Europe's Information Society

SIXTH FRAMEWORK PROGRAMME

INFORMATION SOCIETY TECHNOLOGIES

KNOWLEDGE & CONTENT TECHNOLOGIES

IST-2006-027538 BOEMIE

BOOTSTRAPPING ONTOLOGY EVOLUTION WITH MULTIMEDIA INFORMATION EXTRACTION

FINAL REPORT

Project duration	March 1, 2006 – February 28, 2009
Project coordinator	Constantine D. Spyropoulos (NCSR)
Web site	http://www.boemie.org/

Author of the report	Georgios Paliouras (NCSR)
Date of preparation	July 9, 2009
Version	Final version 1.0

Executive Summary

This document provides an overview of the research project BOEMIE, focusing on the main results that the project has produced.

BOEMIE was an ambitious large-scale research effort that has advanced considerably the state of the art in multimedia content analysis. In order to make multimedia content like videos or images searchable the data must be meaningfully annotated. This is commonly done by humans, but it is a hard and expensive task. Using sophisticated algorithms to extract semantics from multimedia content, BOEMIE annotates content with semantics automatically and provides valuable knowledge for both, content providers and content consumers.

BOEMIE technology fuses knowledge that is automatically extracted by most of the popular media types (audio, video, images and text). In addition to making the content richer, the extracted knowledge is used to expand our understanding of the domain (for example athletics) and extract even more useful knowledge. This knowledge takes the form of ontologies and is represented in a machine-readable format. The synergetic process of extracting semantics from content and enriching the domain knowledge is the fundamental idea of bootstrapping that BOEMIE has pioneered.

In this report, we aim to introduce the reader to the main elements and methods developed in BOEMIE. Furthermore, we illustrate the added value of BOEMIE technologies and applications for multimedia content owners and users. Interested third-parties who would like to acquire more information about the project and potentially commence a scientific or commercial collaboration with BOEMIE partners are welcome to contact the coordinator or the site managers at the addresses provided at the end of the document.

Table of contents

Table of illustrations

List of abbreviations and acronyms

AEO	Athletics Event Ontology
BOEMIE	Bootstrapping Ontology Evolution with Multimedia Information Extraction
BSB	BOEMIE Semantic Browser
BSC	Bootstrapping Controller
BSM	BOEMIE Semantic Manager
BTAT	BOEMIE Text Annotation Tool
DL	Description Logic
DYNAMITE	DYNAMIc InTEractive web pages
Ellogon	The Ellogon text engineering platform
GIO	Geographic Information Ontology
HERITRIX	The Heritrix open-source crawler
HLC	High-Level Concept
HMatch	The HMatch ontology matching tool
HTML	Hypertext Markup Language
IAAF	International Association of Athletics Federations
IMS	Image Mapping Service
MCO	Multimedia Content Ontology
MDO	Multimedia Descriptor Ontology
MLC	Mid-Level Concept
MPEG	Moving Picture Experts Group (used to refer to the MPEG standards)
MPEG-7 MDS	MPEG-7 Multimedia Description Schemes
MSM	Multimedia Semantic Model
OET	Ontology Evolution Toolkit
OWL	Web Ontology Language
PoI	Point of Interest
RacerPro	The RacerPro DL reasoner
RMDF	Recursive Media Decomposition and Fusion
TAHAT	Text and HTML Annotation Tool
VAT	Video Annotation Tool
VIA	Video and Image Annotation tool

Project objectives

BOEMIE aimed towards automation of the knowledge acquisition process from multimedia content which nowadays grows with increasing rates in both public and proprietary webs. Towards this end, it introduced the concept of **evolving multimedia ontologies**. The project was unique in that it linked multimedia extraction with ontology evolution, creating a synergy of enormous yet unrealized potential.

In recent years, significant advances have been made in the area of automatic extraction of low-level features from visual content. However, little progress has been achieved in the identification of high-level semantic features or the effective combination of semantic features derived from different modalities. Driven by domain-specific multimedia ontologies, BOEMIE information extraction systems are able to identify high-level semantic features in image, video, audio and text and fuse these features for optimal extraction. The ontologies are continuously populated and enriched using the extracted semantic content. This is a **bootstrapping process**, since the enriched ontologies are in turn used to drive the multimedia information extraction system. Figure 1 provides a graphical illustration of this iterative bootstrapping process, that is implemened in the BOEMIE prototype.

Through the proposed synergistic approach, BOEMIE aimed at large-scale and precise knowledge acquisition from multimedia content. More specifically, the objectives of the project were:

- A unifying formal representation for domain and multimedia knowledge. This **multimedia semantic model** follows modular knowledge engineering principles and captures the different types of knowledge involved in knowledge acquisition from multimedia. It realises the linking of domain-specific ontologies, which model salient subject matter entities, and multimedia ontologies, which capture structural and low-level content descriptions.
- A **methodology for ontology evolution** to coordinate the various tools that will use the extracted data to populate and enrich the ontologies. The resulting architecture uses the multimedia semantic model and specifies the interfaces for the various types of tools for ontology evolution.
- A **toolkit for ontology evolution**. Based on the ontology evolution methodology, the toolkit provides tools to support ontology learning, ontology merging and alignment, semantic inference and ontology management.

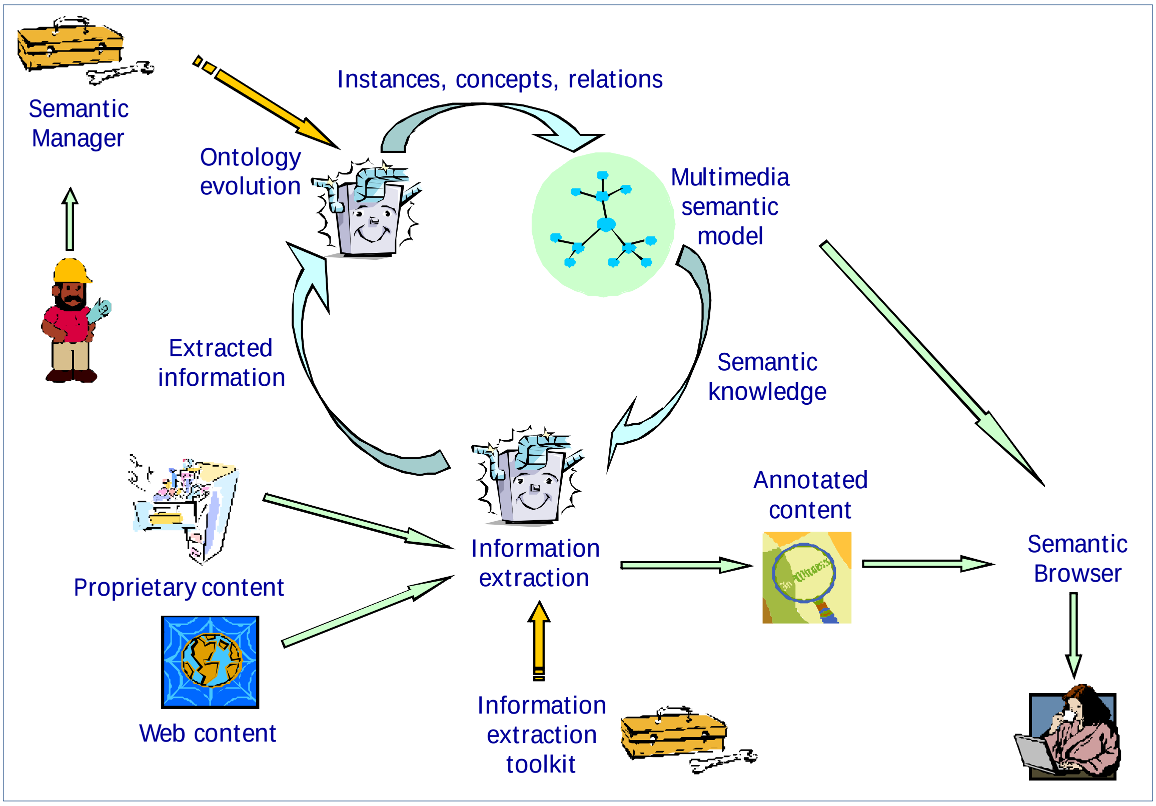


Figure 1. BOEMIE bootstrapping approach.

- A **methodology for information extraction from multimedia** content using data fusion techniques. The methodology specifies how information from the multimedia semantic model can be used to achieve extraction from various media. Additionally, it combines extracted information from multiple media to improve the extraction performance.
- A **toolkit for semantic extraction from multimedia** content. The toolkit comprises tools to support extraction from image, audio, video and text, as well as information fusion.

The resulting technology has a wide range of applications in commerce, tourism, e-science, etc. One of the goals of the project was to evaluate the technology through the development of an automated content collection and annotation service for athletics events in a number of major European cities. The extracted semantic information enriches a **digital map**, which provides an innovative and friendly way for the end user to access the multimedia content.

Who will benefit from BOEMIE

- **Content producers, owners, distributors.** Companies that are involved in the production, acquisition and distribution of multimedia content could add great value to the offering by annotating it semantically. Examples of such organizations are publishers, broadcasters, event organizers, libraries and archives.

- **Content-based service providers.** Companies that collect and analyse multimedia content, in order to draw useful conclusions for their customers could improve their quality and throughput by either receiving semantically annotated content or running BOEMIE technologies themselves. Examples of such organizations are market research and consulting companies.

- **Content management software companies.** The domain-specific semantic analysis of multimedia content could be incorporated in customized content management software by the corresponding companies. The specialization of such companies in market verticals dealing with content from specific domains is becoming increasingly popular.

- **Search engines and portals.** In their quest for new offerings and high-quality search results, search engines are turning to faceted search and multimedia content. BOEMIE technologies support faceted search of multimedia content, within specific domains. Similarly, it supports the automated maintenance of thematic portals that provide multimedia content.

- **End users.** Content consumers are currently overloaded with content, especially multimedia. The semantic annotation of such content enables a number of innovative applications and services that can assist the user in retrieving efficiently appropriate content and using it to perform a range of different tasks.

- **Researchers.** BOEMIE has generated knowledge and infrastructure that is invaluable to researchers working on multimedia analysis and knowledge technologies. Most of this is provided publically, in order to promote research further.

Summary of achievements

BOEMIE has made a significant contribution in many open research issues, introducing innovative methods and developing new tools. The following is a list of the main achievements of the project.

- **Recursive media decomposition and fusion.** A new approach to dealing with documents that combine different media (video, audio, images and text) has been developed. This has been implemented in the form of a toolkit for information extraction, combined with tools that analyse media-specific content, using methods that go beyond the state of the art.
- **Abductive multimedia interpretation.** The interpretation of multimedia documents was performed with a new reasoning method that uses the domain knowledge (ontology) abductively to assume information that may be missing from the analysis of the content.
- **Pattern-based ontology evolution.** A new approach was developed for handling the evolution of domain knowledge. It is based on the separation of the problem into a number of different population and enrichment patterns, which are further supported by an array of learning, discovery and matching tools that form the evolution toolkit of BOEMIE.
- **Multimedia semantic model.** A formal framework has been developed to capture and link knowledge related to the domain and knowledge pertaining to media specific traits in order to support the extraction,and deploymemt of multimedia content semantics. The framework comprises a number of interconnected ontology modules, which form the BOEMIE multimedia semantic model.
- **Integrated bootstrapping prototype.** The basic idea of BOEMIE, i.e. the bootrstrapping of ontology evolution with information extraction was fully realized in the BOEMIE integrated prototype, which implemented the application logic, combining the main components with utilities about content collection and indexing.
- **User-friendly ontology maintenance and editing.** One of the basic assumptions of BOEMIE is that the whole process should be as automated as possible, imposing the minimum requirements on the user. In the case of ontology maintenance, we assumed that the user is not a knowledge engineer, but rather a domain expert. Therefore, we developed innovative presentation and interaction methods to assist in the process of monitoring ontology evolution

and modifying the semantic model. These methods were incorporated in a Web-based application for the domain expert.

- **Context and map-driven browsing of semantically annotated content.** In addition to producing the semantics of multimedia content, we put significant effort into providing added value for the content consumers. We achieved that through the use of digital maps and an innovative method for context-sensitive recommendations for further reading. This was implemented in a Web-based application for the end user.

- **Manual and semi-automated content annotation.** For the purposes of producing training and test material for our methods, we developed innovative tools that facilitate quick and easy annotation of images, video and text. These tools have been made publicly available.

- **Annotated content for training and testing new methods.** A significant amount of content was annotated for training and testing purposes. Discussions with content owners were initiated, in order to make as much of this content as possible publicly available as benchmark corpus.

- **Wide dissemination of results in the research community.** Both intermediate and the final results of the project were disseminated widely. 3 workshops were organized by BOEMIE, in each year of the project. Furthermore, more than 30 scientific articles were presented in international conferences and journals and many more are either under review or are being prepared. A state-of-the-art book is being completed and will be published by Springer. Additionally, the project was showcased through invited talks and demos in major conferences and exhibitions. Finally, two showcase videos were produced and are available through the Web site of the project.

- **Establishment of contacts with interested users of the BOEMIE technology.** BOEMIE has produced commercially interesting technology. Among others this was confirmed in the meetings that we had with some of the potential users during the project. Some of them have expressed particular interest in collaborating with us to bring part of the technology to the market.

Results

BOEMIE has produced more than 100 different modules and components, in order to achieve its ambiguous research goals. Most of these have been incorporated in the integrated prototype that was delivered and evaluated at the end of the project. The BOEMIE integrated prototype implements the bootstrapping process, as illustrated in Figure 1. Figure 2 presents the main components of the prototype.

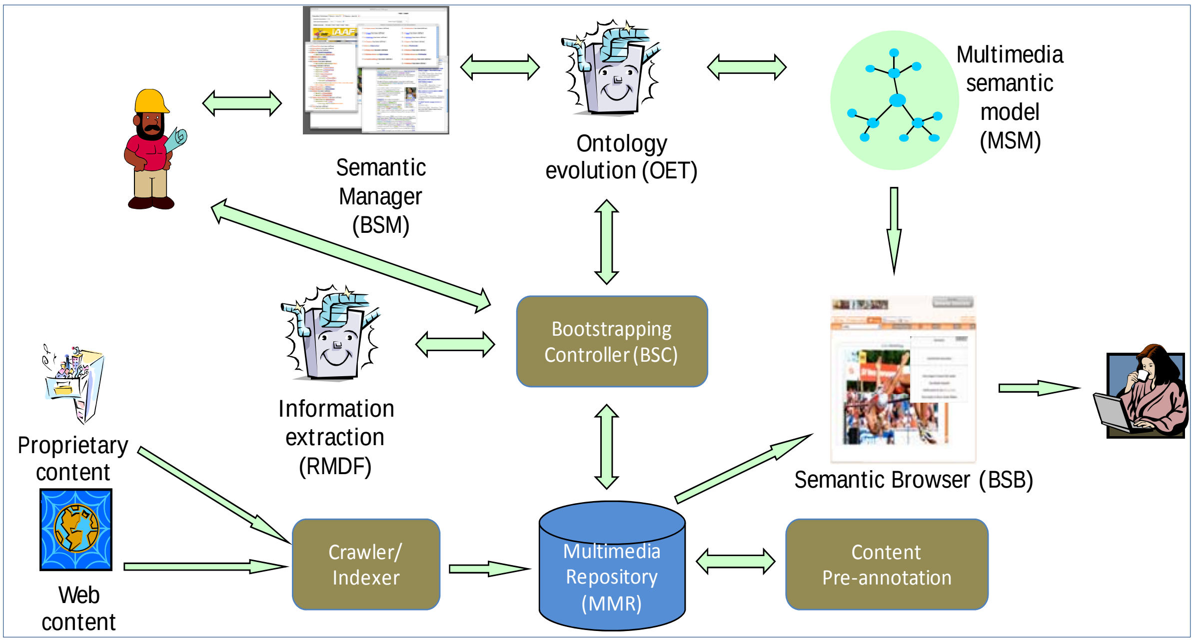


Figure 2. The BOEMIE integrated prototype.

Many of the modules and components that have been developed are also of great value, either as stand-alone tools or as subpackages of the integrated prototype. In the context of the exploitation plan for the project, we identified more than 40 such packages of commercial and social value. In this section of the report, we will present the main components of the BOEMIE prototype, mentioning in brief the tools that are involved in each component.

The extraction toolkit: Recursive media decomposition and fusion

Objective

To jointly process multimedia data in order to extract semantics in greater quantity and/or with higher accuracy than if independently processing the constituent single media data.

Summary of result

The extraction toolkit integrates a number of tools for content analysis and interpretation, using a recursive media decomposition and fusion framework. In the course of the project, innovative methods for the analysis of single-modality content were developed, going in most cases beyond the state-of-the-art. As Figure 3 illustrates, these methods cover most of the currently available types of media and are integrated, using a Web service architecture. Most importantly, they support the bootstrapping process through an evolving cycle of *analysis* of new content, *learning* of improved analysis models and *discovering* interesting objects and entities to add to the domain knowledge.

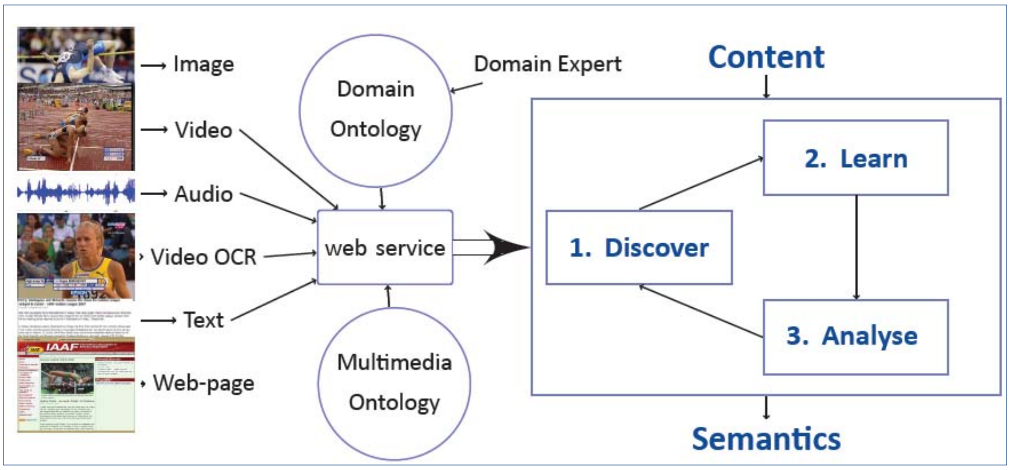


Figure 3. Information Extraction Toolkit.

The coordination of the evolving extraction process is achieved by a new method that was developed in BOEMIE and is called *recursive media decomposition and fusion* (RMDF). The method uses a number of techniques to decompose a multimedia document into its constituent parts, including embedded text in images and speech. It then relies on a synergy of the single-modality modules, each of which produces its results in a common format. The medium-specific analysis results are fused together in a common graph that complies with the domain ontology. In a final step, graph techniques are applied to provide a consistent overall analysis of the multimedia document. Figure 4 illustrates the recursive decomposition process for a Web page, containing a video, containing a static overlaid image, that embeds text, referring to a person.

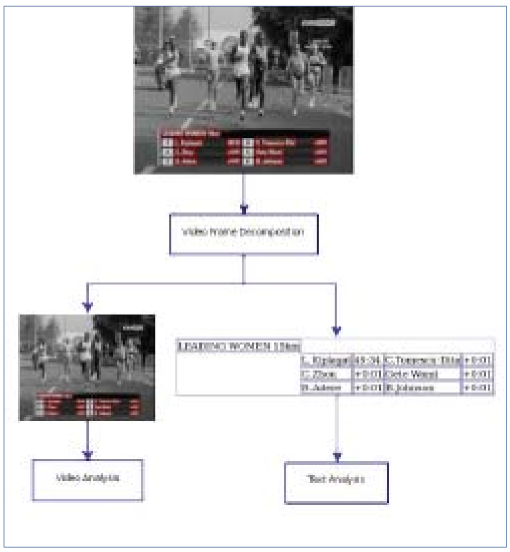


Figure 4. Recursive media decomposition.

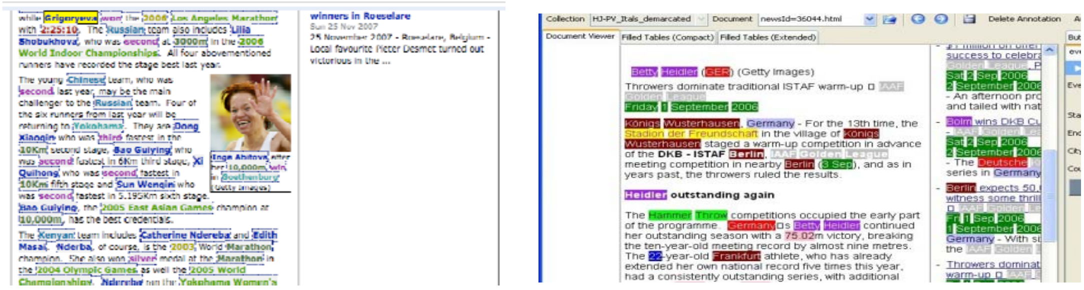
Regarding the single-modality modules, BOEMIE has developed innovative methods to:

- detect and discover objects of various shapes and sizes in images,
- track moving objects in video and classify movement phases,
- identify and discover entities in text and relations amongst them,
- detect overlay and scene text in video and perform optical character recognition on it,
- recognise and discover audio events and interesting keywords in audio.

(a) Object detection in images.

(b) Phase detection in video.

(c) Text recognition in video.



(d) Entity and relation extraction in Web pages.

Figure 5. Sample results of analysis tools.

Figure 5 provides examples of such results. Most importantly, however, through interpretation and fusion the RMDF is able to improve significantly the precision of multimedia analysis, be it in Web pages containing HTML text and images or video footage with audio commentary and overlay text.

Advantages over the state of the art

The BOEMIE approach to information extraction from multimedia content has managed to achieve to a great extent the goals of high accuracy and scalability that we had set at the beginning. It did so, by introducing a number of innovations to the state of the art:

- It can deal with information expressed as a combination of practically any media, with arbitrary structure. Information structure is described by means of suitable ontologies.
- It can handle with unrestricted levels of multimedia information embedding.
- It supports customization to any domain, by allowing the discovery of new semantics in content and learning to identify known objects and entities.
- It is easily distributable and scalable, by dynamically integrating per media analysis techniques in an unrestricted number of servers, communicating through a computer network.
- It uses uniform open interfaces to allow for easy integration of new components.

Abductive multimedia interpretation

Objective

To achieve high-level semantic interpretation of multimedia content, based on the extracted semantics and our knowledge of the domain, represented in the form of ontologies.

Summary of result

The interpretation of multimedia content in BOEMIE goes well beyond the usual extraction of semantics from individual media. Domain knowledge, in the form of ontologies, is being exploited by a reasoning-based interpretation service that operates in two levels: single-media interpretation and fusion. Figure 6 illustrates this multi-level analysis and interpretation process. Both the single-media and the fusion services are supported by the same reasoning apparatus.

Reasoning for multimedia interpretation is based on the RacerPro reasoning engine, which has been extended with many novel methods for the purposes of BOEMIE. One of the main extensions is the use of abduction to generate interpretation hypotheses for what has been “observed” by the extraction tools. The new abductive query answering service of RacerPro is able, during query evaluation, to recognize non-entailed query atoms and hypothesise them. Since there might be more than one hypothesis (i.e. explanations), a set of scoring functions has been designed and implemented in order to prefer certain hypotheses over others. Given the complexity of the interpretation hypothesis (a.k.a. explanation) space, important optimizations have been developed in the reasoner, in order to cut down on the number of consistent and useful interpretations that are produced by the system.

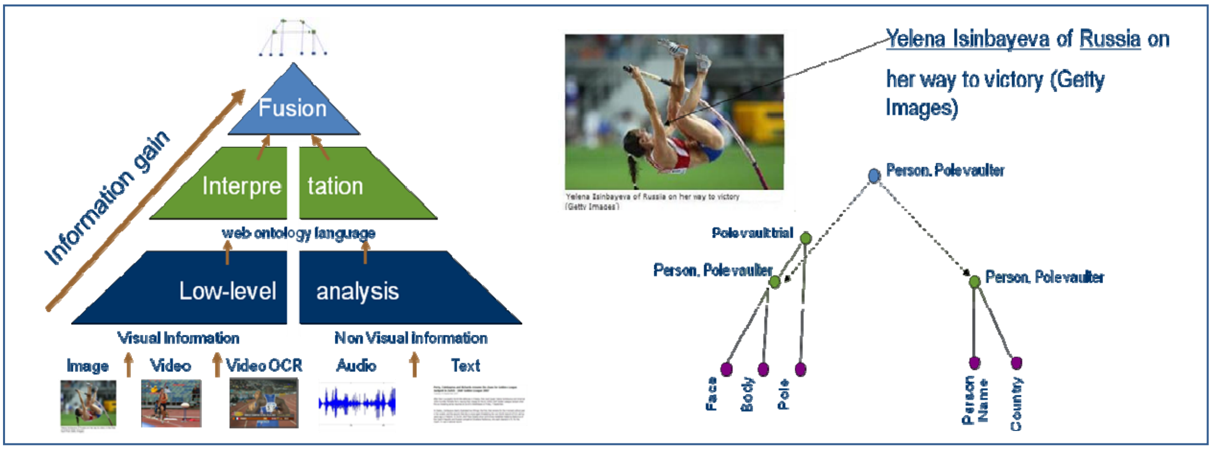


Figure 6. Multimedia analysis and interpretation process.

Abductive multimedia interpretation combines Description Logics, as a representation formalism for ontologies, with DL-safe rules that guide the search for interpretations. In the context of BOEMIE, we have developed methods to learn these rules, as well as to learn an efficient statistical interpreter that simulates the behaviour of the reasoner.

Advantages over the state of the art

The state-of-the-art in high-level multimedia interpretation is very limited. The BOEMIE interpretation service provides among others:

- True reasoning over the extracted information, producing implicit knowledge via the observed relations among extracted semantics.
- A novel abductive approach to producing and ranking potential interpretations.
- Integration of the new methods in a well-known reasoner (RacerPro).
- A Web service interface to the interpretation machinery.

Multimedia Semantic Model

Objective

To provide a unifying representation for thematic ontologies with multimedia content and descriptor ontologies, in order to support multimedia interpretation and ontology evolution.

Summary of result

The BOEMIE Multimedia Semantic Model (MSM) integrates ontologies that capture our knowledge about a particular domain, e.g. athletics, with ontologies that model knowledge about the structure and low-level descriptors pertaining to multimedia documents (Figure 7). Besides addressing the interlinking of a multimedia document segments with the corresponding domain entities, MSM further enhances the engineering of subject matter descriptions by distinguishing between mid-level (MLC) and high-level (HLC) domain concepts and properties, a feature unique to the BOEMIE project. Instances of MLCs represent information that is directly extracted from the multimedia content, using the various analysis tools, e.g. the name of an athlete or her body in a picture. On the other hand, instances of HLCs are generated through reasoning-based interpretation of the multimedia content, using the domain ontology. Such engineering allows incorporating the analysis perspective into the domain conceptualisation, which in turn supports effective logic-grounded interpretation. The developed ontologies allow the utilisation of precise formal semantics throughout the chain of tasks involved in the acquisition and deployment of multimedia content semantics.

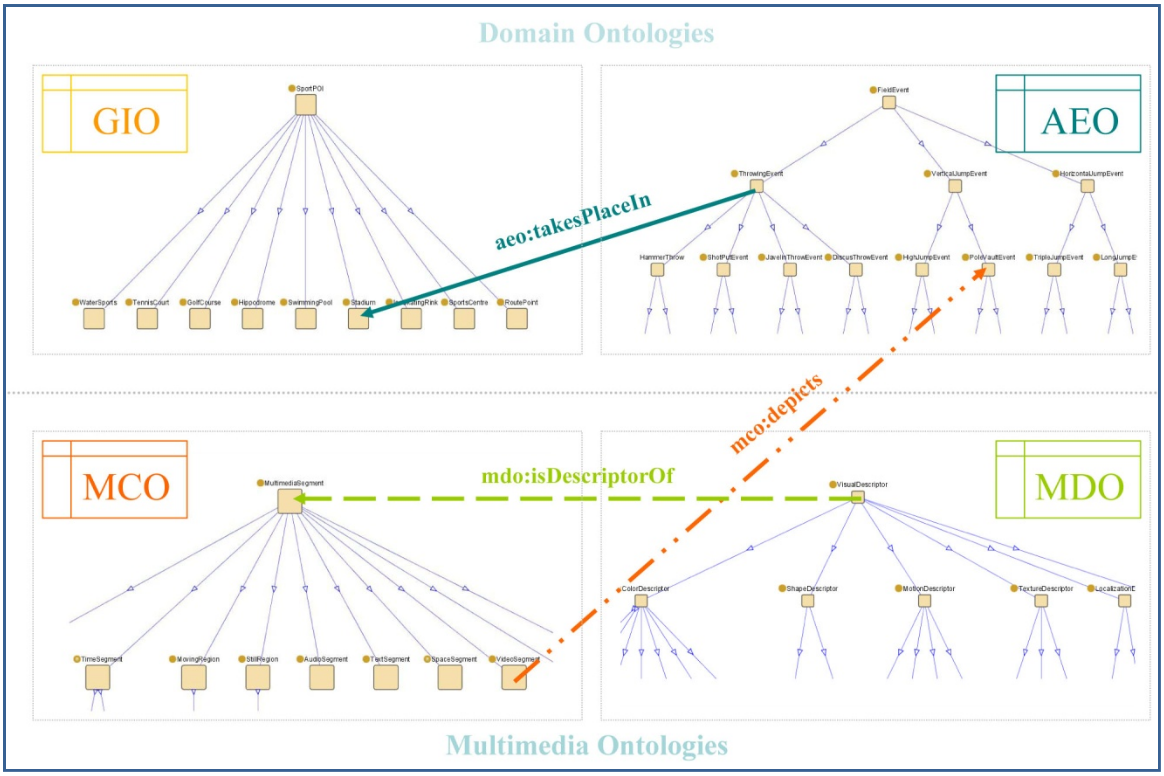


Figure 7. The multimedia semantic model.

As illustrated in Figure 7, the semantic model hosts four different ontologies:

- The Athletics Event Ontology (AEO), which models the BOEMIE scenario domain of interest, i.e. public athletics events. This ontology corresponds to the bulk of the semantic model and is the first comprehensive ontology for athletics, in accordance to the IAAF regulations. AEO provides the domain definitions on which content interpretation is based.
- The Geographic Information Ontology (GIO), which models information relevant to geopolitical, geometry and position attributes of geographic objects. This ontology enables the georeference of objects and events and supports multimedia enriched digital maps.
- The Multimedia Content Ontology (MCO), which models content structure descriptions, by re-engineering the respective MPEG-7 MDS definitions to alleviate semantic ambiguities and scalability of descriptions. MCO provides the means to link the extracted domain specific semantics to the respective content parts for enhanced retrieval and presentation, while providing structural and modality information in content interpretation.
- The Multimedia Descriptor Ontology (MDO), which models the MPEG-7 visual and audio part descriptors. This ontology supports the exchange and reuse of low-level information between analysis modules, capturing restrictions and valid descriptor structure.

The four OWL DL ontologies are linked in a way that supports the purposes of BOEMIE for semantics extraction, interpretation, evolution, as well as retrieval and representation of the acquired semantics. Figure 8 presents a simple example of this interlinking between the ontologies.

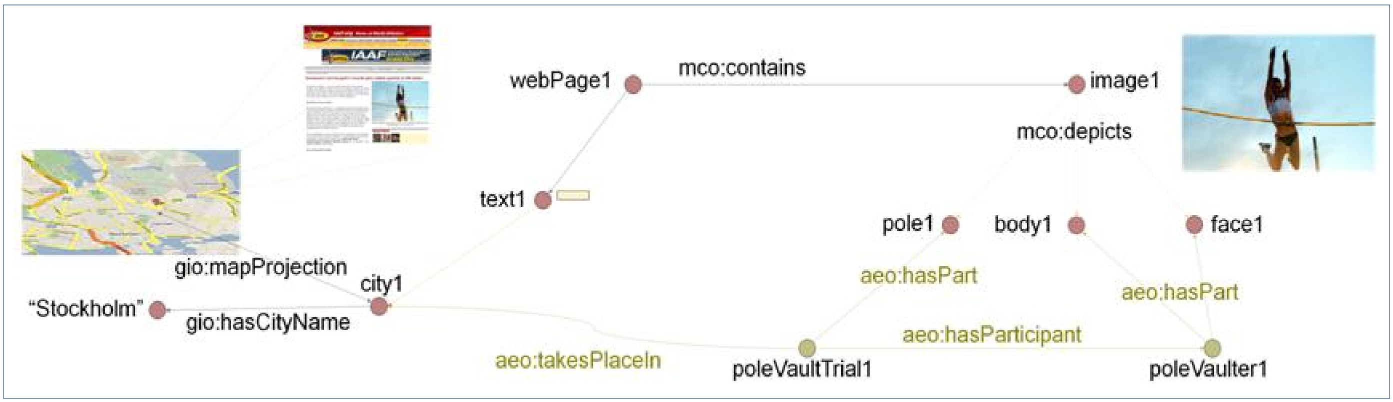


Figure 8. Linking knowledge between ontologies in the semantic model.

Advantages over the state of the art

The main novelty of the BOEMIE Multimedia Semantic Model (MSM) is the interlinking of carefully developed ontologies that provide the required domain knowledge, while at the same time supporting the linking of extracted information with high-level semantics, geographic information and content structure. More specifically the MSM has the following advantages:

- It follows principles of modularised and extensible ontology engineering.
- It provides clean modelling and sound semantics of the knowledge.
- It supports seamless communication and sharing of analysis / interpretation information.
- It provides comprehensive yet compact modeling to keep reasoning efficient.
- It incorporates expressive semantics for the representation of multimedia feature information.
- It supports the handling of fuzzy information in the extraction results.
- It provides the first comprehensive ontology of athletics, following IAAF regulations.

Ontology Evolution Toolkit

Objective

To provide methods for semi-automated population and enrichment of the domain ontologies, based on the results of the extraction and discovery from multimedia content.

Summary of result

The ontology evolution toolkit (OET) implements a pattern-based approach to the population and enrichment of the ontology, which is unique to BOEMIE. In particular, two different cases have been identified for the population process, one in which a single interpretation is produced for a document and one in which more than one candidate interpretations are provided. Furthermore, two cases are defined for the enrichment process, one in which a HLC and one in which a MLC is added. Each of those cases requires different handling in terms of the interaction with the domain expert and the modules that are employed for the semi-automated generation of new knowledge, e.g. concept enhancement, generation of relations and interpretation rules, etc. Figure 9 provides a high-level overview of these four cases (patterns P1 to P4) and the modules that are involved.

The simplest of the four cases occurs when a single interpretation is generated (P1). In that case, OET is responsible for matching the named entities that take part in the interpretation with the names of entities that have been identified in the past and populate already the ontology. Once instance matching and grouping has taken place, the new instances can be associated correctly with concepts of the ontology. Instance matching and grouping are even more important in P2, where multiple interpretations have been generated and one of them needs to be selected as the most appropriate. This is based on statistical and semantic similarity of the instances involved in the interpretation with instances already in the ontology. Novel methods have been developed for the purposes of instance matching and grouping, in order to take advantage of the rich semantics of the BOEMIE semantic model and scale efficiently to large document sets. These methods have been incorporated in the HMatch ontology matching software, which is publicly available.

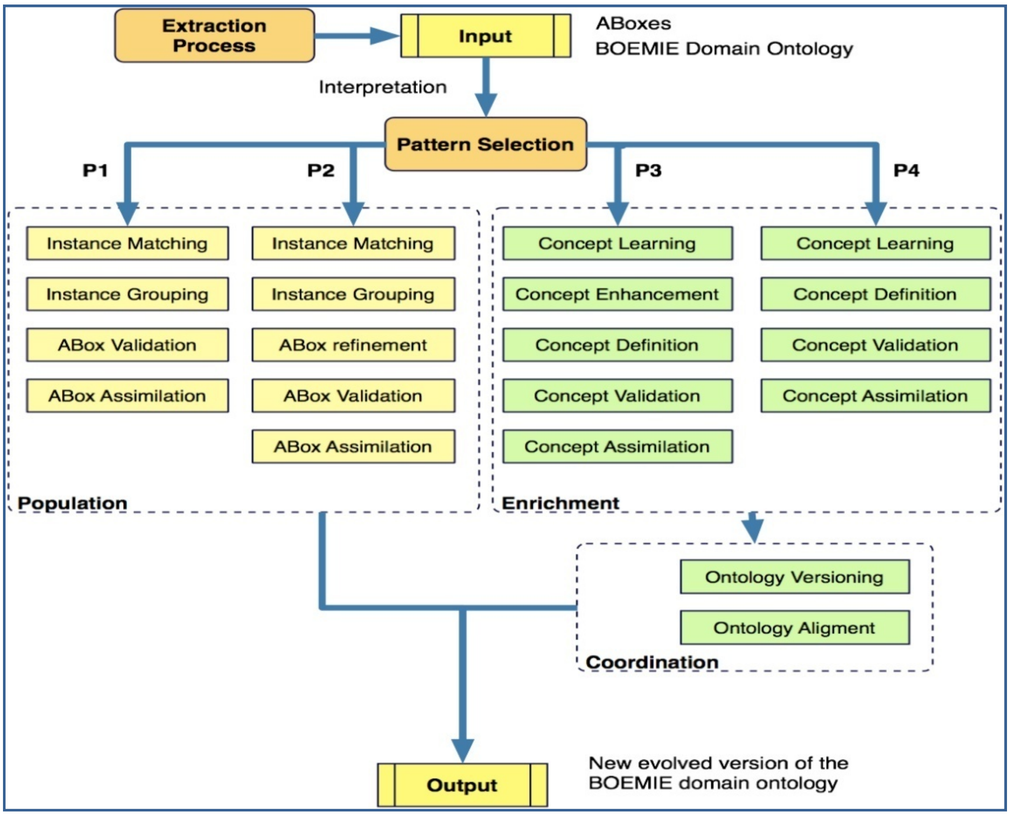


Figure 9. Pattern-based ontology evolution.

A number of innovations have been made also in the area of ontology enrichment. The discovery of new high-level concepts and properties is based on a new methodology that incorporates a set of ontology modification operators, e.g. add/remove property, generate aggregate concept, specialize a concept, etc. The applicability of the operators and the statistical significance of the resulting modifications is assessed logically and statistically and the most promising changes are kept. Further to this data-driven enrichment, a concept enhancement methods is incorporated, which matches new constructs to knowledge in external resources, e.g. on the Web, in order to provide a richer description for a new concept. The discovery of new mid-level concepts is even more based on the data and the discovery methods implemented in the extraction tools. Since mid-level concepts are primitive, matching with existing knowledge, internal or external to the BOEMIE semantic model, is not suitable.

Advantages over the state of the art

The BOEMIE ontology evolution toolkit provides a very comprehensive set of evolution functionalities for semantically rich ontologies that are to be used for multimedia interpretation. In order to achieve that, it contributes to the state of the art in a number of different ways:

- A pattern-based approach that triggers automatically different types of population and enrichment, based on the results of multimedia interpretation.
- Efficient instance and concept matching, using statistical evidence and semantics.
- A flexible enrichment process, based on a set of modification operators and evaluation measures, covering a variety of changes to the ontology.
- Extensive use of automated learning and reasoning, in order to minimise and simplify the involvement of the user in the population and the enrichment process.

BOEMIE Semantic Manager

Objective

To assist domain experts and content managers in managing domain knowledge (ontologies) and semantically annotated multimedia content.

Summary of result

The BOEMIE Semantic Manager (BSM) is a Web-based application that provides user-friendly access to the semantically annotated content and allows the enhancement of the ontologies, through the recommendations of the ontology evolution toolkit (OET). Particular emphasis has been given to make the BSM available to non-experts in knowledge technologies. As an interface to the OET, the BSM provides three primary functionalities:

- Population of the ontology with semantically annotated multimedia content.
- Enrichment of the ontology with new knowledge that has been learned from data.
- User-friendly interactive enhancement of new knowledge by the domain expert.

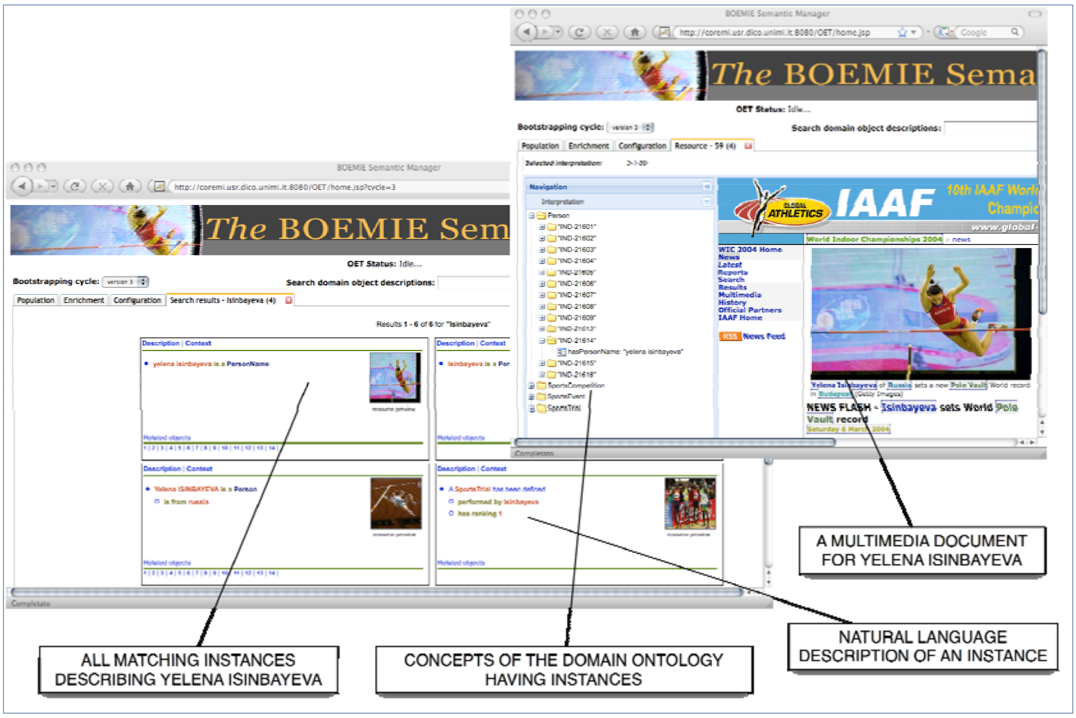


Figure 10. Ontology population, using the semantic manager.

Figure 10 provides a view of the ontology population functionality of BSM, where new concept instances are added to the ontology. The BSM provides interactive selection/approval/rejection of the multimedia content interpretations automatically produced by the BOEMIE system, as well as (similarity-based) document browsing facilities. The BSM processes concept instances in order to create a natural language description of the underlying logic representation of the extracted knowledge. This description is more comprehensible to the non-skilled in knowledge engineering.

Figure 11 illustrates the process of using the recommendations of the OET to add new knowledge to the domain ontology. Pattern structures are reported in the form of proposals for insertion to or modification of the ontology. The BSM shows both terminological and structural suggestions to support the domain expert in performing ontology enrichment. Suggestions are dynamically extracted from knowledge chunks similar to a given concept proposal by exploiting ontology matching techniques. A repository of knowledge chunks is created and maintained through a knowledge harvesting process that periodically searches knowledge of interest from other ontologies, web directories, and, in general, external knowledge repositories. As illustrated in Figure 12, a simple ontology editor has been incorporated in the BSM, which uses natural language patterns and autocompletion techniques to facilitate the incorporation of new knowledge to the domain ontology. Through simple dragging and dropping, the user can add and modify knowledge chunks that are either learned from data by the OET or have been harvested by external knowledge repositories.

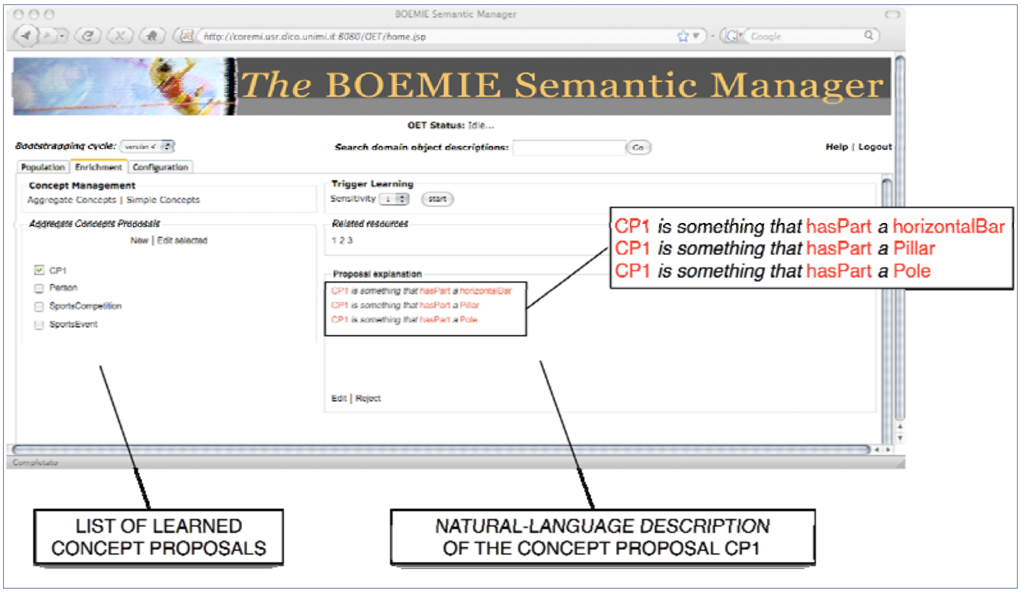


Figure 11. Ontology enrichment process, using the semantic manager.

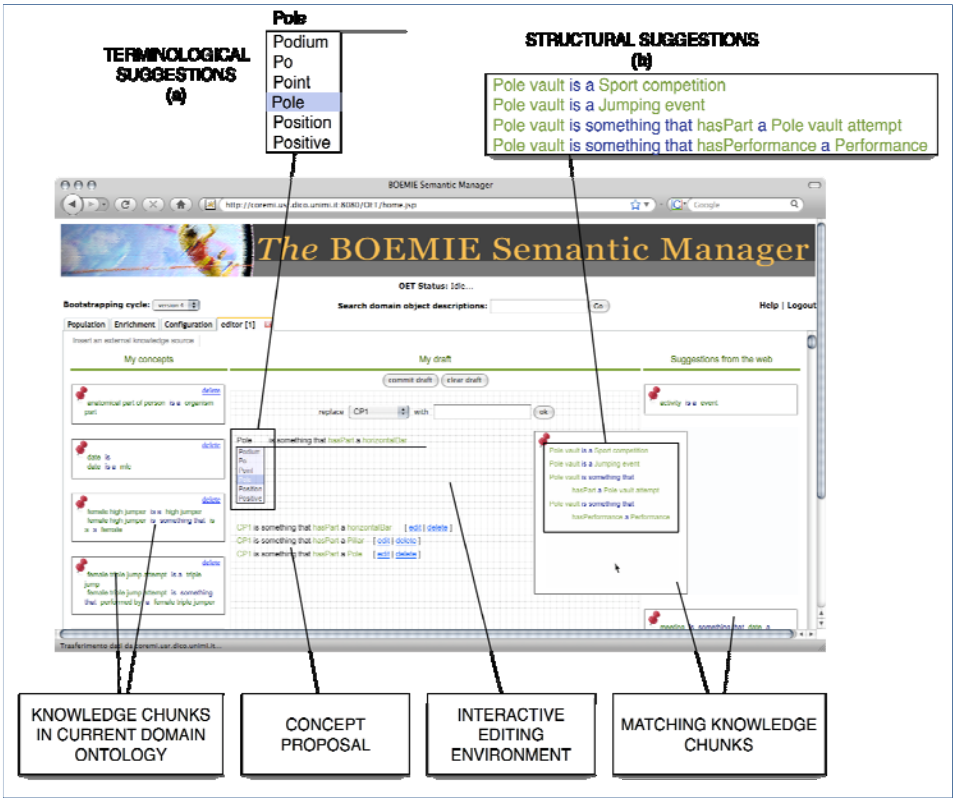


Figure 12. Concept definition and enhancement, using the semantic manager.

Advantages over the state of the art

The BOEMIE Semantic Manager is unique in its simplification of a complex and demanding process, i.e., that of adding semantics to multimedia content and maintaining the associated domain knowledge. In order to achieve that, it contributes to the state of the art in a number of different ways:

- Natural language translation of the logical expressions in the interpretation of multimedia content. This process involves also a significant and difficult task of data reduction, in order to focus to the information that is most important for the domain expert to make decisions.
- Intelligent coordination of knowledge coming from disparate external and internal sources, such as data clustering results and information from Web directories. The common ground for all this is the domain knowledge and its desired modifications.
- Simple natural language ontology editor with a user-friendly interface and extensive automation of the knowledge editing process.

BOEMIE Semantic Browser

Objective

To demonstrate the added value for the simple user of interactive and geo-referenced multimedia content, facilitated by the semantic analysis and annotation technologies.

Summary of result

The BOEMIE Semantic Browser (BSB) is the main end-user application of BOEMIE, which demonstrates the usefulness of annotated multimedia content for the construction of dynamic interfaces to support the retrieval of the content. It does so by providing three models of interaction with the multimedia content:

- Interactive maps for multimedia retrieval.
- Interactive content of media objects.
- Dynamic suggestion of related information.

Regarding the interactive maps scenario, BSB demonstrates the feasibility of using the BOEMIE technologies to support geographically aware information retrieval for professional settings. Typically, such a setting is confronted by large information providers, e.g. owners of large amounts of data, such as news agencies. Such providers wish to provide geographically aware information retrieval, while avoiding manual annotation costs and keeping their rights over the multimedia content. Based on the information extraction technologies of BOEMIE, which are capable of recognizing explicit references to geographic information, the BSB can associate multimedia content with geopolitical areas and specific Points of Interest (PoIs) on digital maps. In this manner it provides direct access to the multimedia, through what we call “BOEMIE PoIs”. Figure 13 illustrates this idea.

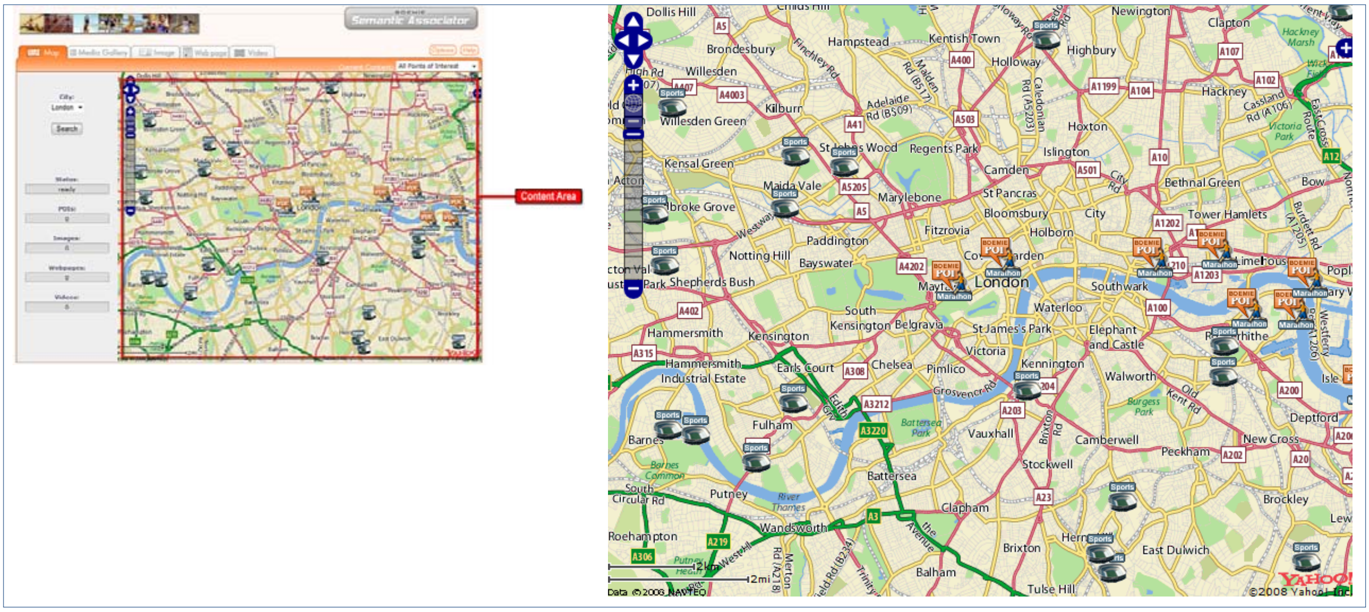


Figure 13. Navigating to multimedia content using digital maps.

Further to the map-based navigation, BSB uses the semantic annotations generated automatically by BOEMIE to make media objects interactive. More specifically, it automatically highlights relevant content of a specific domain on top of text or images to prepare the interface for further interaction possibilities. Figure 14 shows the kind of information that is highlighted and made active for further navigation through the multimedia content.

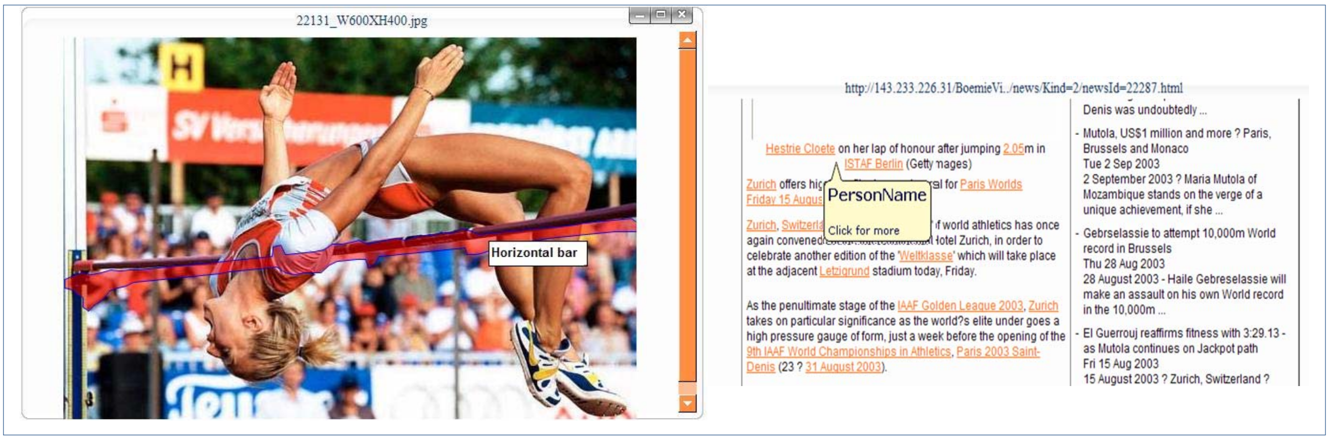


Figure 14. Making media objects active in the semantic browser.

The highlighted information illustrates only the surface-level capabilities of BOEMIE information extraction technologies. However, through interpretation, supported by reasoning on domain ontologies, BOEMIE is able to generated deeper semantic information, e.g. the type of sport that an image depicts. Using this implicit knowledge, BSB provides context-sensitive advertisement and suggests related information. This is realized by the idea of context menus illustrated in Figure 15.

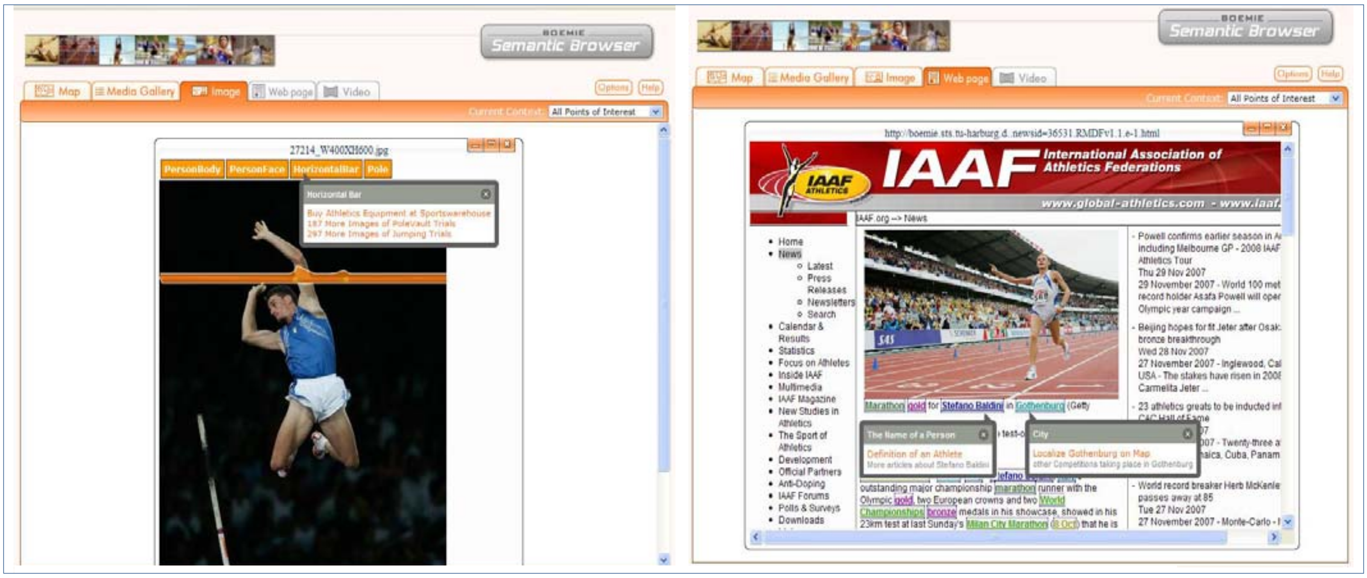


Figure 15. Suggesting information related to active media objects.

Finally, in order to make all this advanced interaction capabilities real-time several tools have been developed in BOEMIE to pre-compute information for BSB. One of this tools that is integrated at the server side of BSB is called DYNAMITE (DYNAMIc InTEractive web pages). DYNAMITE annotates the content of HTML files to augment it with deep-level semantics resulting from the interpretation. Other complementary tools to DYNAMITE are provided with the Bootstrapping Controller of BOEMIE, described in the following section.

Advantages over the state of the art

The BOEMIE Semantic Browser provides an innovative interaction experience with multimedia content. It showcases the value added to multimedia content, through semantic annotation and facilitates the developer of related Web applications in several ways. More specifically:

- It facilitates developers to create interfaces where users interact with the identified objects.
- It facilitates developers to seamlessly integrate standard interaction technologies (e.g., interaction with maps, context menus, etc.) with media objects on a semantic basis.
- It illustrated how intelligent interfaces can be created that identify the set of commands that will drive their functionality according to the semantic context.
- It allows developers to build such intelligent interfaces with much less effort than with standard development techniques known today.

BOEMIE Bootstrapping Controller

Objective

To support the addition of new content and implement the bootstrapping of the analysis and evolution processes.

Summary of result

The Bootstrapping Controller (BSC) is the main application logic component that implements the iterative extraction and evolution process. In order to achieve that it acts as a controller of the information extraction toolkit (RMDF) and the ontology evolution toolkit, introducing also a number of other components that are necessary: (a) a Web crawler and content indexer, (b) the BOEMIE multimedia content repository (MMR), (c) content preparation tools, in order to provide optimised viewing of semantically annotated content for the user applications of BOEMIE, i.e. the semantic manager (BSM) and the semantic browser (BSB).

Using the Web user interface of the BSC (Figure 16), the content owner can add content to the Multimedia Repository (MMR) and then send it for processing through predefined workflows. The content is added to MMR, either by uploading specific files or by crawling the Web, using the BOEMIE wrapper over the HERITRIX (http://crawler.archive.org) crawler. Once uploaded or downloaded, the content is indexed according to the needs of BOEMIE. Typically, a new document

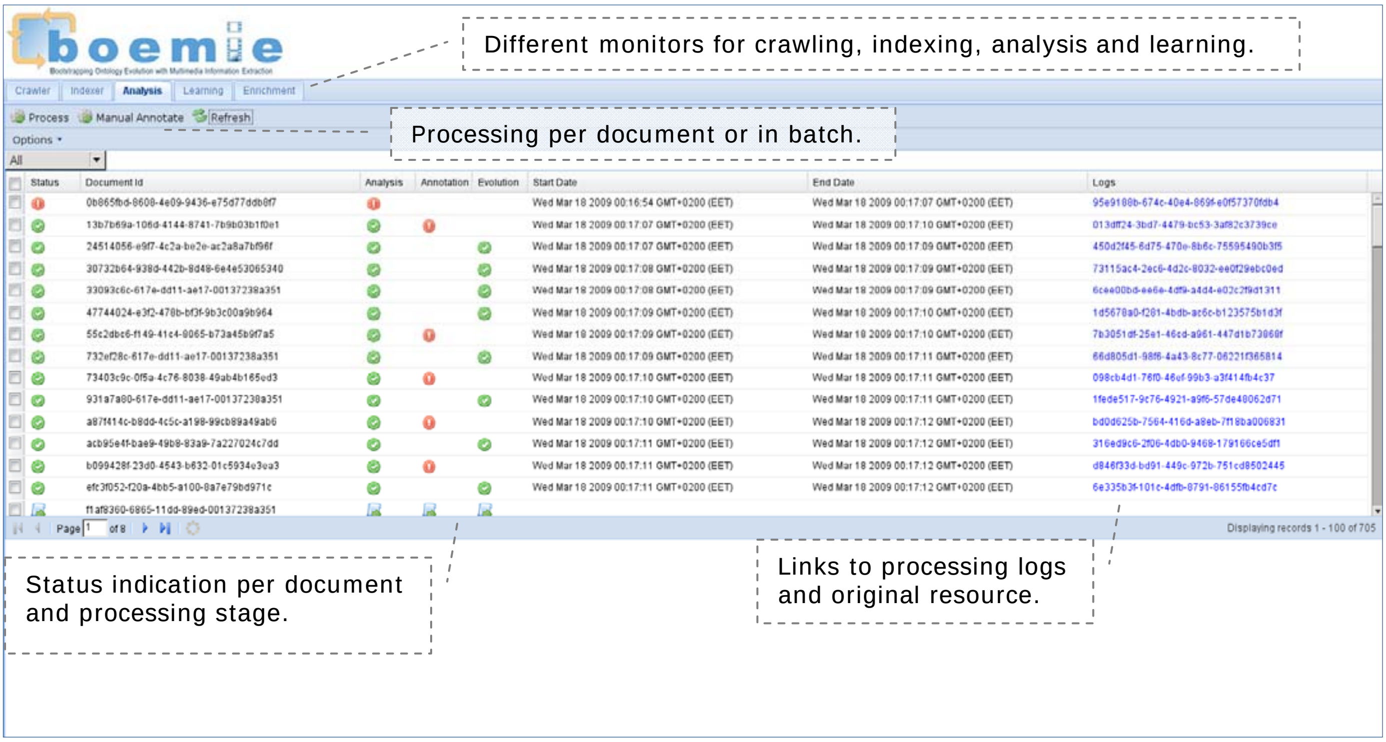


Figure 16. Bootstrapping controller's Web interface.

will be sent to RMDF for extraction and the results of its interpretation will be populated into the ontology, using the OET. When sufficient evidence is accumulated, the OET will generate proposals for changes to the ontology, i.e. new HLCs. The domain expert will use these recommendations to change the ontology and the content will be sent again for processing by the RMDF. In some cases, new mid-level concepts (MLCs) will be generated based on the analysis of the multimedia content so far. In these cases, in addition to the extension of the ontologies, the BSC will send sufficient training data to the RMDF, asking for the re-training of the analysis modules.

Regarding its support for the Web applications of BOEMIE, the BSC integrates a number of modules that visualize the semantic annotations to be shown to the user. By pre-computing this information, the visualization of multimedia content is made faster. Such pre-annotation tools have been constructed for text and HTML (TAHAT), as well as for images (IMS). The results of the pre-annotation are stored separately in the MMR repository.

Advantages over the state of the art

The main innovation of the BSC is in the implementation of the bootstrapping logic, including the various process workflows that are needed. Furthermore, BSC implements a number of project-specific processing components.

Manual Annotation Tools & Manually Annotated Corpus

Objective

To create the necessary training material for the information extraction methods of BOEMIE and provide publicly available tools and valuable benchmark corpora to the research community.

Summary of result

The information extraction methods developed in BOEMIE are trainable and therefore require training material, in order to learn to identify interesting entities, objects and relations among them in multimedia content. The BOEMIE bootstrapping process generates semi-automatically such training data. However, for the purposes of training and evaluating the initial extractors, we generated significant quantities of training data for all types of media: image, video, audio, text. For these data we used interactive tools for manual annotation. Most of these tools were also developed in BOEMIE and improve significantly the state of the art in the field. Based on previous experience of the partners in developing manual annotation tools and the challenging needs of BOEMIE, we developed user-friendly applications for the manual annotation of image and video (VIA), as well as text and HTML (BTAT). These tools are distributed publicly, through the BOEMIE Web site.

The VIA tool can be used for high-level and low-level video and image annotation. In both cases, annotation is aligned with concepts of the domain ontologies. In the case of image annotation, either image regions and complete images are linked with concepts (high-level annotation) or visual descriptors are extracted per annotated region and associated with the corresponding concept (low-level annotation). To reduce the manual annotation burden, VIA supports the automatic segmentation of a still image into regions and region-merging. In case the user is not satisfied with the outcome of automatic segmentation, new regions can be defined by free-hand, shape-based or polygon approximation drawing tools. Regarding video annotation, VIA supports input in MPEG1/2 video format and frame accurate video playback and navigation. Video annotation can take place either in a frame-by-frame style or as live annotation during playback. Furthermore, it supports annotation at all granularity levels (i.e. global, shot, frame and region). Figure 17 illustrates the use of VIA.

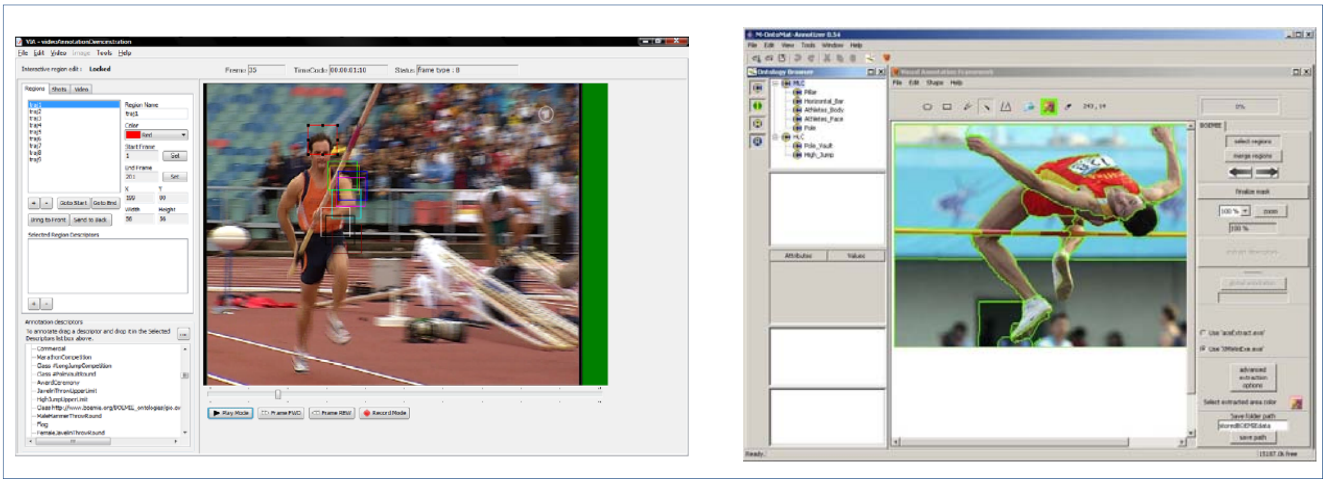


Figure 17. Video and image annotation with the VIA tool.

The BTAT tool has been developed over the Ellogon open-source text engineering platform (http://www.ellogon.org/), but runs as a stand-alone application. BTAT supports the annotation of named entities, the so-called Middle level concepts (MLCs), as well as relations between those named entities. The relations are grouped in tables of specific types. Tables correspond to High Level Concepts (HLCs) having their fields filled with MLC instances. Furthermore, the tool enables the annotation of relations between HLC instances by creating linkages between tables in an effective and easy way. An additional property of the tool is that it is able to annotate blocks of texts in a web page that refer to the same topic defined by the presence of predefined types of MLC instances. One of the innovations in BTAT is its dual manual and automated annotation functionality. Manual annotation is facilitated by a smart text-marking system, where the user selects with a mouse click words, instead of single characters. Automatic annotation works by matching user-defined regular expressions. Figure 18 illustrates the use of BTAT.

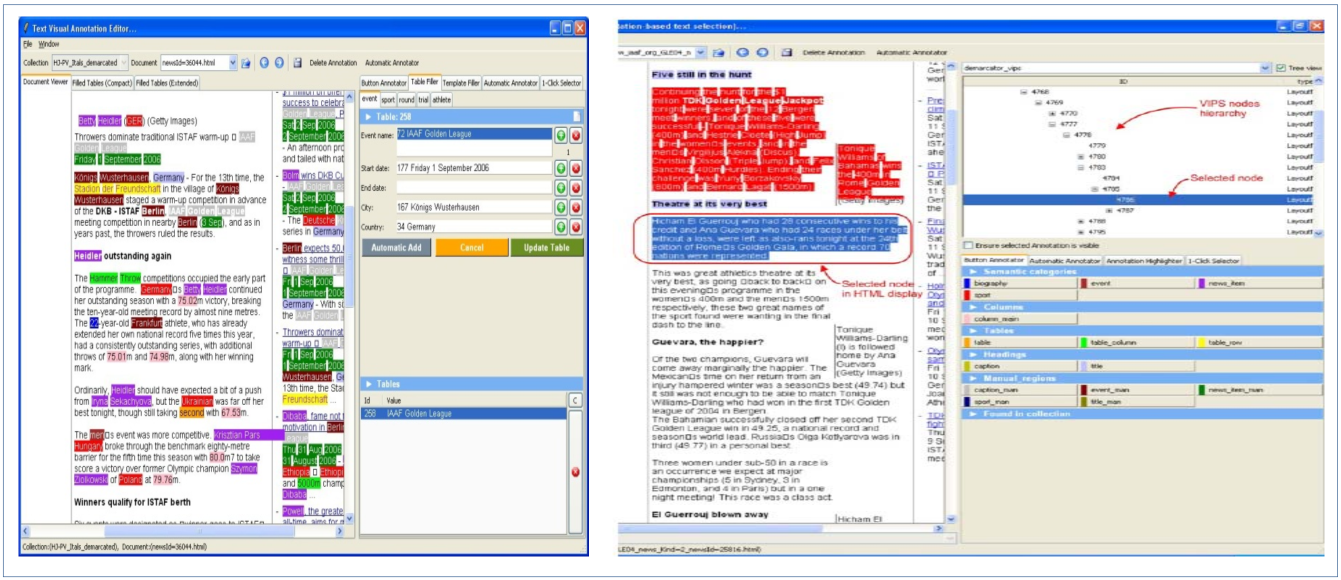


Figure 18. Text and HTML annotation with the BTAT tool.

Using the manual annotation tools, a large corpus of annotated documents have been produced in BOEMIE. These documents include thousands of HMTL pages, hundreds of images and several hours of video and audio footage. Efforts are made to make some of this content publicly available for research purposes, in collaboration with the content owners.

Advantages over the state of the art

The requirements of BOEMIE for deep semantic annotation of multimedia content have led to the introduction of many innovations in the manual annotation tools for video, image and text, going well beyond the current state of the art. Among these innovations are:

- The handling of segments in image annotation by VIA.
- The live annotation of video by VIA.
- The pattern-based semi-automated annotation of BTAT.
- The level at which semantic annotations are linked to concepts of the domain ontology.

By making these tools freely available, BOEMIE provides very valuable resources to the research community. Similarly useful will be the publishing of the annotated corpora as benchmarks for further research.

Scientific and commercial use of results

BOEMIE has made a number of its results already **publicly available** through its Web site: http://www.boemie.org/

Regarding software tools the following are available:

- BTAT (BOEMIE Text Annotation Tool): The tool used for manual annotation of text and HTML pages. Provided under an open-source licence.
- VIA (BOEMIE Video and Image Annotation Tool): The tool used for manual video and image annotation. Provided freely in the form of an executable.
- VAT (BOEMIE Video Annotation Tool): A specialized version of VIA. Provided freely in the form of an executable.
- Hmatch (Ontology Matching Tool). A suite of tools for ontology matching. Provided as a Protege plugin and as a stand-alone application, under an open-source licence.
- RacerPro (DL Reasoner). A general-purpose DL reasoned, incorporating the multimedia interpretation technology of BOEMIE. Provided under special research licence.

More of the tools developed in BOEMIE are expected to be packaged and made available in the future.

Further to software tools, BOEMIE has made publicly available the ontologies that have been developed in the project, including:

- AEO (BOEMIE Athletics Ontology)
- MCO (BOEMIE Multimedia Content Ontology)
- MDO (BOEMIE Multimedia Descriptor Ontology)
- GIO (BOEMIE Geographic Information Ontology)

Additionally, for the purposes of acquaintance with the project and in particular the Web applications that have been developed, a number of videos and related presentations are available at the Web site:

- A short showcase video, targeting the general public.
- A longer showcase video, aiming in particular at the domain experts, e.g. content providers and news agencies.
- A demo presentation and video for the BOEMIE Semantic Browser.
- A demo presentation for the BOEMIE Semantic Manager.

- A demo presentation for the Bootstrapping Controller.

Furthermore, a long list of publications and public deliverables are available on the Web site.

Beyond the publicly available resources, however, the BOEMIE consortium has particular interest in the **commercial exploitation** of its results. Towards this end the BOEMIE management board has decided in detail the **property rights** of different partners on the technologies that have been produced and has drafted a plan for exploiting the most promising ones. These exploitation plans foresee the involvement of commercial third-parties who can add value to products and services based on BOEMIE technologies. Companies who are interested in collaborating with the BOEMIE partners on such commercial activities are welcome to contact the coordinator or other site managers, the contact details of which are provided in the following section.

Finally, access to the Web applications of BOEMIE for testing can be arranged also by contacting the project coordinator.

Partner contacts

NCSR "DEMOKRITOS" Institute of Informatics and Telecommunications http://www.iit.demokritos.gr/	Constantine D. Spyropoulos costass@iit.demokritos.gr
Fraunhofer-Gesellschaft (FhG) Institut für Intelligente Analyse- und Informationssysteme http://www.iais.fraunhofer.de/	Joachim Koehler joachim.kohler@iais.fraunhofer.de
Informatics and Telematics Institute (ITI) of CERTH http://www.iti.gr/	Yannis Kompatsiaris ikom@iti.gr
Techniche Universität Hamburg - Harburg (TUHH) Institut für Softwaresysteme http://www.sts.tu-harburg.de/	Ralf Möller r.f.moeller@tuhh.de
Università degli Studi di Milano (UniMi) Dipartimento di Informatica e Comunicazione http://www.dico.unimi.it/	Silvana Castano castano@dico.unimi.it
TELE ATLAS N.V. http://www.teleatlas.com/	Linde Vande Velde Linde.VandeVelde@teleatlas.com