques, thus allowing editors the possibility to reassemble a piece of video according to their needs. It will also be possible to search on the text extracted from a speech in a video using the most technological

approach of speech to text technology, and returning the meaningful frames that relate to the search argument. All of these tasks shall be carried out automatically without the user noticing it.

Overall, the goal of the M-ADVANTAGE platform is to offer users a multimedia access experience that combines all that is needed to one time visitors and to professional users.

3. OUTLINE IMPLEMENTATION PLAN

M-ADVANTAGE platform intends to provide an integrated solution for the B2B value chain starting from the content owners (image archives, public domain digital libraries, multimedia online deposits, etc), passing through the added-value content creators (press agencies, publishers, creative sector, etc) and arriving to the service providers (Internet portals, broadcasters, news, etc).

This can be broken down into the segmentations described in Figure 1 representing an in depth view of the content value chain that M-ADVANTAGE intends to address.

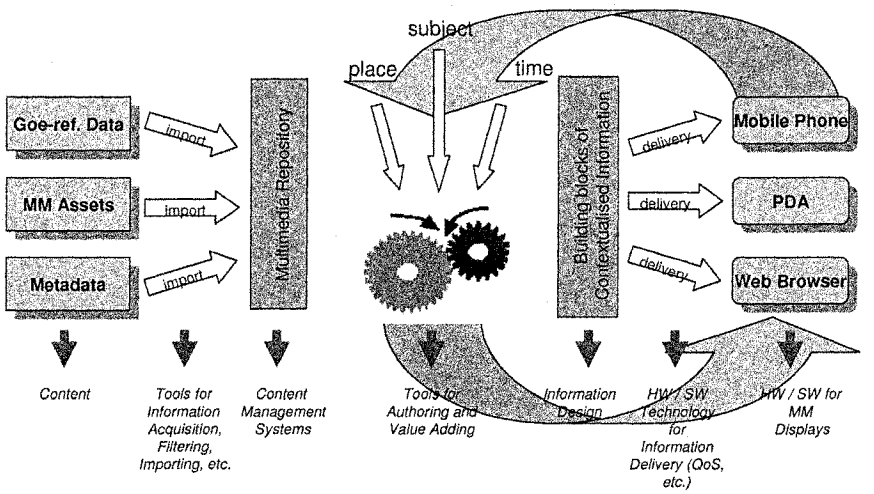

Figure 1. M-ADVANTAGE services from a technical point of view. In this sketch we see from the left to the right the workflow of the project, from input (content providers) to output (new services, enriched content and so on).

From the point of view of the users, on the other hand, what M-ADVANTAGE mainly contributes is the consideration of knowledge in the

process of access to multimedia content, as described in Figure 2. The integrated wizard helps the online editor or content creator to enrich its valuable multimedia material with innovative and unique features.

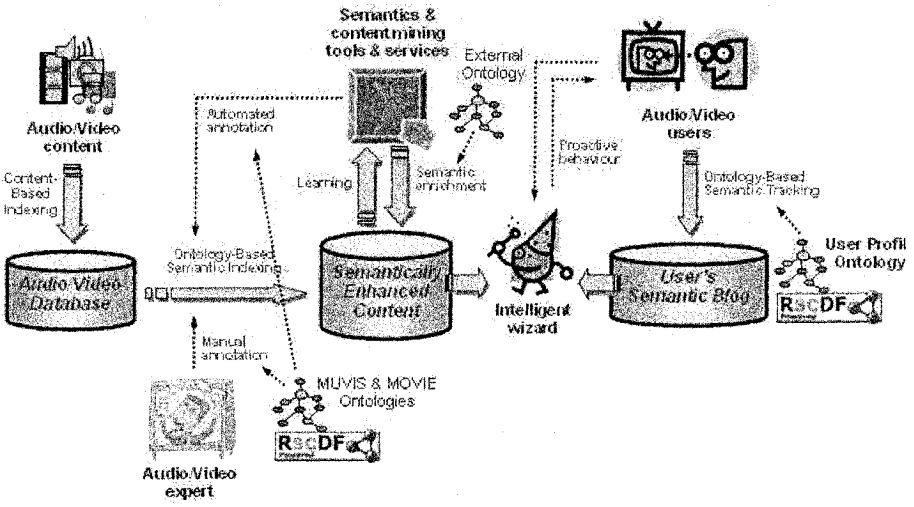

Figure 2. M-ADVANTAGE services from the users' point of view

M-ADVANTAGE platform is targeted to improve the industrial processes of the various actors involved in the generation of multimedia information services, by:

- Providing access to a larger amount on multimedia information,
- Simplifying search activities,
- Offering an integrated Digital Rights Management System (DRMS)
- Providing a Customized Multi Licensee Service for different level of subscribers, partners and end users with different kind of features and permissions available, according to the subscription level
- Offering a secure online payment mechanism, including a strong mechanism to guarantee the user's privacy
- Pay per View
- Automatic tasks to enrich "poor" or unclassified items with enriched multimedia features
- Offering customized and personalized search environments, using, for example:
    - Online wizard: product creation,

- Smart on line assistant: 3D tutorial aid character,
- software tracking for user behaviour

The M-ADVANTAGE platform is intended to be a basis for a wide set of tools. These tools will satisfy the different needs of the different actors involved, taking in consideration their different business approaches (profit private organization V/S public non-profit bodies), their different technological situation (on-line, fully digital, advanced business V/S traditional archives with limited informatics support), their different vocations (highest protection of rights V/S widest accessibility to contents).

M-ADVANTAGE aims to research, develop, implement, integrate, and test with users, the enabling technologies required to realize the M-ADVANTAGE concept. In order for all these aspects, as seen from the users' and technical points of view, to be supported, M-ADVANTAGE needs to integrate and/or develop a wide range of tools and services, as briefly summarized in Figure 3. As most of them (one may consider automatic indexing of any type of media as an example) are complex services the implementation of which remains in some cases an open research issue, the work is organized in a set of work packages, some of which have a research oriented nature, rather than a pure development and integrating one.

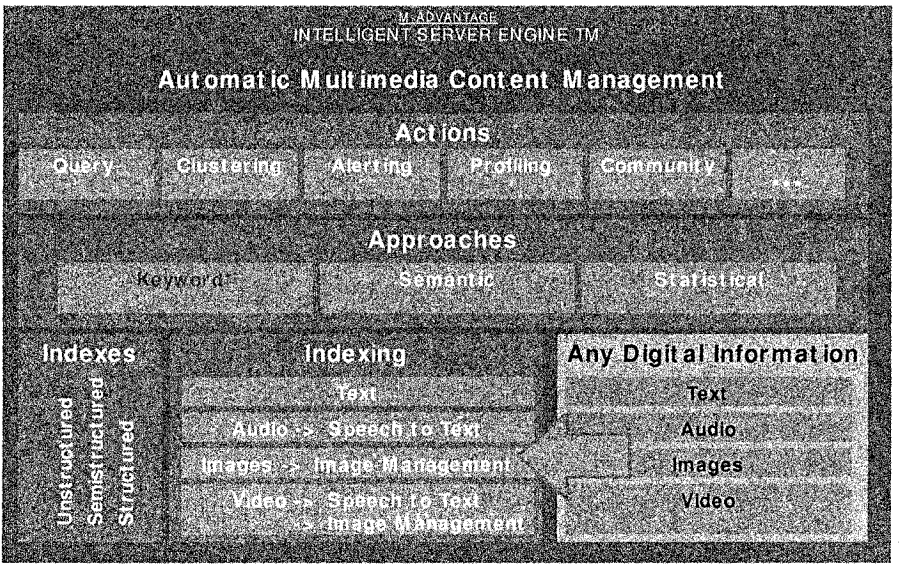

Figure 3. Business services included in M-ADVANTAGE available for the members of the system

Figure 4 below shows the overall architecture proposed for the M-ADVANTAGE platform, and the way in which the results of the various work packages come together in order to meet the projects objectives. In the following sections we will explain how the architecture's components (A-D) will be developed.

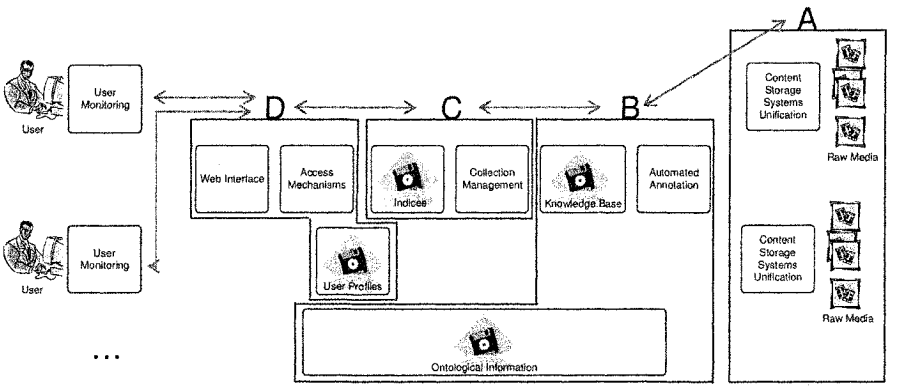

Figure 4. M-ADVANTAGE architecture and its related workflow.

Development of component, service 1

One of the main objectives of the concerned research is to make possible the access and consideration of heterogeneous archives, in order to provide a platform able to unify the wealth and diversity of archives currently existing and/or under development. Thus, the first service considered in the architectural design of the platform is that of the analysis of existing content storage systems and the specification of a generic querying and access interface, able to support and serve all existing content. In this process the needs of subsequent services will also be considered in order to make sure that no constraints on the ability to provide efficient and intelligent services is generated as an artifact. Based on this generic interface, it will be possible to create software interfaces, custom to each archive, that will allow for the automatic connection of the archive with the overall M-ADVANTAGE system.

Development of component, service 2

The most important and challenging objective of M-ADVANTAGE is to contribute to the effort to bridge the semantic gap by following innovative, knowledge based approaches to semi-automatic and fully automatic media annotation. For this purpose, complex ontological data models will be developed, and organized in Mathematical feature ontology and Mathematical feature ontology, as well as a reasoning framework capable to

consider and reason with such input. This infrastructure will be used in order to automatically analyze the content acquired from the various media archives and generate the knowledge base. Various methodologies will be supported in this process, ranging from simple manual annotation to semi-automatic, retainable and adaptive computer assisted annotation and fully automated knowledge driven annotation.

Specifically for the manual annotation approach, the concept of visual representation of content is an important element and it is known that high data dimensionality (as that typically defined in multimedia processing) hinders its visualization and hence its handling. M-ADVANTAGE will perform data-adaptive dimension reduction, as opposed to classical homogeneous dimension reduction. The above structural characterization will guide dimension grouping and selection to arrive to a data representation suitable for visualization and interaction.

Development of component, service 3

The meta-publication description format used to store the analysis results in the knowledge base will also be designed in way that provides for optimal balance between effectiveness (in terms of descriptive power) and efficiency (in terms of space and processing power) for storage and consequent processes.

Clearly, M-ADVANTAGE is not designed to serve a single or a mall number of archives. Thus, characterizing the diversity of the multimedia collection at hand will be essential for obtaining a reliable cue of the available multimedia space. Typically, measuring such a parameter includes locating dense and sparse parts of the feature space, as populated by the multimedia collection. Several studies in the fields of data mining and related provide efficient tools for doing so. In M-ADVANTAGE we will also experiment with radically different approaches based on discrete global optimization procedures inherited from the field of Operations Research. We assert that not only this field is providing tools for operating transforms on the multimedia collection, as required by our context but also helps in designing gauging tools that will provide measures on the opaque collection to enhance subsequent operations.

While the above provides fundamental measurements of the collection properties and structures, M-ADVANTAGE will also address the issue in concrete terms of clustering and informative sampling, underlying the classical tasks of filtering and summarizing multimedia information collections. Again, solutions defined will directly apply to concrete issues such as that of proposing interactive multimedia catalogues of document collections. In particular, the aim is to instantiate the concept of collection guiding that extends classical browsing by creating exploration strategies

around the document collection and therefore literally guide the user through it.

The result of the abovementioned work will be the generation of information rich indices, ideal for the subsequent access procedures.

Development of component, service 4

Finally, M-ADVANTAGE aims to offer state of the art and innovative, intelligent, personalized multimedia search and access services to end users. In this respect, a state of the art content management system will be integrated in the overall platform, allowing for simple (keyword based), semantic and statistical search in the available indices; in all of these search approaches, knowledge contained in the ontological databases will also be considered. Additionally, user interactions with the M-ADVANTAGE platform will be logged and analyzed in order to extract user profiles that can be fed back into the system thus enhancing the quality of services offered to each specific user; the representation of user profiles will be based on a properly designed profile ontology, thus allowing for the optimal consideration of available ontological information in the extraction, as well as in the utilization, of the user profiles. Finally, additionally to the conventional access through web interfaces, a tool capable of operating totally unsupervised, monitoring local user activity, utilizing it to extract preference information and query for interesting documents will be integrated in the platform. M-ADVANTAGE knowledge bases will be accessible not only via the browser or the M-Assistant but also via a virtual assistant (such as on www.alfaromeo.it or www.agenttalk.nl).

4. CONCLUSIONS

The M-ADVANTAGE project aims to deliver a first version of an infrastructure capable of delivering multimedia information and content customized to the needs of end-users. It focuses on building some specific components to provide the functionalities necessary to facilitate the construction of advanced multimedia content applications and the use of structured and unstructured multimedia information.

More specifically, it will develop new formal models for knowledge representation with major focus being placed on multimedia ontological knowledge representation. Specifically, a multimodal data model will be constructed. Moreover, a whole work package will be devoted to the generation of an ontology infrastructure containing all the knowledge needed for the analysis in three main ontologies: the mathematical feature ontology (MFO) containing the knowledge about the mathematical features (low-level descriptors), the semantic feature ontology (SFO) containing the semantic

information concerning the multimedia content (the actors, directors, etc) and the user profile ontology (UPO) covering the information about the user preferences and the usage history of a specific user. Finally, merging multiple types of digital data into an integral representation is one of the main objectives. Thus, a formal data model for integration of diverse multimedia content (meta-publication) will be designed.

Also, it will provide new tools to support automatic analysis, annotation, filtering and visualization of multimedia content to the extend that this is possible. Specifically, tools will be developed for semantic annotation of the existing multimedia by human experts, collaborative online and offline learning of document concepts and (semi-)automated annotation. Tools will also be developed for management and presentation of multimedia meta-publications.

The project maximizes cross-fertilization between several different areas, including knowledge technologies, database technology, multimedia processing and so on. M-ADVANTAGE brings together a diverse consortium that, as a whole, holds know how and acknowledged scientific expertise in a variety of areas. Within the project, research results, technical practices and tools provided by the partners will be integrated. In this process, among other things, the communication between the semantic approach, the statistical approach, and the Ontology definition approach will be studied and implemented. Statistical search will be used as a super set of the conventional methods in the sense that where these should fail it will always be possible with this methodology to grasp concepts embedded in images, text and videos together and deliver a complete content overview on somebody's search.

REFERENCES:

[1] J. Hunter, "Adding Multimedia to the Semantic Web - Building an MPEG-7 Ontology", International Semantic Web Working Symposium (SWWS), Stanford, July 30 - August 1, 2001

[2] RDF Schema Specification 1.0, W3C Candidate Recommendation 27 March 2000. http://www.w3.org/TR/rdf-schema/

[3] TV-Anytime Forum, http://www.tv-anytime.org/

[4] MPEG-21 Multimedia Framework, http://www.cselt.it/mpeg/public/mpeg-21_pdtr.zip

[5] NewsML http://www.newsml.org/

[6] ISO/IEC 15938-5 FCD Information Technology - Multimedia Content Description Interface - Part 5: Multimedia Description Schemes, March 2001, Singapore

[7] DAML+OIL Revised Language Specification, March 2001.
http://www.daml.org/2001/03/daml+oil-index

[8] J. Hunter, J. Drennan, S. Little "Realizing the Hydrogen Economy through Semantic Web Technologies", IEEE Intelligent Systems Journal - Special Issue on eScience, January 2004

[9] Little, S., and Hunter, J., "Rules-B-Example – a Novel Approach to Semantic Indexing and Querying of Images", In 3rd International Semantic Web Conference (ISWC2004), Hiroshima, Japan, November 2004.

[10] Tsinaraki, C., Polydoros, P., Christodoulakis, S., "Interoperability support for Ontology-based Video Retrieval Applications", In Proceedings of Third International Conference on Image and Video Retrieval (CIVR), Dublin, Ireland, July 21-23, pp 582-591, 2004.

[11] Baeza-Yates, R.A., Ribeiro-Neto, B.A. (1999) Modern Information Retrieval. ACM Press / Addison-Wesley.

[12] Zhao, R., W.I. Grosky (2002) Narrowing the Semantic Gap-Improved Text-Based Web Document Retrieval Using Visual Features. IEEE Transactions on Multimedia 4(2).